Deutscher Ethikrat

Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung

STELLUNGNAHME

+++ Vorabfassung – wird durch die endgültige Fassung ersetzt. +++

Deutscher Ethikrat

Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung

Stellungnahme

30. November 2017

Deutscher Ethikrat
Geschäftsstelle
Jägerstraße 22/23
10117 Berlin

Telefon: +49 30 20370-242
Fax: +49 30 20370-252
E-Mail: kontakt@ethikrat.org

Danksagung

Der Deutsche Ethikrat dankt allen, die schriftlich oder mündlich Inspiration und Informationen zu dieser Stellungnahme beigetragen haben, herzlich für ihr Engagement.

Zum Projektauftakt lieferte die Jahrestagung „Die Vermessung des Menschen – Big Data und Gesundheit“ am 21. Mai 2015 mit Vorträgen und Podiumsbeiträgen der eingeladenen Referenten[1] Elisabeth André, Anke Domscheit-Berg, Arno Elmer, Nils Hoppe, Christof von Kalle, Peter Langkafel, Klaus Mainzer, Wolfgang Marquardt, Günther Oettinger, Thomas Petri, Frank Rieger, Florian Schumacher, Stefan Selke und Henry Völzke sowie zahlreichen Diskussionsbeiträgen der über 500 Tagungsteilnehmer eine exzellente Grundlage.

Gemeinsam mit Christoph Kucklick stand Stefan Selke dem Deutschen Ethikrat im Rahmen einer öffentlichen Sitzung am 23. März 2016 ein zweites Mal Rede und Antwort. Jan Philipp Albrecht, Urs-Vito Albrecht, Luciano Floridi, Thomas Hofmann, Sascha Lobo, Indra Spiecker genannt Döhmann und Bart de Witte brachten ihren Sachverstand im Rahmen interner Anhörungen ein.

Von Februar bis März 2017 nahmen zahlreiche interessierte Personen und Organisationen an einer öffentlichen Befragung zu Big Data und Gesundheit teil, deren Ergebnisse in die weiteren Beratungen Deutschen Ethikrates eingeflossen sind und begleitend zu dieser Stellungnahme separat veröffentlicht werden.[2]

In der mit der Erarbeitung der Stellungnahme befassten Arbeitsgruppe wirkten bis zum Ende ihrer Amtszeit (2016) auch die Ethikratsmitglieder Wolf-Michael Catenhusen, Thomas Heinemann, Anton Losinger, Ulrike Riedel, Eberhard Schockenhoff, Jochen Taupitz, Christiane Woopen und Silja Vöneky mit. Wertvolle Unterstützung als kooptierte AG-Mitglieder leisteten Sascha Lobo (2015) und Thomas Hofmann (ab 2016).

Besonderer Dank gebührt den Studierenden, die während der Projektlaufzeit im Rahmen von Praktika beim Deutschen Ethikrat mit ihrer intensiven Zuarbeit entscheidend zur Fertigstellung dieser Stellungnahme beigetragen haben.

[1] Aus Gründen der besseren Lesbarkeit wird auf eine geschlechterspezifische Differenzierung verzichtet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung für alle Geschlechter.
[2] Alle öffentlich verfügbaren begleitenden Informationen und Dokumentationen des Deutschen Ethikrates zum Thema sind unter http://www.ethikrat.org/themen/forschung-und-technik/big-data abrufbar.

Inhaltsverzeichnis

Zusammenfassung

Grundlagen: Big Data und Gesundheit

1) Big Data gehört zu den Schlüsselbegriffen der gegenwärtigen Debatte über die technologisch induzierte gesellschaftliche Veränderung. Das Stichwort beschreibt einen Umgang mit großen Datenmengen, der darauf abzielt, Muster zu erkennen und daraus neue Einsichten zu gewinnen. Dazu sind angesichts der Fülle und Vielfalt der Daten sowie der Geschwindigkeit, mit der sie erfasst, analysiert und neu verknüpft werden, innovative, kontinuierlich weiterentwickelte informationstechnologische Ansätze notwendig.

2) Die systematische Erhebung und Auswertung von Daten ist spätestens seit Beginn der Neuzeit ein bedeutender Faktor zivilisatorischer Entwicklung und schließt auch den Menschen und seine Lebensumgebung ein, zum Beispiel in der Biologie und Medizin, der Psychometrie, der Epidemiologie und den Sozialwissenschaften. Der Einsatz von modernen Computern, Speichertechnologien und schnellen Netzwerken erlaubt eine enorme Steigerung des handhabbaren Datenvolumens, aber auch vielfältige qualitative Verbesserungen, wie die Verwendung komplexerer Rechenvorschriften (Algorithmen) in rechenintensiven Computersimulationen und eine Rationalisierung, Standardisierung und Qualitätssteigerung vieler Arbeitsprozesse.

3) Mit der Entwicklung zu Big Data geht eine Transformation aller Phasen der Datenverarbeitung einher, die von zunehmender Automatisierung, Vernetzung und Durchdringung geprägt ist. Volumen und Tempo der voll automatisierten Datenerfassung sind in wenigen Jahren exponentiell gestiegen, und die rasche Verbreitung und Vernetzung von Geräten, die in allen Sphären der menschlichen Lebenswelt zur Datenerhebung genutzt werden können, eröffnet ständig neue Datenquellen.

4) Dies zeigt sich besonders anschaulich im Gesundheitsbereich. Dort nutzen immer mehr Forscher, Firmen und Ärzte Informationen, die aus der Verarbeitung riesiger Datenmengen entstanden sind. Zudem nimmt die individuelle Erfassung gesundheitsrelevanter Daten zu, zum Beispiel über die Apps von Mobiltelefonen und am Körper getragene Sensoren. Wenn solch vielfältige Daten verknüpft und analysiert werden, ermöglicht dies tiefe Einblicke in den aktuellen Gesundheitszustand, die Persönlichkeit sowie den Lebenswandel und erlaubt teilweise sogar Vorhersagen, etwa zur Krankheitsentwicklung.

5) Sind Daten einmal erhoben, sorgen Datennetzwerke und vernetzte Softwaresysteme mitunter in Echtzeit für ihren Austausch und ihre Verknüpfung, oft auch über Staatsgrenzen hinweg. Hierfür werden technische Standards für den Datenaustausch über

Schnittstellen zur Anwendungsprogrammierung entwickelt, die auch die Festlegung bestimmter Nutzungsregeln und die Nachverfolgung von Daten erleichtern.

6) Die effiziente Erfassung, Speicherung und Verarbeitung von Daten benötigt eine leistungsfähige Rechenmaschinerie. Sie wird meist in Datenzentren mit vielen vernetzten Servern bereitgestellt und vielfach von kommerziellen Anbietern offeriert. Die Verlagerung von lokalen Rechnern in die Virtualität solcher Datenzentren wird als Cloud-Computing bezeichnet.

7) Wesentlich für die Beurteilung von datenbasierten Aussagen, Schlussfolgerungen oder Vorhersagen sind die Objektivität, Reliabilität, Reproduzierbarkeit und Validität der verwendeten Daten bzw. Analyseverfahren. Mit der Menge der Daten steigen die Aussagekraft der Analyse für einzelne untersuchte Faktoren und die Möglichkeiten, zusätzliche, auch schwach wirkende Faktoren und ihre Interaktionen zu berücksichtigen. Die unabhängige Überprüfung und Verifizierung von Datenanalysen bleibt gleichwohl von zentraler Bedeutung.

8) Aus statistischen Zusammenhängen zwischen Variablen (Korrelationen) kann nicht ohne Weiteres auf Ursachen (kausale Effekte) oder Wirkmechanismen geschlossen werden. Letztere gilt es mittels zusätzlicher Argumente und Annahmen oder mittels Gewinnung zusätzlicher Daten, zum Beispiel aus Langzeit- oder experimentellen Studien, zu klären.

9) Besondere Bedeutung für den Einsatz und die weitere Entwicklung von Big-Data-Anwendungen hat das maschinelle Lernen. Hier „erlernen“ statistische Modelle anhand von Trainingsdatensätzen Berechnungsvorschriften, mit denen Daten in bestimmter Weise klassifiziert oder kategorisiert werden können. Eine zentrale Frage dabei ist, in welchem Umfang solche Techniken zur Entwicklung von entscheidungsfähigen und -befugten maschinellen Agenten führen, die beispielsweise auch an der Therapiegestaltung oder gesundheitspolitischen Entscheidungsprozessen beteiligt werden könnten.

10) Selbstlernende Systeme können anhand von Daten einer großen Gruppe von Menschen maßgebliche Faktoren, wie etwa gesundheitsrelevante Verhaltensweisen, ermitteln und einzelne Personen und Inhalte in diesem Koordinatensystem verorten. Solche Ansätze erlauben schnelle individualisierte Empfehlungen und Interaktionen mit maschinellen Assistenten. Sie gehen allerdings notgedrungen mit der Preisgabe persönlicher Informationen einher und erleichtern gegebenenfalls Täuschungen und die Manipulation persönlicher Entscheidungen.

11) Big-Data-gestützte Verfahren erkennen bei der Analyse von Zusammenhängen immer feinere Unterschiede zwischen Personen, wodurch eine stärkere Berücksichtigung

höchstpersönlicher Eigenschaften und Umstände möglich wird – etwa in der Diagnostik, Prognose und Therapie oder im Versicherungswesen hinsichtlich der Einstufung in Prämiengruppen. Bei der Bildung solcher Gruppen (Stratifizierung) durch komplexe Big-Data-Algorithmen ist es allerdings wichtig, mögliche Fehlerquellen zu berücksichtigen und zu minimieren.

12) Gesundheitsbezogene Daten, die einer bestimmten Person zugeordnet werden können, sind besonders sensibel, weil sie tiefe Einblicke in einen sehr intimen Bereich ermöglichen. Personenbezogene Daten können aus einer immer größeren Zahl von Quellen gesammelt und miteinander verknüpft werden, wobei im Verlauf des Auswertungsprozesses auch solche Daten Gesundheitsrelevanz erlangen können, die einen entsprechenden Anschein zunächst nicht erwecken, zum Beispiel Bewegungsdaten oder Einkaufsdaten.

13) Gesundheitsrelevante Daten fallen in verschiedenen, einander teilweise überschneidenden Kontexten an, von der medizinischen Praxis und gesundheitsbezogenen Forschung über Behörden und Versicherer bis hin zur aktiven oder unbeabsichtigten Datengenerierung durch Bürger bzw. Patienten. Big-Data-Technologien ermöglichen darüber hinaus eine umfassende Dekontextualisierung und Rekontextualisierung von Daten, die zu unterschiedlichen Zwecken erfasst, analysiert und neu verknüpft werden. Dies führt zu einer Entgrenzung des gesundheitsrelevanten Bereichs. Zudem erleichtert es die Deanonymisierung von Daten bzw. die Reidentifizierung einzelner Personen.

14) Weil alle Daten, die in irgendeiner Form erhoben werden, in Relation zur persönlichen Gesundheit interpretiert werden *können*, ist es prinzipiell möglich, all diese Daten auch als gesundheitsrelevant einzuschätzen. Ob bestimmte Daten als sensibel oder gesundheitsrelevant zu betrachten sind, lässt sich angesichts dieser Entwicklungen somit oft nicht mehr zum Zeitpunkt ihrer Erhebung bestimmen, sondern hängt in erster Linie vom Kontext ab, in dem sie verwendet werden. Dieser Kontext kann sich im Laufe der Zeit ändern.

15) An der Erhebung, Verarbeitung und Nutzung großer Datenmengen sind verschiedene Akteure mit unterschiedlichen Funktionen und zumindest teilweise gegenläufigen Interessen in vielfältigen Handlungskontexten beteiligt. Dabei lassen sich fünf ausgewählte Anwendungsbereiche von Big Data exemplarisch auf ihre jeweiligen Chancen und Risiken untersuchen: erstens die biomedizinische Forschung, zweitens die Gesundheitsversorgung, drittens Datennutzung durch Versicherer und Arbeitgeber, viertens die kommerzielle Verwertung gesundheitsrelevanter Daten durch global agierende IT- und Internetfirmen und fünftens ihre Erhebung durch Betroffene selbst.

16) In der biomedizinischen Forschung (Anwendungsbereich 1) soll die Auswertung großer Mengen gesundheitsrelevanter Daten zu einem besseren Verständnis wissenschaftlich relevanter Zusammenhänge und Prozesse führen. Zu den datenintensivsten Anwendungen gehören moderne bildgebende und molekularbiologische Verfahren, wie sie etwa in der Neurowissenschaft und den sogenannten Omik-Disziplinen (zum Beispiel Genomik, Proteomik, Metabolomik) eingesetzt werden.

17) Zentrale Akteure im wissenschaftlichen Bereich sind Forschungsinstitutionen und deren Mitarbeiter, aber auch Probanden und Patienten. Die Arbeit mit großen Datenmengen erfolgt in der Forschung in der Regel nach hohen und gut kontrollierbaren Standards der Datenerhebung, -verwendung und -sicherheit und häufig institutionenübergreifend. Wissenschaftsorganisationen machen sich die neuen technischen und infrastrukturellen Möglichkeiten von Big Data zunutze und vernetzen sich zum Zweck des Datenaustauschs und der gemeinsamen Analyse und Auswertung.

18) Bei vielen Erkrankungen sind die krankheitsbedingenden und -modulierenden Zusammenhänge sehr komplex. Big Data eröffnet Chancen, verschiedene Informationen integrativ in umfangreichen und quellenübergreifenden Analysen zusammenfassen. Für diese Integrationsleistung ist neben der bloßen Menge der einbezogenen Daten auch die Qualität ihrer interpretatorischen Aufbereitung von entscheidender Bedeutung.

19) Die Zusammenführung von Daten, die von mehreren Institutionen in oft unterschiedlichen Kontexten erhoben werden, bringt besondere Herausforderungen für den Einsatz von Big Data in der medizinischen Forschung mit sich. Vielfach fehlen einheitliche Standards zur Erfassung, Annotation und Qualitätssicherung von Daten ebenso wie gut funktionierende Regeln für den Datenaustausch. Das liegt zum einen an Datenschutzbedenken und einem Mangel an geeigneten Kontaktaufnahmemöglichkeiten und Einwilligungsmodellen für Patienten und Probanden zur Sekundärnutzung von Daten. Zum anderen gibt es Unsicherheiten und unterschiedliche Vorstellungen darüber, wer in welchem Ausmaß das Recht hat, über die generierten Daten zu verfügen.

20) Lösungsansätze bieten neben neuen Einwilligungsmodellen vor allem technische Maßnahmen für einen standardisierten Datenaustausch, der sowohl Datenqualität als auch hohe Schutzstandards garantiert, aber auch unterstützende regulatorische und Fördermaßnahmen sowie Initiativen für einen offenen Datenaustausch.

21) In der Gesundheitsversorgung (Anwendungsbereich 2) eröffnet der Einsatz von Big Data Chancen auf stärker personalisierte Behandlungskonzepte sowie Effektivitäts- und Effizienzsteigerungen. Der Rückgriff auf große Datenmengen ermöglicht eine bessere Stratifizierung von Patienten, sodass zum Beispiel Nebenwirkungen reduziert werden und unnötige Therapieversuche unterbleiben können. Die Sammlung und Auswertung

gesundheitsbezogener Daten erschließt zudem neue Potenziale bei der Früherkennung und Prävention von Erkrankungen.

22) Der Gesundheitssektor wird von einer Vielzahl von Akteuren mit teilweise divergierenden Interessen geprägt. Dazu gehören die Erbringer, Kostenträger und Empfänger von Gesundheitsleistungen, aber auch Behörden, Interessenverbände und Forscher mit einem unmittelbaren Bezug zur klinischen Praxis.

23) Den Chancen datenintensiver Ansätze stehen Risiken für Patienten gegenüber, etwa Kontrollverluste über die eigenen Daten, der immer weitergehend eröffnete Zugriff auf intime Informationen durch Leistungsanbieter („gläserner Patient"), sowie erleichterter Datenmissbrauch. Hinzu kommen Sorgen, dass eine verstärkte Nutzung Big-Data-gestützter Ansätze die persönliche Zuwendung zum Patienten weiter reduzieren und ihr unkritischer oder unsachgemäßer Einsatz zu Diagnose- und Behandlungsfehlern führen könnte.

24) Für Versicherer und Arbeitgeber (Anwendungsbereich 3) eröffnet Big Data umfangreiche neue Zugriffs- und Auswertungsmöglichkeiten, die von den geltenden rechtlichen Bestimmungen nicht durchgehend erfasst werden. Immer umfangreichere Datenmengen und -verknüpfungsoptionen ermöglichen zunehmend feinkörnige Profile einzelner Personen oder Personengruppen.

25) Damit verbunden ist eine Sorge vor Diskriminierung, etwa mit Blick auf Szenarien, in denen Versicherer und Arbeitgeber mithilfe der Analyse kommerziell verfügbarer, Big-Data-generierter persönlicher Verhaltensprofile gezielt risikoarme Antragsteller bzw. Bewerber auswählen oder diesen bessere Konditionen anbieten.

26) Auch innerhalb bestehender Verträge haben Arbeitgeber und Krankenversicherungen ein Interesse an der Gesundheit ihrer Vertragspartner, da im Krankheitsfall hohe Kosten entstehen können. Die Überwachung des Patienten- bzw. Arbeitnehmerverhaltens lässt Anreize für eine gesunde bzw. Sanktionen auf eine ungesunde Lebensführung zu. Wo es mit solchen Programmen gelingt, Krankenstände zu reduzieren, eröffnet dies für alle Beteiligten Chancen. Die Risiken dürfen gleichwohl nicht ignoriert werden. Prämienanpassungen oder Abmahnungen wegen gesundheitsschädlichen Verhaltens beispielsweise liegen nicht im Interesse der jeweiligen Datengeber.

27) Global agierende IT- und Internetfirmen (Anwendungsbereich 4) treten in erster Linie als Dienstleister auf. Auf der Grundlage ihres Zugangs zu riesigen Datenmengen und der geeigneten Dateninfrastruktur stellen sie Suchmaschinen, interaktive Informationsplattformen und Angebote wie Online-Shopping, aber auch eine breite Auswahl an multifunktionalen Geräten bereit. Dabei werden unterschiedliche Nutzerdaten in großem

Stil gesammelt, gespeichert und verwertet. Solchen Unternehmen, die zunehmend auch in gesundheitsrelevanten Bereichen agieren, ist es daher in besonderer Weise möglich, primär gesundheitsrelevante Daten mit zahlreichen anderen Informationen in Verbindung zu setzen. Hier besteht ein großes Missbrauchspotenzial.

28) Unternehmen bieten Software, Hardware, Technologieentwicklung und Online-Dienste für Big-Data-Anwendungen an. Sie stellen datenorientierten Institutionen Systeme, Algorithmen, Geräte und Infrastruktur zur Datenerhebung, Auswertung, Verwaltung und Speicherung zur Verfügung, mit denen Prozesse beschleunigt und verbessert werden sollen, um eine hocheffiziente Nutzung jeweils relevanter Informationen zu gewährleisten.

29) Die zunehmenden Aktivitäten digitaler Firmen im Gesundheitsbereich bieten Chancen für Forschung und Medizin, da große Internetkonzerne im Vergleich zum öffentlichen Sektor Zugriff auf wesentlich größere Datenmengen haben und oft mit leistungsfähigeren Analysemöglichkeiten sowie besseren technischen und finanziellen Ressourcen ausgestattet sind. Auf der anderen Seite stellen Einschränkungen beim Datenzugang für Datengeber und Nutzungsinteressenten aus Medizin und Forschung jedoch mitunter auch Hindernisse für den medizinischen Fortschritt dar.

30) Für die Erhebung gesundheitsrelevanter Daten durch Betroffene selbst (Anwendungsbereich 5) stehen viele tragbare Geräte mit Sensoren und Apps zur Verfügung, mit denen immer mehr individuelle Gesundheitsdaten sowie tägliche Aktivitäts- und Umweltdaten erfasst, aufbereitet und mit vorhandenen Datenbeständen verknüpft werden können. Die Digitalisierung der Lebenswelt ist zudem so weit fortgeschritten, dass alltägliche Verhaltensweisen und Kommunikationsformen häufig auch jenseits sozialer Netzwerke, Lifestyle-Apps und Ähnlichem eine automatische Datenproduktion nach sich ziehen.

31) Geräte und Apps zur Erhebung gesundheitsrelevanter Daten können den zeit- und ortsunabhängigen Zugang des Betroffenen zu seinen Gesundheitsinformationen und eine faktengestützte Gesundheitsversorgung erleichtern sowie einen gesundheitsbewussten Lebensstil und das persönliche Wohlergehen fördern. Sie eröffnen zudem Chancen für die Forschung, wenn sie als wichtige quantitative und qualitative Erweiterung der Datengrundlage verwendet werden.

32) Andererseits kann eine überzogene Selbstkontrolle mithilfe solcher Angebote zu einem übertriebenen, der Gesundheit abträglichen Optimierungsstreben sowie der Medikalisierung „natürlicher“ Lebensvorgänge beitragen. Zudem ist zweifelhaft, ob Selbstvermessung tatsächlich immer Ausdruck persönlicher Souveränität oder eher eine Form selbstinduzierter Fremdbestimmung ist. Befürchtet wird ferner die Diskriminierung

von Personen, die sich an solchen Messungen nicht beteiligen können oder wollen. Auch die bisherige Orientierung vieler Angebote an den wirtschaftlichen Interessen der Hersteller sowie Mängel bei Nutzerfreundlichkeit, Transparenz und Datenschutz lösen Kritik aus.

33) Zusammenfassend lassen sich anwendungskontextübergreifend die folgenden Stärken, Schwächen, Chancen und Risiken von Big Data in gesundheitsrelevanten Bereichen identifizieren: Zu den Stärken gehören die wachsende Datenbasis, die damit verbundene Entwicklung innovativer digitaler Instrumente sowie der hohe Grad der Vernetzung der Akteure. Zu den Schwächen gehören Schwankungen bei der Datenqualität, Intransparenz von Datenflüssen, Kontrollverluste sowie erhöhte Koordinations-, Regulierungs- und Qualifikationsanforderungen.

34) Als Chancen von Big Data sind vor allem bessere Stratifizierungsmöglichkeiten bei Diagnostik, Therapie und Prävention und damit verbundene Effizienz- und Effektivitätssteigerungen sowie die Unterstützung gesundheitsförderlichen Verhaltens zu nennen. Risiken bestehen hinsichtlich Entsolidarisierung, Verantwortungsdiffusion, Monopolisierung, Datenmissbrauch und informationeller Selbstgefährdung.

35) Die konkrete Beurteilung von Big-Data-Anwendungen mit Gesundheitsbezug hängt maßgeblich von den jeweils beteiligten Akteuren mit ihren unterschiedlichen Interessen und eigenen Chancen- und Risikoeinschätzungen sowie dem jeweiligen Anwendungskontext ab.

Rechtliche Vorgaben für Big Data

36) Big Data stellt eine erhebliche Herausforderung für das Rechtssystem dar. Zu berücksichtigen sind dabei vor allem verfassungsrechtliche Vorgaben, das allgemeine Datenschutzrecht, die speziellen Datenschutzbestimmungen des Gesundheitssektors sowie das Medizinprodukterecht, aber auch die zugrunde liegenden Anreizmechanismen und selbstregulative sowie hybride Steuerungsmechanismen.

37) Die wesentlichen Elemente des Datenschutzrechts sind grundrechtskonstituiert. Die zentrale verfassungsrechtliche Maßstabsnorm auf nationaler Ebene ist das Recht auf informationelle Selbstbestimmung, das vom Bundesverfassungsgericht im Volkszählungsurteil als spezifische Ausprägung des allgemeinen Persönlichkeitsrechts entwickelt worden ist. Es flankiert und erweitert den grundrechtlichen Schutz von Privatheit und Verhaltensfreiheit.

38) Diese Entfaltungsfreiheiten können mit wichtigen Gemeinwohlbelangen kollidieren wie der Förderung des wissenschaftlichen Fortschritts oder der Gewährleistung einer effektiven Gesundheitsversorgung. Konflikte können aber auch mit den Grundrechtspositionen anderer Privatrechtssubjekte bestehen, die ihnen zugängliche Informationen aufgreifen und verarbeiten wollen.

39) Das Datenschutzrecht orientiert sich an den verfassungsrechtlichen Vorgaben. Es wurde allerdings nicht für Verwendungskontexte geschaffen, die erst durch die neuen technischen Möglichkeiten relevant werden, und ist auch nach seinen jüngsten, durch die europäische Datenschutz-Grundverordnung (DSGVO) veranlassten Veränderungen auf das Phänomen Big Data unzureichend eingestellt. Dies gilt ungeachtet der klaren Fortschritte, die diese neuen Vorgaben etwa mit Blick auf die Etablierung grenzüberschreitender Standards sowie die stärkere Einbeziehung des Konzepts von *privacy by design* bedeuten.

40) Grundlegende Annahmen, zentrale Prinzipien und Zielvorgaben des überkommenen Datenschutzrechts sind mit den Besonderheiten von Big-Data-Anwendungen kaum in Einklang zu bringen. Die traditionellen datenschutzrechtlichen Grundsätze des Personenbezugs, der Zweckbindung und Erforderlichkeit der Datenerhebung, der Datensparsamkeit, der Einwilligung und Transparenz stehen der spezifischen Eigenlogik von Big Data entgegen. Will man weder den Einsatz von Big Data grundsätzlich untersagen noch relevante Einbußen am Schutzniveau hinnehmen, müssen neue Gestaltungsoptionen und Regelungsmechanismen entwickelt werden.

41) Das geltende Datenschutzrecht knüpft an den Personenbezug von Daten an und legt besonderen Wert auf die damit einhergehenden spezifischen Zweckbindungen. Für Big Data ist demgegenüber entscheidend, dass bei der Erfassung der Daten die künftigen Anwendungen nicht vorhersehbar sind und auch der Personenbezug bzw. der Bezug zu ihrer Gesundheit unter Umständen erst nachträglich hergestellt wird. Daten, die zu anderen Zwecken gespeichert wurden, werden oft für neue Zwecke ausgewertet oder es werden Daten für noch unbestimmte Zwecke erhoben.

42) In augenfälligem Widerspruch zu Big Data steht ferner der Grundsatz der Datensparsamkeit bzw. Datenminimierung, nach dem so wenig personenbezogene Daten wie möglich erhoben, verarbeitet oder genutzt werden sollen. Das führt leicht zu einem weitgehenden Ausschluss der Möglichkeiten von Big Data. Weil aber mit der Menge an gespeicherten Daten zugleich das Gefährdungspotenzial für das Recht auf informationelle Selbstbestimmung wächst, bedarf es wirksamer alternativer Schutzmechanismen.

43) Auch bei dem im Datenschutzrecht normierten Erfordernis der Einwilligung, wonach eine Datenverwendung nur erlaubt ist, wenn der Betroffene bei Abgabe seiner Einwilligung die Bedeutung und Tragweite der beabsichtigten Datenverwendung überblickt, zeigen sich Inkompatibilitäten mit Big Data. Schon jetzt ist häufig zweifelhaft, dass Datengeber insbesondere die Verwendungszwecke und die damit verbundenen Implikationen tatsächlich verstehen. Big Data verstärkt diese allgemeine Problematik noch einmal erheblich, da künftige Verwendungsarten zum Zeitpunkt der Datenerhebung oftmals unbekannt sind.

44) Das geltende Datenschutzrecht bietet zudem jenseits der Einwilligung nur wenige Möglichkeiten, auf das weitere Schicksal der Daten Einfluss zu nehmen. Jede weitere Verwendung bedarf einer neuen Einwilligung, und sind Daten einmal mit Einwilligung erhoben, können sie von dem Betroffenen nicht mehr weiterverfolgt werden. Die Dynamik von Big Data passt nicht in dieses Regelungskonzept. Gerade wenn man die Zustimmung der Betroffenen für ein zentrales Erfordernis des Datenschutzes erachtet, ist deshalb nach Wegen zu suchen, wie dies auch unter Big-Data-Bedingungen funktional sinnvoll möglich ist.

45) Big Data intensiviert zudem gerade durch die Verknüpfung vielfältiger Daten dsie Möglichkeiten der Reidentifizierung und verstärkt damit Zweifel an der Effektivität des Anonymisierungs- bzw. Pseudonymisierungsgebots. Die Frage, inwieweit und ab welchem Grad die Gefahr einer Reidentifizierung für sich genommen anonymisierter Daten für die Annahme eines Personenbezugs der Daten ausreichend ist und wie das gemessen werden kann, verschärft die Problematik um den ohnehin schon umstrittenen Begriff des Personenbezugs im Datenschutzrecht.

46) Die Rechte auf Auskunft, Berichtigung, Löschung und Sperrung dienen der Transparenz, bieten aber häufig keinen effektiven Schutz. Gerade im Kontext von Big Data wird der Datengeber kaum alle potenziellen Anspruchsgegner kennen. Auch die von den Auskunftsrechten umfasste Nachvollziehbarkeit des Datenverarbeitungsprozesses gestaltet sich angesichts komplexer und selbstlernender Algorithmen schwierig. Damit läuft auch das Recht auf Berichtigung und Löschung leer, da der Betroffene diese Rechte ohne eine umfassende Auskunft nicht wahrnehmen kann.

47) Diese auf das allgemeine Datenschutzrecht bezogene Defizitanalyse kann mit gewissen Einschränkungen auf das besondere Gesundheitsdatenschutzrecht übertragen werden, das das zum Teil bereichsspezifisch ausgestaltete Datenschutzrecht um die zivil-, straf- und berufsrechtlichen Vorgaben der ärztlichen Schweigepflicht ergänzt. Im Kern bleiben auch die normativen Lösungsansätze des Gesundheitsdatenschutzrechts weitgehend einer Problemperspektive aus der „Vor-Big-Data-Zeit“ verhaftet.

48) Eine kompensatorische Wirkung könnten die Bestimmungen des Medizinprodukterechts entfalten, das den freien Verkehr mit Medizinprodukten regelt und dabei gleichzeitig die Sicherheit, Eignung und Leistung der Medizinprodukte zum Schutz der Patienten, Anwender und Dritter zu gewährleisten versucht. Anders als Arzneimittel bedürfen Medizinprodukte keiner staatlichen Zulassung, wohl aber der Zertifizierung nach einer produktspezifischen Risikobewertung, Risikominimierung und Risiko-Nutzen-Analyse sowie einem dem Risiko des Produkts angemessenen Verfahren der Konformitätsbewertung.

49) Software kann als Medizinprodukt zu klassifizieren sein, wenn sie eine medizinische Zweckbestimmung hat. Ob dies der Fall ist, hängt maßgeblich von den Angaben des Herstellers ab. Die Abgrenzung zwischen medizinischen Anwendungen und bloßen Lifestyle- oder Fitness-Apps gestaltet sich allerdings in der Praxis oft schwierig.

50) Die Vorgaben des Krankenversicherungsrechts erweisen sich ebenfalls als relevant für Big Data. Die Einordnung von M-Health-Applikationen in die Vergütung der gesetzlichen wie privaten Krankenversicherung könnte zum Beispiel finanzielle Anreize für Entwickler solcher Angebote und damit ein Gegenmodell zum „Zahlen mit Daten" schaffen. Dabei sind jedoch Wirksamkeitsnachweise zu erbringen. Ebenso sind Diskriminierungen zu vermeiden, auch bei der Berücksichtigung solcher Daten bei der Beitragsgestaltung.

51) Angesichts der jüngst erfolgten umfassenden Neuordnung des Datenschutzrechts durch die DSGVO und das BDSG n. F. ist zwar abzuwarten, ob und wie sich die neuen Normen und Mechanismen bewähren. Indes dürfte feststehen, dass einige Grundprinzipien des geltenden Datenschutzrechts mit dem Konzept von Big Data kaum in Einklang zu bringen sind. Dieser Spannung kann im Rahmen der vom Verfassungsrecht gewährten Handlungsspielräume mit flexiblen, innovationsoffenen Regelungen Rechnung getragen werden, die auch die Verwendung komplexerer, privatrechtlicher wie privat-staatlich kooperativer Steuerungsbeiträge mitberücksichtigen.

52) Insbesondere wäre zu prüfen, ob der Mangel an Konkretheit von gesundheitsrelevanten Big-Data-Anwendungen durch zusätzliche technisch-organisatorische sowie materiell- und verfahrensrechtliche Sicherungen kompensiert werden kann. Im Zuge der Weiterentwicklung des Datenschutzrechts könnte vor allem eine stärker ausdifferenzierte, den Besonderheiten eines Regelungsbereichs und den Präferenzen der Betroffenen Raum gebende Konzeption von Einwilligungsmodellen oder eine verstärkte Erhebung und Nutzung von Daten auf Basis gesetzlicher Erlaubnisnormen in den Blick genommen werden. Auch dem Privatrecht kommt große Bedeutung für die Weiterentwicklung des

Datenschutzes zu, vor allem dem Verbraucherrecht, dem Haftungsrecht sowie den Regelungen für die Zuordnung von Daten und die Befugnis, über ihre Verwendung zu bestimmen („Eigentum“ an Daten).

53) Sämtliche Steuerungsansätze für Big Data haben mit dem Problem zu kämpfen, mit einer territorial begrenzten Rechtsetzung auf ein seiner Natur nach globales Phänomen zu reagieren. Die jeweiligen Datenschutzrechte sind international gesehen sehr unterschiedlich, was sowohl die Betroffenen als auch die Regulierer vor besondere Herausforderungen stellt. Trotz vielfältiger Harmonierungsbemühungen gibt es nach wie vor zahlreiche praktische Hindernisse, die einer effektiven grenzüberschreitenden Rechtsverfolgung im Wege stehen.

54) Angesichts der spezifischen Dynamik und Volatilität des Regelungsbereichs gewinnen zudem nicht hoheitliche und kooperative Steuerungsmechanismen an Bedeutung, zum Beispiel Zertifizierungen mit Datenschutz- bzw. Datensicherheitssiegeln oder Handlungsregeln und Kodizes für Wissenschaft und Wirtschaft.

Zur Ethik von Big Data und Gesundheit

55) Von Big Data sind sowohl ethische Orientierungsmuster betroffen, die normativ und evaluativ die Rolle, Funktion und Stellung des datengebenden Individuums thematisieren, als auch Maßgaben sozialer Orientierung. Zu den relevanten Begriffen gehören Freiheit und Selbstbestimmung, Privatheit und Intimität, Souveränität und Macht, Schadensvermeidung und Wohltätigkeit sowie Gerechtigkeit, Solidarität und Verantwortung.

56) Der Ausdruck Freiheit wird in vielen Bedeutungen verwendet. Dabei ist zu unterscheiden zwischen Handlungsurheberschaft als grundsätzlicher Freiheitsbedingung und Selbstbestimmung als Praktisch-Werden von Freiheit in Abhängigkeit von mehr oder weniger deutlich erfahrbaren Umständen. Selbstbestimmt sind Handlungsurheber in unterschiedlichen Graden.

57) Der Begriff der Selbstbestimmung bezeichnet sowohl die Fähigkeit einer Person, ihr Leben nach ihren eigenen Vorstellungen zu gestalten, als auch die tatsächliche Ausübung dieser Fähigkeit und einer als ideal vorgestellten Form der Lebensführung. Von diesen Formen personaler Selbstbestimmung ist der rechtliche Schutz ihrer Ausübung zu unterscheiden. Formen und Grade der Ausübung von Selbstbestimmung sind von erheblicher praktischer Bedeutung. So kann man in bestimmten Zusammenhängen sein Recht auf Selbstbestimmung delegieren oder können Einschränkungen der Selbstbestimmungsfähigkeit teilweise durch Vertreter kompensiert werden.

58) Im Kontext von Big Data sind vor allem für Biobanken in den letzten Jahren neue Einwilligungsmodelle entwickelt worden, die mit Blick auf die Selbstbestimmung der Datengeber eine Balance zwischen einer unrealistisch engen Zweckbestimmung und einer einmaligen, allzu breiten Freigabe garantieren sollen. Hierbei werden dynamische Modelle, bei denen mehrfach in jeweils einzelne Elemente eingewilligt werden kann, um weitere Optionen ergänzt, etwa um Möglichkeiten zur Delegation. Teilnehmer können zudem entscheiden, welche Form der Einwilligung sie grundsätzlich bevorzugen.

59) Für die Beurteilung von Selbstbestimmung ist auch der soziale Kontext des Handelnden einzubeziehen. Frei zu sein und selbstbestimmt handeln zu können, bedeutet vor diesem Hintergrund zumindest die realistische Möglichkeit, die eigene Identität zu bewahren und zu gestalten sowie die eigenen Handlungen vor sich und anderen zu verantworten. Dazu sind verlässliche und faire rechtsstaatliche Standards notwendig, die ohne Ansehen der Person gelten.

60) Privatheit bezeichnet klassischerweise das Recht, in Ruhe gelassen zu werden bzw. eine Lebenssphäre, in der eine ungewollte Kontrolle durch die Öffentlichkeit und Rechtfertigungsnotwendigkeiten weitgehend zurückgedrängt sind. Eng mit Privatheit verbunden ist der Begriff der Intimität. Er kennzeichnet Lebensbereiche, die ausschließlich den unmittelbar Betroffenen vorbehalten bleiben und deren Details diese – wenn überhaupt – nur ausgewählten anderen selbstbestimmt zugänglich machen.

61) Was als privat und intim gilt oder gelten sollte, ist in erheblichem Umfang kulturvariant. Dessen ungeachtet lässt sich die Wahrung der Privatsphäre jedoch normativ mit ihrer sozialanthropologischen Bedeutsamkeit begründen. Nur in der Sphäre des Privaten können sich soziale Nahbeziehungen wie auch die Entwicklungsbedingungen personaler Identität ausbilden. Privatheit eröffnet Räume von Intimität und Vertraulichkeit, in denen Personen Beziehungen pflegen und unbefangen und unverstellt sie selbst sein können – nach außen abgeschirmt, nach innen aber offen.

62) Mit Blick auf Big Data ergeben sich mögliche Privatheitsgefährdungen aus den vielfältigen neuen Gelegenheiten zur Erfassung, Analyse und neuen Verknüpfung von Daten und Informationen sowie der damit einhergehenden erschwerten Anonymisierung und Pseudonymisierung. Je mehr intime Details digital preisgegeben werden können, desto eher droht zudem eine selbstinduzierte Fremdbestimmung bzw. informationelle Selbstgefährdung im Rahmen einer persönlichen Lebensführung, die sich maßgeblich von äußeren Einflussfaktoren abhängig macht.

63) Auch wenn in der digitalen Gesellschaft eine vollständige Kontrolle der eigenen Datenspuren unmöglich geworden sein mag, legen Menschen Wert darauf, kontextabhängig mitbestimmen zu können, wie ihre Daten gebraucht und weiterverwendet werden.

Gleichzeitig gewinnt die Erwartung an Bedeutung, dass Datennutzer die ihnen zur Verfügung gestellten Daten auch im Rahmen von De- und Rekontextualisierungen vertraulich und vertrauenswürdig behandeln.

64) Wie Privatheit unter den Bedingungen von Big Data zu schützen ist, betrifft nicht nur Individuen, sondern auch Gruppen. Die Analyse großer Datenmengen erlaubt es oft, auf Merkmalskombinationen zahlreicher Personen zu schließen. Betroffene werden von Algorithmen zu Gruppen zusammengefasst, mit möglicherweise stigmatisierenden, diskriminierenden oder exkludierenden Folgen. Eine solche Zuordnung ist für den Einzelnen oft nicht erkennbar.

65) Zentrale Bedeutung im Kontext von Big Data erhält der Begriff der Souveränität. Er entstammt kulturhistorisch vornehmlich dem religiös-politischen Bereich und wird in zahlreichen Lebensbereichen unterschiedlich konkretisiert. Souveränität galt als jene Eigenschaft Gottes oder eines absolutistischen Herrschers, kraft derer er absolut und unbedingt von anderen Mächten alles zu tun oder zu lassen imstande sei. Andere Konzepte von Souveränität betonen anstelle einer vermeintlichen absoluten Ungebundenheit des souveränen Subjekts die Abhängigkeiten seiner physischen wie sozialen Leiblichkeit.

66) Nach einem Souveränitätsverständnis, das jedenfalls eine Verfügungsgewalt von Menschen über andere Menschen grundsätzlich ausschließt, sind personenbezogene Daten für die Sammler und Nutzer nur Leihgabe, niemals frei und willkürlich verfügbares Eigentum. Das bedeutet zwar umgekehrt nicht, dass damit der Datengeber automatisch Eigentümer seiner Daten ist oder selbst seinen Souveränitätsanspruch unter allen Umständen realisieren kann, begründet jedoch im Prinzip weitreichende Kontrollmöglichkeiten des Individuums.

67) Der Begriff der Souveränität ist eng mit dem der Macht verbunden. Souveränität verwirklicht sich im Modus der Ausübung von Macht und wird umgekehrt begrenzt durch die Ausübung souveräner Macht anderer. Im Kontext von Big Data werden spezifische Formen der Machtausübung ethisch bedeutsam: erstens solche, mit denen Präferenzen und Überzeugungen anderer manipuliert werden können; und zweitens solche, die darüber hinaus sogar eine subtile Formung, Veränderung und damit mögliche Beherrschung ihrer Charaktere ermöglichen.

68) Der Einsatz von Big-Data-Algorithmen eröffnet Anbietern von Internetdiensten neue Möglichkeiten gezielter Einflussnahme auf das Denken, Fühlen und Handeln der Nutzer solcher Dienste. Das Spektrum reicht von offenem Nudging, mit dem gesundheitsförderliches Verhalten subtil angeregt werden soll, bis hin zu verdeckten und vor allem fremdnützigen manipulierenden Interventionen. Letztere sind ethisch zumindest besonders rechtfertigungsbedürftig. Denn sie entziehen sich der kognitiven Kontrolle

durch den Betroffenen, umgehen damit seine Möglichkeiten zur Beherrschung der Bedingungen seines Handelns und untergraben so seine Selbstbestimmtheit.

69) Ein weiterer relevanter normativer Bezugspunkt ergibt sich aus der moralischen Verpflichtung zur Wohltätigkeit, wonach das eigene Handeln in vielen Situationen über die bloße Schadensvermeidung hinaus auch Vorteile für andere, insbesondere für hilfsbedürftige Menschen erbringen soll. Für das Thema Big Data und Gesundheit sind vor allem zwei Aspekte von Wohltätigkeit von besonderem Interesse: zum einen der Wissens- und Erkenntniszuwachs und zum anderen der therapeutische Mehrwert, der aus neuen Möglichkeiten der digitalen Informationsgewinnung und -verarbeitung großer Datenmengen im Gesundheitsbereich für unterschiedliche Beteiligte resultiert.

70) Wissen und Erkenntnis sind von großer Bedeutung für die Selbstkonstitution des Individuums und seine Befähigung zur autonomen Lebensführung. Darüber hinaus kommt der kritischen Überprüfung, der Sicherung und der Ausweitung von Wissensbeständen eine wichtige gesellschaftliche Funktion zu.

71) Um die mit Wissenszuwachs verbundenen Ziele zu erreichen, bedarf es des Schutzes einer der Wahrhaftigkeit verpflichteten Kommunikation. Zu deren Sicherung insbesondere auf dem Feld der Wissenschaften haben sich differenzierte methodologische und wissenschaftstheoretische Maßgaben entwickelt. Daher ist darauf zu achten, dass neue digitale Verfahren der Datensammlung, -auswertung und -verknüpfung nicht zur Absenkung epistemischer Standards oder zu Einbußen der Zuverlässigkeit daraus gewonnener Aussagen führen.

72) Zu klären ist auch, welchen Personengruppen die durch Big Data erzielten Erkenntnisfortschritte jeweils primär zugutekommen sollen, wie sich derzeit bestehende Hindernisse auf dem Wege einer effizienteren Gestaltung des Datennutzungsprozesses beseitigen lassen und wie eine gerechte Verteilung jener positiven Effekte erreicht werden kann, die aus dem zu erwartenden Wissenszuwachs resultieren.

73) Die Sammlung und Weitergabe großer Mengen gesundheitsbezogener Daten berührt grundlegende Fragen der Gerechtigkeit. Als normierendes Prinzip sozialer Beziehungen gebietet die Gerechtigkeit, willkürliche Privilegierungen Einzelner oder bestimmter Gruppen zu vermeiden. Vielmehr ist das jedem Einzelnen Angemessene auf rationale Weise zu bestimmen. Das setzt voraus, dass einheitliche Kriterien Verwendung finden und Unterschiede in der Behandlung Einzelner normativ konsensfähig begründet werden.

74) Mit Blick auf Big-Data-Anwendungen im Gesundheitsbereich sind vor allem vier Problemfelder besonders gerechtigkeitsrelevant: erstens der Zugang zu Datensammlungen

für den Forschungsbereich, zweitens die schleichende Etablierung monopolartiger Strukturen, drittens die Einbeziehung von Gesundheits-Apps und verschiedenen, der privaten Selbstvermessung dienenden Geräten in die Tarifgestaltung von Krankenversicherungen und viertens Aspekte der Befähigungsgerechtigkeit im Hinblick auf einen verantwortlichen Umgang mit gesundheitsbezogenen Daten.

75) Der Begriff der Solidarität bezeichnet prosoziale Handlungen, Praktiken und Dispositionen sowie institutionelle, politische und vertragliche Regelungen, die dazu dienen sollen, andere zu unterstützen. Solidarität wird vielfach als komplementär – und oft auch subsidiär – zur Gerechtigkeit verstanden. Sie entsteht regelmäßig vor dem Hintergrund gemeinsamer Ziele einer Gruppe, angesichts einer gemeinsamen Herausforderung oder auch aus der geteilten Vorstellung vom guten Leben in einer Solidargemeinschaft.

76) Solidarität gründet häufig in Reziprozitätserwartungen. Die Bereitschaft zur Solidarität kann nachlassen, wenn Zweifel an der Einlösbarkeit solcher Erwartungen entstehen, etwa wenn auf Dauer der Eindruck entsteht, die Hilfs- und Unterstützungsbedürftigkeit anderer werde durch deren fahrlässige Selbstschädigung oder mangelnde Eigeninitiative verursacht und das Solidaritätsgefüge damit überstrapaziert.

77) Die durch Big Data ermöglichte Auswertung umfänglicher und vielfältiger gesundheitsrelevanter Daten erlaubt die Erstellung genauerer Risikoprofile. Damit verbindet sich die Sorge, dass die Annahme einer allen gemeinsamen Vulnerabilität gegenüber Krankheitsrisiken, die nicht sicher antizipierbar sind, als Grundlage der Solidargemeinschaft in der gesetzlichen Krankenversicherung und der fairen Vertragsgestaltung in der privaten Krankenversicherung infrage gestellt werden könnte. Dann könnten Niedrigrisikogruppen verstärkt die Solidargemeinschaft verlassen, wodurch für Letztere erhebliche Mehrbelastungen entstünden.

78) Innerhalb der gesetzlichen Krankenversicherung unterlaufen verhaltensdatenbasierte Versicherungstarife den Solidargedanken, der die Absicherung gegen krankheitsbedingte Vulnerabilität weitgehend ohne Ansicht individueller verhaltensbedingter Risiken fordert. Die private Krankenversicherung arbeitet hingegen mit risikoäquivalenten Prämien. Auch hier kann sich eine Umverteilung von Risiken zuungunsten der Versicherten ergeben, falls Prämien künftig auf Grundlage der durch Big Data ermöglichten kontinuierlichen Erhebung und Auswertung individueller Daten auch nach Abschluss der Versicherung regelmäßig angepasst würden. Dies würde das Versicherungsprinzip, nach dem Risiken von einer größeren Gruppe gemeinsam getragen werden und Tarife auch nicht individualisiert angepasst werden dürfen, gänzlich aushebeln. Es könnten zunehmend kleine Tarifgruppen entstehen, bei denen Schadensfälle dann umso schneller zu Beitragserhöhungen führen.

79) Zudem könnten privat Versicherten, die nicht bereit oder in der Lage sind, in einem verhaltensbasierten Versicherungsmodell mitzuwirken, finanzielle Vorteile vorenthalten werden, was auf lange Sicht zu Prämiennachteilen führen muss. Unabhängig davon, ob sie sich gesundheitsförderlich verhalten oder nicht, würden sie dafür bestraft, dass sie ihre Daten nicht der Versicherung überlassen, und somit durch die Ausübung ihres Rechts auf informationelle Selbstbestimmung benachteiligt.

80) Grundsätzlich hat die Freiheit zur Lebensgestaltung und Selbstentfaltung Vorrang vor einer strikten und permanenten Pflicht zur Vermeidung aller Gesundheitsrisiken. Dies gilt zwar nicht unbegrenzt, doch ließen sich die dauernde gezielte Sammlung von Daten über die individuelle Lebensführung und die Nutzung Big-Data-gespeister Risikoprofile, die alle Lebensbereiche umfassen, schwerlich als zumutbare Erwartung an die Mitverantwortung für die eigene Gesundheit qualifizieren.

81) Ob und wie gesetzliche Krankenkassen gesundheitliche Eigenverantwortung berücksichtigen und das Gesundheitsverhalten ihrer Versicherten beeinflussen dürfen, ist umstritten. Datenbasierte Anreizsysteme könnten eine sehr intensive und invasiv-überwachende Wirksamkeit entfalten. Die differenzierte Offenlegung von Risikofaktoren über Big-Data-Analysen, die Daten aus allen Lebensbereichen integrieren, könnte künftig aber auch ergeben, dass der weitaus überwiegende Teil der Bevölkerung gemischte Risikoprofile hat, die protektive und günstige Faktoren ebenso einschließen wie negative Faktoren körperlicher, mentaler, verhaltensbedingter und anderer Art.

82) In verschiedenen Bereichen der Medizin hat der Einsatz von Big-Data-Technologien bereits zur Entwicklung neuer pro-sozialer Unterstützungspraktiken geführt, beispielsweise zur Bildung kleinerer Gruppen von Patienten, die insbesondere seltene Krankheitsrisiken oder -erfahrungen teilen und ihre Daten und Bioproben in gemeinschaftlichen Pools zusammenführen, um sie für die Forschung an ihrem Krankheitsbild zur Verfügung zu stellen.

83) Andere Solidaritätsgewinne sind gegenwärtig in Online-Foren zu beobachten, in die Patienten ihre Erfahrungen und Krankheitsdaten aus Klinik und Selbstvermessung einspeisen, sie dort austauschen, gemeinsam diskutieren und für das individuelle Krankheitsmanagement nutzen. Mit der zunehmenden Entwicklung von online vernetzten Instrumenten für die Patienten-Selbsthilfe steht zu erwarten, dass derartige Praktiken zunehmen werden.

84) Verantwortung als moralische Kategorie lässt sich nach Handlungs- und Entscheidungstypen, aber auch nach der Ausgestaltung institutioneller Strukturen differenzieren. Sie kann moralisch, rechtlich, politisch und vertraglich sowie vor und nach einer

Handlung oder Entscheidung eingefordert und übernommen werden. Die damit verbundenen unterschiedlichen Typen von Verantwortung stehen oft in einem sachlichen Wechselverhältnis: Man erwartet genau von demjenigen die Übernahme von Verantwortung für die Zukunft, den man in einem tatsächlichen Schadensfall zur Rechenschaft ziehen würde. Das komplexe Zusammenspiel zwischen Einzelnen, Institutionen und Technik beim Einsatz von Big Data gewinnt im gesundheitsrelevanten Bereich besondere Bedeutung. Vermieden werden sollte eine undurchsichtige Diffusion von Verantwortung, die dort droht, wo viele Akteure und hoch technisierte Prozesse zusammenwirken.

85) Damit individuelle Datengeber auch im Big-Data-Zeitalter Verantwortung für ihre Daten übernehmen können, bedarf es bestimmter Rahmenbedingungen, die sich technisch wie organisatorisch leicht und effektiv nutzen lassen. Im sensiblen Gesundheitsbereich gelten zudem erhöhte Sorgfaltspflichten, etwa für Forscher oder Ärzte.

86) Zu den Möglichkeiten von Unternehmen, Big-Data-Prozesse verantwortlich zu gestalten, gehört es vor allem, Bedingungen dafür zu schaffen, gegebene Zustimmungen widerrufbar zu machen und die Verwaltung von Daten auf Abruf zu gestalten. Davon ausnehmen könnte man hinreichend aggregierte Daten, abgeleitete Daten oder Modelle, die nachweislich keinen Rückschluss auf den Einzelnen erlauben. Mit solchen Ansätzen die Big-Data-spezifischen De- und Rekontextualisierungen bei gleichzeitiger Wahrung hoher Anonymisierungsstandards zu ermöglichen und Institutionsvertrauen zu schaffen, dürfte eine der entscheidenden Aufgaben der Zukunft sein.

87) Eine weitere Möglichkeit, Verantwortung für die Rechte des Individuums zu übernehmen und dabei dennoch legitime Geschäftsinteressen zu wahren, wären Stellvertretersysteme an den programmatischen Schnittstellen in Datennetzwerken. Solche Schnittstellen könnten als „Datenagenten“ Präferenzen von Datengebern für die Datenhandhabung umsetzen. Hierdurch würde eine individuelle Datenverwaltung durch eine programmatische Verwaltung ersetzt, die dem Einzelnen eine technisch niedrigschwellige und reliable Möglichkeit gäbe, Verantwortung für die Wahl eigener kurz-, mittel- und langfristiger Strategien der Datenhandhabung zu übernehmen, ohne jede Einzelfrage selbst entscheiden zu müssen.

88) Unternehmen können Verantwortung auch übernehmen, indem sie ihre Verfahren besser überprüfbar machen, etwa mit Blick auf die verwendeten Algorithmen, den Ausschluss systematischer Benachteiligungen, die Einhaltung von Regeln zur Datenaufbewahrung, Anonymisierung oder Datenlöschung und die lückenlose und manipulationssichere Protokollierung der Herkunft, Verarbeitung, Verwendung und des Austauschs von Daten.

89) Neben staatlicher Regulierung gibt es weitere Möglichkeiten, die Übernahme von Verantwortung durch institutionelle Akteure zu gewährleisten bzw. zu fördern. Zertifizierungen, Qualitätssiegel oder Selbstverpflichtungen, die von Interessen- oder Berufsverbänden bereitgestellt und überprüft werden, können beispielsweise Vertrauen in die jeweiligen Organisationen und Prozesse stärken.

90) Eine weitere Verantwortungsfrage betrifft mögliche Eingriffe von Organisationen in die persönliche Kommunikation zwischen Nutzern, beispielsweise in Form gesundheitsförderlicher Hinweise oder Hilfsangebote. Dagegen spricht einerseits die Ablehnung offensichtlicher Eingriffe in die Privat- oder Intimsphäre. Wäre die Funktionssicherheit solcher Algorithmen aber wissenschaftlich gut belegt, müsste man andererseits aus ethischer Perspektive auch berücksichtigen, dass ihr Einsatz gegebenenfalls schweres Leid oder sogar Todesfälle verhindern könnte, beispielsweise bei Hilfsangeboten für suizidgefährdete Personen in sozialen Netzwerken.

91) Der Staat kann auf nationaler Ebene, im Verbund der EU, aber auch als völkerrechtlicher Akteur Verantwortung übernehmen. Mit Blick auf die angedeutete Problematik der Rechtsumsetzung sollte allerdings ein regulatorischer Subsidiaritätsgrundsatz gelten, der Selbstverpflichtungen und Zertifikaten den Vorzug vor detaillierten rechtlichen Regelungen lässt, sofern und solange diese effektiv funktionieren.

92) Angesichts der drei Ebenen möglicher Verantwortungszuschreibung im Bereich gesundheitsbezogener Big-Data-Anwendungen (Individuen, Organisationen, Staat) bleiben Individuen zwar in der Pflicht, Verantwortung für die Nutzung ihrer Daten zu übernehmen. Vornehmlich tragen jedoch die Daten sammelnden, verarbeitenden und weitergebenden Organisationen Verantwortung dafür, Rahmenbedingungen für die verantwortliche informationelle Freiheitsgestaltung der Datengeber zu gewährleisten.

93) Je weniger Organisationen willens oder fähig sind, technische Möglichkeiten bereitzustellen, die dem Einzelnen die Kontrolle über seine Daten erleichtern, desto mehr drängt sich in verantwortungsethischer Perspektive die Notwendigkeit für den Staat auf, gewährleistend, überwachend und gegebenenfalls auch regulierend und sanktionierend einzugreifen. Das Ziel, dem Einzelnen die Möglichkeit zum souveränen Umgang mit seinen Daten zu geben, ist nur erreichbar, wenn dazu auf allen Seiten die jeweils gebotene Verantwortung übernommen wird.

Datensouveränität als informationelle Freiheitsgestaltung

94) Datensouveränität, verstanden als eine den Chancen und Risiken von Big Data angemessene verantwortliche informationelle Freiheitsgestaltung, sollte das zentrale ethische und rechtliche Ziel im Umgang mit Big Data sein.

95) Der Begriff der informationellen Freiheitsgestaltung entwickelt das Konzept der informationellen Selbstbestimmung weiter. Er gründet nicht in einem eigentumsanalogen Ausschlussrecht, sondern in der Befugnis, selbst zu bestimmen, mit welchen Inhalten jemand in Beziehung zu seiner Umwelt tritt. Informationelle Freiheitsgestaltung in diesem Sinne meint interaktive Persönlichkeitsentfaltung unter Wahrung von Privatheit in einer vernetzten Welt und ist gekennzeichnet durch die Möglichkeit, auf Basis persönlicher Präferenzen effektiv in den Strom persönlich relevanter Daten eingreifen zu können. Verantwortlich ist eine solche Freiheitsgestaltung dann, wenn sie sich gleichzeitig an den gesellschaftlichen Anforderungen von Solidarität und Gerechtigkeit orientiert.

96) Mit Datensouveränität im hier vertretenen Sinne werden weder die tradierten, letztlich kaum veränderten Regulierungsansätze des Datenschutzes nur unter neuem Namen fortgeschrieben, noch wird damit eine vollständige Neuorientierung oder gar eine Aufgabe des herkömmlichen Datenschutzgedankens oder die generelle Absenkung des bestehenden Schutzniveaus gefordert. Vielmehr geht es darum, die benannten normativen Grundanforderungen, einschließlich der ethisch wie grundrechtlich fundierten informationellen Selbstbestimmung und damit auch des Datenschutzes, unter den Bedingungen von Big Data zur Geltung zu bringen.

97) Datenschutz war und ist kein Selbstzweck, sondern dient dem Schutz der Person: ihrer Privatsphäre ebenso wie der freien Entfaltung ihrer Persönlichkeit in der Öffentlichkeit. Mit dem Begriff der Datensouveränität wird aber zugleich die Absicht betont, den souveränen, also selbstbestimmten und verantwortlichen, Umgang des Einzelnen mit seinen eigenen personenbezogenen Daten mit einer Realisierung der Potenziale zu verknüpfen, die Big Data sowohl gesellschaftlich als auch für die individuelle Lebensgestaltung eröffnet.

98) Das Ziel einer verantwortlichen informationellen Freiheitsgestaltung im Gesundheitsbereich besteht darin, die Big-Data-spezifischen Potenziale für die medizinbezogene Forschung, die klinische Anwendung und das individuelle Gesundheitsverhalten zu nutzen und die damit einhergehenden Risiken auf ein Minimum zu reduzieren.

99) Bei der Wahrnehmung und Gestaltung von Datensouveränität lassen sich zwei einander zunehmend annähernde und bereits jetzt teilweise überschneidende Sphären unterscheiden: erstens die Sphäre der bislang schon durch vergleichsweise klare und strikte Datenschutz-, Qualität- und Sicherheitsstandards gekennzeichneten Datennutzung in der medizinbezogenen Forschung und klinischen Praxis. Zweitens die Sphäre der zunehmend den Gesundheitssektor mitbestimmenden, allerdings sehr heterogenen Angebote des freien Marktes. Letztere reichen von Anwendungskonzepten, die nahe an der

ersten Sphäre und den mit ihr verbundenen Standards liegen, bis hin zu ersichtlich unseriösen, nicht auf nachhaltige Gesundheitsförderung angelegten Angeboten.

100) Big-Data-Entwicklungen lassen sich nicht aufhalten, sehr wohl aber gestalten. Da die Handlungsformen und Schutzmechanismen des traditionellen Datenschutzrechts für eine solche Gestaltung nicht ausreichen, gilt es ein verändertes, die Komplexität und Entwicklungsdynamik von Big Data stärker spiegelndes Gestaltungs- und Regelungsmodell zu erarbeiten. Dieses sollte Datensouveränität als informationelle Freiheitsgestaltung multidimensional und mit Blick auf unterschiedliche Akteursgruppen und Handlungskontexte reflektieren und dabei die zuvor skizzierten Verantwortungsmöglichkeiten und -zuschreibungen aufgreifen.

101) Unter den Bedingungen von Big Data ist es notwendig, sich von überholten Vorstellungen einer spezifischen, vorgegebenen Sensibilität bestimmter Daten und hierauf rekurrierender besonderer Schutzmechanismen zu lösen. Datenschutz kann nicht mehr statisch an bestimmten Daten und Datennutzungskategorien ansetzen, sondern muss sich auf ständige Rekombinationen und Rekontextualisierungen einstellen.

102) Ein auf Datensouveränität ausgerichtetes Gestaltungs- und Regelungsmodell nimmt dabei vor allem den Datengeber als entscheidend zu schützenden und zu achtenden Zweck in den Blick. Ziel ist es, über eine gleichermaßen kontextsensible wie falladäquate Regulierung und Institutionengestaltung diese Subjekte, aber auch die mit ihnen in Verbindung stehenden Organisationen, zu einem souveränen Umgang mit ihren Daten zu befähigen. Vereinfachende Pauschallösungen sollten aufgegeben werden zugunsten komplexerer, aber auch flexiblerer und problemadäquater, institutionell diversifizierter Kombinationsmodelle.

103) Die heterogene zweite Sphäre gilt es dabei nach folgender Grundregel zu gestalten: Je näher einzelne Anwendungen an die erste Sphäre heranreichen, desto mehr besteht ethisch und rechtlich die Aufgabe, ihre Gestaltung multiakteursbezogen in die Richtung der dort generell vorherrschenden Qualitäts-, Schutz- und Vertraulichkeitsstandards zu entwickeln.

Empfehlungen

104) Der Deutsche Ethikrat empfiehlt ein Gestaltungs- und Regelungskonzept, das sich am zentralen Ziel der Datensouveränität orientiert. Ein solches Konzept verlangt eine umfassende gesamtgesellschaftliche Anstrengung, die rechtliche wie außerrechtliche Elemente einbezieht, technische Weiterentwicklungen aufnimmt und deren grundrechtswahrende Verfügbarkeit für alle gesellschaftlichen Akteure gewährleistet.

105) Das vorgeschlagene Gestaltungs- und Regelungskonzept enthält konkrete Handlungsempfehlungen zu vier Themenbereichen, die darauf abzielen, erstens die Potenziale von Big Data zu erschließen, zweitens individuelle Freiheit und Privatheit zu wahren, drittens Gerechtigkeit und Solidarität zu sichern und viertens Verantwortung und Vertrauen zu fördern. Die empfohlenen Maßnahmen sollten zeitnah verwirklicht und finanziert werden.

106) Empfehlungen, um die Potenziale von Big Data in gesundheitsbezogenen Bereichen zu erschließen (Themenbereich A), berühren die folgenden Punkte:

- Infrastrukturelle Grundvoraussetzungen schaffen (A1)
- Datenaustausch und -integration erleichtern (A2)
  - Standardisierte Verfahren der Interoperabilität von Daten entwickeln und bereitstellen (A2.1)
  - Kooperatives Forschungsdatenmanagement weiterentwickeln (A2.2)
- Daten- und Forschungsqualität fördern und schützen (A3)
  - Epistemische Standards einhalten, insbesondere die der evidenzbasierten Medizin (A3.1)
  - Einheitliche Daten- und Dokumentationsstandards einführen (A3.2)
  - Datengütesiegel etablieren (A3.3)
- Rechtliche Rahmenbedingungen für die Datennutzung zu Forschungszwecken anpassen (A4)
  - Sekundärnutzung von Forschungsdaten weiterentwickeln (A4.1)
  - Rechtliche Möglichkeit für Individuen schaffen, die umfassende Nutzung ihrer Daten für die medizinische Forschung zu erlauben („Datenspende“) (A4.2)
- Digitale Entscheidungshilfesysteme in der klinischen Praxis fördern (A5)
- Internationale Anschlussfähigkeit fördern (A6)

107) Empfehlungen zur Sicherung individueller Freiheit und Privatheit (Themenbereich B), umfassen folgende Aspekte:

- Datenhoheit bewahren (B1)
  - Programmatische Schnittstellen für Datengeber öffnen („Datenagenten“) (B1.1)
  - Mitbestimmung bei der Datenweitergabe erleichtern (B1.2)
  - Rechtsprobleme eines vermeintlichen Eigentums an Daten klären (B1.3)
- Kaskadisch strukturierte Einwilligungsmodelle etablieren (B2)
- Privatsphärenfreundliche Grundeinstellungen gewährleisten (B3)
- Einsatz von Algorithmen transparent machen und erläutern (B4)
- Täuschung und Manipulation entgegenwirken (B5)

- Digitale Bildung fördern (B6)
- Diskurs und Teilhabe stärken (B7)

108) Um Gerechtigkeit und Solidarität auch unter Big-Data-Bedingungen zu sichern (Themenbereich C), empfiehlt der Deutsche Ethikrat folgendes:

- Fairen Zugang zu digitalen Angeboten schaffen (C1)
- Diskriminierung und Stigmatisierung aufdecken bzw. verhindern (C2)
- Widerspruch bei automatisierten Entscheidungen ermöglichen (C3)
- Vulnerable Gruppen und Individuen schützen (C4)
  - Einwilligungserfordernisse bei Kindern und Jugendlichen streng beachten (C4.1)
  - Schutzmechanismen für die Datenerhebung an sonstigen Personen mit eingeschränkter Einwilligungsfähigkeit entwickeln (C4.2)
  - Einsatz von Chatbots restriktiv regeln (C4.3)
- Zuwendungsorientierte Medizin gewährleisten (C5)
- Wirksame Haftung von Unternehmen, die im Gesundheitsbereich mit Daten arbeiten, sicherstellen (C6)

109) Folgende Empfehlungen sollen Verantwortung und Vertrauen beim gesundheitsbezogenen Einsatz von Big Data fördern (Themenbereich D):

- Schutz- und Qualitätsstandards garantieren (D1)
  - Bestmögliche Schutzstandards gegen unbefugte Identifizierung von Individuen aus anonymisierten, pseudonymisierten oder aggregierten Datensätzen etablieren (D1.1)
  - Anonymisierungsdefizite durch kontrollierten Zugang zu Daten kompensieren (D1.2)
  - Umsetzung von Schutzvorgaben gewährleisten und nachweisen (D1.3)
  - Informationspflicht bei Pannen und Fehlverhalten etablieren (D1.4)
- Kontrollmechanismen verbessern (D2)
  - Datenschutzbeauftragte stärken (D2.1)
  - Datenprüfer etablieren (D2.2)
  - Datentreuhandmodelle einführen (D2.3)
- Kodizes für Forschung, Klinik und Wirtschaft erarbeiten (D3)
- Gütesiegel für Anbieter und Anwendungen unterstützen und ausbauen (D4)
- Kompetenz im verantwortungsvollen Umgang mit Daten für alle, die professionell mit Big Data zu tun haben, stärken (D5)

1 Einleitung

Big Data gehört zu den Schlüsselbegriffen der gegenwärtigen Debatte über die technologisch induzierte gesellschaftliche Veränderung. Obwohl die als Big Data zusammengefassten Technologien oft selbst keiner festgelegten, eindeutigen Definition folgen, entfaltet der Begriff in der Öffentlichkeit große Wirkmacht. Hinter dem Stichwort Big Data verbirgt sich ein zentraler Mechanismus der Datenwelt: die Erfassung, Analyse und neue Verknüpfung wachsender Datenmengen auf Grundlage einer aufwendigen Infrastruktur. Die Menge der weltweit kursierenden Daten wächst stetig und rasant an. Durch Rekontextualisierungen und Rekombinationen lassen sich immer weitergehende Erkenntnisse aus den vorhandenen und neu erhobenen Daten gewinnen. Die mit Big Data verbundenen Prozesse gehen dabei über die bisherige, auf einen bestimmten Verwendungszweck konzentrierte Datenerfassung weit hinaus. Sie fordern deshalb unseren bisherigen Umgang mit Daten heraus und konfrontieren uns mit Fragen zu einer Reihe gesellschaftlicher Praktiken, wie die Gestaltung von Versicherungsverträgen, und zu angemessenen rechtlichen Standards, insbesondere den Datenschutz. Die Diskussion um Big Data betrifft im Kern unser Selbstverständnis als Einzelne sowie als Gesellschaft unter den Bedingungen der Digitalisierung und der zunehmenden Vernetzung.

Die transformativen Potenziale von Big Data für Individuen und Gesellschaft zeigen sich in besonders anschaulicher und eindrücklicher Weise im Gesundheitsbereich. Schon jetzt arbeiten immer mehr Forscher, Firmen und Ärzte mit riesigen Datenmengen. Gesundheitsdaten werden aber längst nicht mehr nur in Arztpraxen und Studien gesammelt, sondern auch von Bürgern selbst erfasst – etwa über die Sensoren und Apps von Mobiltelefonen und am Körper getragenen Geräten (Fitness-Tracker, Smartwatches). Gleichzeitig wachsen dank neuer Entwicklungen in der Datenwissenschaft und Dateninfrastruktur die Möglichkeiten, die so gewonnenen, vielfältigen Daten schnell und effektiv auszuwerten, auszutauschen und sie miteinander sowie mit anderen Daten zu verknüpfen, die gerade in der Zusammenschau ebenfalls gesundheitsrelevant werden können, zum Beispiel Informationen über das Einkaufsverhalten, Suchanfragen im Internet, Ortsdaten oder die Analyse von Text-, Sprach- und Videomaterial. Solche Analysen ermöglichen nicht nur tiefe Einblicke in den aktuellen Gesundheitszustand, die Persönlichkeit und den Lebenswandel, sondern erlauben mitunter sogar entsprechende Vorhersagen. Insgesamt fließt eine immer größere Zahl verschiedener Datenströme in Big-Data-basierte Auswertungshorizonte ein.

Die beschriebene Verknüpfung unterschiedlicher Datenarten verspricht neue Erkenntnisse für wissenschaftliche Forschung und medizinische Behandlungsstrategien. Die gemeinsame Auswertung von zum Beispiel klassischen medizinischen Daten, Forschungsdaten, Daten öffentlicher Gesundheitsversorgung, Bewegungsdaten, Fitnessdaten, Daten aus sozialen Netzwerken und Versicherungsdaten ermöglicht ganz andere Einblicke und Eingriffe beim Umgang mit

Gesundheitsrisiken und Krankheiten als bisher. Dass solche Visionen besserer, schnellerer und präziserer Diagnose, Prävention und Behandlung überhaupt möglich erscheinen, hängt jedoch auch mit der Entgrenzung von Gesundheitsfragen zusammen, in deren Zuge alles, was wir tun und erleben, nicht nur in Bezug auf unsere Gesundheit analysiert wird, sondern beispielsweise auch in die Berechnungen des Krankheits- oder Gesundheitsstatus einfließen kann. Dies gilt für medizinische Diagnosen, Therapie- und Präventionsoptionen ebenso wie für die Kalkulationen von Krankenversicherungsprämien oder von Berufsrisiken.

Gesundheitsdaten gelten herkömmlich als besonders sensible Daten und sind entsprechend geschützt. Hierfür gibt es mindestens zwei Gründe: Wo in einem sozialen Kontext (Beruf, Verträge, Beziehungen) als ungünstig erachtete Gesundheitszustände bekannt werden, kann dies für die Betroffenen zu Benachteiligungen – bis hin zu Diskriminierungen und Stigmatisierungen – mit der Folge sozialer und gegebenenfalls vertraglicher und finanzieller Exklusion führen. Zweitens und unabhängig von solchen Folgenabwägungen gilt die leibliche Konstitution der Person in unserer Gesellschaft als etwas zutiefst Intimes. Über sie muss bis auf im Einzelnen zu begründende Ausnahmefälle keine Rechenschaft gegenüber anderen abgelegt werden. Ob bestimmte Daten als sensibel oder gesundheitsrelevant zu betrachten sind, lässt sich jedoch angesichts der neuen Kombinations- und Auswertungsmöglichkeiten von Big-Data-Anwendungen oft nicht mehr bei der Erhebung bestimmen, sondern hängt zunehmend vom Kontext ab, in dem die Daten verwendet werden.

Auch die massive Einflussnahme auf diesen Sektor durch große IT- und Internetfirmen, die ihren Firmensitz außerhalb von Europa haben und derzeit noch immer, auch innerhalb Europas, schwer zu kontrollieren sind, verschärft die Herausforderung für den Einzelnen sowie für die Gesellschaft. Zentrale Prinzipien wie Privatheit und informationelle Selbstbestimmung erscheinen nicht nur gegen Eingriffe von außen zunehmend schwer zu verteidigen zu sein, sondern auch wegen des Verhaltens der Betroffenen selbst, die für die Servicevorteile vieler Apps, Programme und internetbasierter Technologien Eingriffe in den intimsten Persönlichkeitsbereich billigend in Kauf nehmen. Damit deutet sich ein weiterer Problembereich an: Individualisierung, die bei medizinischer Prävention und Behandlung möglicherweise segensreich ist, kann im Versicherungswesen und im sozialen Miteinander bewusste oder unbewusste Dynamiken der Entsolidarisierung hervorrufen. Was der Prämienvorteil für den einen wäre, könnte letztlich einen finanziellen Nachteil für den anderen bedeuten. Wer verbreiteten Verhaltensratschlägen zukünftig nicht folgt, könnte Gefahr laufen, Versicherungsschutz nur zu schlechteren Konditionen oder gar nicht zu erhalten.

Durch Big Data im Gesundheitsbereich eröffnen sich also einerseits vielversprechende neue Perspektiven und Chancen für die Erforschung, Prävention, Diagnose und Behandlung von Krankheiten. Andererseits sind aber auch ernst zu nehmende Herausforderungen und Risiken

für verschiedene gesellschaftliche Bereiche zu erkennen. Angesichts der starken Veränderungen, die diese Chancen und Risiken sowohl für den Einzelnen als auch die Gesellschaft bewirken könnten, stellen sich für den Einsatz von Big Data im Umgang mit Gesundheit zahlreiche Fragen: Wie berührt zum Beispiel die immer engmaschigere und oft kaum merkbare Sammlung gesundheitsrelevanter Daten unsere Selbstwahrnehmung, Freiheit und Selbstbestimmung? Entpuppt sich das, was unter den Bedingungen von Big Data Selbstbestimmung zu sein scheint, am Ende möglicherweise als Selbstentmündigung, oder handelt es sich dabei nur um eine jener Veränderungen des eigenen Selbstverständnisses, wie sie im Laufe eines Lebens durchaus öfter stattfinden? Wie können Privatpersonen, Forscher und Firmen verantwortungsbewusst mit Big Data umgehen, und wie lässt sich die Qualität und Zuverlässigkeit komplexer Datenauswertungen sichern? Welche Herausforderungen stellen sich für Solidarität und Gerechtigkeit, zum Beispiel mit Blick auf Krankenversicherungen, aber auch mit Blick auf die Forschung, die für Big-Data-Anwendungen viele Daten benötigt? Wie kann das Vertrauen in die Forschung gewahrt bleiben, wenn auch anonymisierte Probandendaten zumindest im Prinzip immer wieder entschlüsselt werden können? Welche regulatorischen Mechanismen und Anreize bieten sich, um die Chancen und Risiken von Big Data im Gesundheitsbereich angemessen zu handhaben? Und wo sollen aus der Sicht des Individuums oder des regulierenden Staates Grenzen für die Erhebung, Verknüpfung und Nutzung von Daten gezogen werden? Sollten wir überhaupt Grenzen ziehen oder stattdessen die bisherigen Grenzen, wie gegenwärtig gültige Standards des Datenschutzes, lockern, um die gewünschten Effekte von Big Data noch besser nutzen zu können?

Angesichts solcher und anderer Fragen und Herausforderungen gilt es, zumindest aufmerksam wahrzunehmen, wie weit Big Data Einfluss auf das Verständnis von Gesundheit, Selbstbestimmung und sozialem Miteinander hat. Ziel muss es sein, diese Entwicklung so zu gestalten, dass sie nicht nur – was zunächst durchaus legitim ist – zu Gewinnen diverser beteiligter Unternehmen führt, sondern auch die reale Freiheitsgestaltung der Menschen in der Gesellschaft sichert und fördert. Die vorliegende Stellungnahme zielt vor diesem Hintergrund darauf, Sensibilität für das drängende Thema Gesundheit und Big Data zu wecken und zugleich ethische und rechtliche Standards auf diesem Felde zu setzen.

Sie skizziert im Folgenden zunächst die wissenschaftlich-technischen wie auch ökonomischen Grundlagen von Big Data im gesundheitsrelevanten Bereich (Kapitel 2). Es schließt sich eine Darstellung des rechtlichen Rahmens und der dabei identifizierten Regelungslücken an (Kapitel 3). Anschließend werden ethische Grundfragen behandelt (Kapitel 4). Die Antworten darauf münden in die Forderung, mittels eines den neuen Big-Data-Bedingungen angemessenen Gestaltungs- und Regelungskonzepts das zentrale Ziel der Datensouveränität zu erreichen (Kapitel 5). Dieser Begriff ist im Zusammenhang mit Big Data in unterschiedlichen Bedeutungen geläufig; hier wird er verstanden als eine den Chancen und Risiken von Big Data angemessene ver-

antwortliche informationelle Freiheitsgestaltung. Dieses Konzept nimmt damit basale rechtliche und ethische Überlegungen auf und entwickelt sie bereichsspezifisch weiter. Datensouveränität umzusetzen, ist eine komplexe Aufgabe, in der eine Vielzahl von Akteuren und unterschiedliche Interessen berücksichtigt werden müssen. Die Grundlagen, die jetzt hierfür technisch, ökonomisch, rechtlich, politisch und kulturell geschaffen werden, entscheiden langfristig über die gesellschaftliche Bedeutung von Gesundheit und Selbstbestimmung. In diesem Sinne formuliert der Deutsche Ethikrat abschließend Empfehlungen zum Umgang mit den erkannten ethischen, rechtlichen und sozialen Herausforderungen, die sich durch Big Data in gesundheitsbezogenen Bereichen ergeben (Kapitel 6).

2 Grundlagen: Big Data und Gesundheit

Die systematische Erhebung und Auswertung von Daten ist spätestens seit Beginn der Neuzeit ein bedeutender Faktor zivilisatorischer Entwicklung. Systematische Experimente, Messungen und Beobachtungen bilden das Fundament der empirischen Wissenschaften. Die Wissenschaftsgeschichte ist geprägt von Erfindungen neuer Instrumente und verbesserter Verfahren, die zur Vermessung unserer Welt herangezogen werden können. Ohne diese Erfindungen ist weder eine erfolgreiche Theoriebildung noch ein anhaltender wissenschaftlicher, technologischer und wirtschaftlicher Fortschritt denkbar. Datenerhebung und Informationssammlung spielen aber auch eine wichtige Rolle bei der Bewältigung von Problemen gesellschaftlicher Selbstorganisation, wie sie sich auch im Gesundheitssystem stellen.

Die Vermessung der Welt schließt auch den Menschen und seine Lebensumgebung ein. Dies reicht im medizinischen Kontext etwa von der Bestimmung von Körpermaßen wie Puls, Temperatur, Atmung oder Blutdruck bis hin zu moderner Labordiagnostik, bildgebenden Verfahren und Genomanalysen. Ebenso werden Persönlichkeitseigenschaften, emotionale Zustände und kognitive Fähigkeiten oder Motivationen zum Gegenstand einer empirischen Forschung, die darauf ausgerichtet ist, menschliches Verhalten zu erklären, besser vorhersagbar zu machen oder sogar zu beeinflussen. Die Psychometrie etwa verfolgt das Ziel der Messung psychischer Phänomene und versucht dabei, (experimentell) messbare Daten auf vermutlich zugrunde liegende Faktoren zurückzuführen. Die Epidemiologie und die Sozialwissenschaften weiten die empirische Methode auf Gruppen bzw. soziale Kollektive aus, wobei oft auf vorhandene Daten aus Registern, Archiven und klinischen Datenbanken zurückgegriffen wird, Daten aber auch durch Befragung oder Beobachtung von Probanden neu erhoben werden.

Der Computer trat zunächst als Instrument für mathematische Rechenoperationen und später der elektronischen Datenverarbeitung (EDV) in Erscheinung. In den Wissenschaften erlaubte der Einsatz von Computern eine Steigerung des handhabbaren Datenvolumens, aber auch vielfältige qualitative Verbesserungen, wie die Verwendung komplexerer Rechenvorschriften (Modelle) in rechenintensiven Computersimulationen. Hinzu kam die mit der Einführung des Computers vorangetriebene Revolution von Geschäftsprozessen. EDV erschöpfte sich zunächst oftmals in der manuellen Erfassung von Daten sowie ihrer Aufbewahrung, Pflege und Aggregation. Kennzeichen der klassischen EDV waren – und sind noch immer – Arbeitsprozesse und eine Geschäftslogik, die menschliche Tätigkeit mit begrenzt automatisierten Schritten verschränkt, in der Regel mit spezieller Software (zum Beispiel zur Buchhaltung, Inventarisierung, Kundenverwaltung usw.). Dies gilt auch im Gesundheitswesen, etwa beim Einsatz von Informationstechnik (IT) in Krankenhäusern, in der ärztlichen Praxis oder bei Versicherungen. Daten sind in diesem Zusammenhang fast immer strukturiert, das heißt interpretierbar in einem

Datenbankschema, das Messprotokolle, Datenformate und Datensemantik festlegt. Rationalisierung, Standardisierung und Qualitätssteigerung sind die Hauptziele dieser technologischen Entwicklung.

Seit den 1990ern werden Entscheidungsprozesse immer stärker durch systematische Analyseverfahren und hierfür optimierte zentrale Datenlager (*data warehouses*) unterstützt. Mit ihrer Hilfe werden unterschiedliche Daten aus ihren ursprünglichen Quellen und Verarbeitungsprozessen herausgelöst, im großen Maßstab zusammengeführt und verdichtet, und schließlich einer Datenanalyse zugänglich gemacht. Dabei kommen zunehmend komplexere Algorithmen[3] zum Einsatz.

2.1 Charakteristika von Big Data

Die Entwicklung hin zu Big Data geht mit einer Transformation aller Phasen der Datenverarbeitung einher, die von zunehmender Automatisierung, Vernetzung und Durchdringung geprägt ist. Den Begriff Big Data definitorisch exakt zu bestimmen, ist schwierig; die Unschärfe des Begriffs gilt mitunter sogar als ein charakteristischer Aspekt des Phänomens.[4] Big Data bezieht sich jedenfalls nicht nur auf die reine Quantität der verarbeiteten Datenmengen, sondern auch auf die damit verbundenen qualitativen Veränderungen bei den Anforderungen an die Datenauswertung und deren Potenziale. Der Deutsche Ethikrat legt folgende Arbeitsdefinition zugrunde:

Big Data ist der Umgang mit großen Datenmengen, der darauf abzielt, Muster zu erkennen und daraus neue Einsichten zu gewinnen, und der hierzu angesichts der Fülle und Vielfalt der Daten sowie der Geschwindigkeit, mit der sie erfasst, analysiert und neu verknüpft werden, innovative, kontinuierlich weiterentwickelte informationstechnologische Ansätze nutzt.

Diese Definition enthält drei Schlüsselbegriffe, die in vielen Begriffsbestimmungen von Big Data vorkommen[5] und aufgrund ihrer Alliteration im Englischen als „die drei Vs" bezeichnet werden: *volume* (Volumen bzw. Datenmenge), *variety* (Vielfalt) und *velocity* (Geschwindigkeit). Als weitere Vs werden häufig *validity* (Sicherstellung der Datenqualität, auch *veracity* im Sinne von Glaubwürdigkeit) sowie *value* (unternehmerischer Wert) als wichtige Zielgrößen von Big Data genannt.[6] Von besonderer Bedeutung für den hier behandelten, ethisch besonders sensib-

[3] Algorithmen sind eindeutige mathematische Handlungsvorschriften und Vorgehensweisen, die aus definierten Schritten bestehen und zu einem bestimmten Ziel führen; sie stellen die Grundlage von Computerprogrammen dar.
[4] Vgl. Mayer-Schönberger/Cukier 2013, 6.
[5] Vgl. zum Beispiel die Definition im Gartner-IT-Glossar: „Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation." (http://www.gartner.com/it-glossary/Big-Data [17.10.2017]). Siehe auch Laney 2001.
[6] Siehe beispielsweise Amma 2016, 54. Insgesamt lassen sich jedoch viele weitere Vs in der Debatte identifizieren. Vgl. hier zum Beispiel Cartledge 2017.

len Gesundheitsbereich ist die durch Big-Data-Technologien ermöglichte umfassende Dekontextualisierung und Rekontextualisierung von Daten, die zu unterschiedlichen Zwecken erfasst, analysiert und neu verknüpft werden (siehe dazu Abschnitt 2.4). Das Konzept von Big Data unterliegt aufgrund der stetig zunehmenden Datenmengen und Verknüpfungsoptionen sowie der dynamischen Weiterentwicklung der technischen Möglichkeiten einem ständigen Wandel, beispielsweise im Bereich des maschinellen Lernens. Die in der obigen Arbeitsdefinition enthaltenen Elemente dürften aber begriffsbestimmend bleiben.

2.2 Erhebung und Handhabung großer Datenmengen

Infolge des technischen Fortschritts sind immer mehr Daten aus immer mehr unterschiedlichen Quellen immer schneller verfügbar. Dies wird wesentlich durch zunehmende Automatisierung ermöglicht. Der sinnbildliche, vor einer Eingabemaske sitzende EDV-Experte, der sich von Eingabefeld zu Eingabefeld hangelt, ist in der Big-Data-Welt verschwunden. Vielfältige wissenschaftlich-technologische Errungenschaften in Feldern von den Grundlagentheorien in der Physik über die Materialwissenschaft bis hin zur Nanotechnologie und verschiedenste neue Fertigungstechniken fließen in die Entwicklung von neuartigen Instrumenten, Messapparaturen und -verfahren ein. Diese haben zumindest eine Gemeinsamkeit, nämlich in vormals unbekannte Größenordnungen der Datenerfassung vorzustoßen. In kürzester Zeit erhobene Datensätze von einigen Petabyte[7] sind keine Seltenheit mehr und der Trend des Datenwachstums scheint ungebrochen.[8] Volumen und Tempo der voll automatisierten Datenerfassung sind in wenigen Jahren um viele Größenordnungen gestiegen und scheinen nun weit jenseits der Vorstellungskraft des Menschen zu liegen. So gehen Schätzungen davon aus, dass das globale Datenvolumen 2016 bei etwa 16,1 Zettabyte (entspricht 16,1 Billionen Gigabyte) liegt und mit einer jährlichen Wachstumsrate von 30 Prozent auf etwa 163 Zettabyte im Jahr 2025 ansteigen wird.[9] Davon sind bis zu 80 Prozent sogenannte unstrukturierte Daten, wie zum Beispiel Videos, Tonaufnahmen, E-Mails, Recherchedaten und Beiträge aus sozialen Medien.[10]

Das enorme Anwachsen der Datenbestände bringt es mit sich, dass der Mensch oft die Qualität der Daten nicht mehr direkt kontrollieren kann. Dies gilt beispielsweise für Messfehler, Ausreißer und andere irreguläre Daten in großen Datensätzen, etwa Datenbanken in Krankenhäusern, die nach neuen Analysemethoden und automatisierten Verfahren der Qualitätssicherung verlangen. Andererseits ist es mitunter auch erforderlich, den Menschen in den Prozess der Datengewinnung oder Analyse zu integrieren – man bedient sich dann des sogenannten Crowdsourcing, bei dem man wohldefinierte Fragen, Tests oder einfache Analyseaufgaben gezielt an

[7] Ein Petabyte entspricht 10^{15} Byte oder einer Million Gigabyte.
[8] Siehe etwa Cisco 2017.
[9] Vgl. Reinsel/Gantz/Rydning 2017, 3.
[10] Vgl. Stone 2014, 2.

menschliche „Experten“ weiterleitet. Man bezeichnet diese Umkehrung des Verhältnisses zwischen Mensch und Maschine, bei der der Mensch zum Helfer eines Computerprogramms wird, auch als *human computation*.

Neben dem zunehmenden Einsatz von leistungsfähigen Großgeräten in Medizin und Forschung trägt zu den großen Datenmengen auch der Umstand bei, dass die menschliche Lebenswelt in wachsendem Maß durch digitale Technologien geprägt wird. Die rasche Verbreitung und Vernetzung von Geräten, die in der Haupt- oder Nebensache zur Datenerhebung dienen oder dazu genutzt werden können, eröffnet ständig neue Datenquellen. Hierzu gehören neben der Datensammlung mithilfe des kompletten Spektrums an Computern (PC, Notebook, Spielkonsole, Tablet, Smartphone) auch die Erfassung von Körperfunktionen und Lebensgewohnheiten mittels tragbarer Geräte (Fitness-Tracker, Smartwatches) und vernetzter Haushaltsgeräte, die Überwachung des öffentlichen Raums zum Beispiel durch Kameras sowie diverse Sensoren zur Instrumentierung von Maschinen (vom Auto bis zur Industrieanlage), die Messung von Umweltgrößen (durch Verfahren der Fernerkundung), die Bewegung von Waren[11] und dergleichen mehr. Das Smartphone allein kann man mit Recht als eines der vielseitigsten und am weitesten verbreiteten Messinstrumente der Menschheitsgeschichte ansehen. Täglich zeichnen Smartphones Bewegungsdaten (GPS, WPS), Web- und App-Nutzung, Suchanfragen (Google, Siri), Kommunikation (Anruf, SMS, Messenger, E-Mail, soziale Netzwerke), Transaktionen (Käufe, Buchungen) und vieles mehr milliardenfach auf. Dies alles sind Daten mit einem klaren Personenbezug. Sie erlauben also Rückschlüsse auf die Person, von der die Daten stammen. Aus Nutzersicht wird es wohl meist so sein, dass viele dieser Daten nur als Neben- oder Abfallprodukt anfallen und eine Datenspeicherung oder -weitergabe weder beabsichtigt ist, noch überhaupt in den Blick kommt.

Sind Daten einmal erhoben, sorgen Datennetzwerke und vernetzte Softwaresysteme für ihren Austausch und ihre Verknüpfung, häufig in Echtzeit und auf skalierbare Weise, und oft in einem transnationalen oder gar globalen Umfang. Dabei gibt es weitverbreitete Standards des Austauschs von Daten und Nachrichten über das Internet, meistens mittels Schnittstellen zur Anwendungsprogrammierung (*programming interface*, API), die auf dem Hypertext Transfer Protocol (HTTP) aufbauen. Die Standards erlauben den programmatischen Austausch und die Verknüpfung von Daten zwischen verschiedenen Computersystemen, auch unter Beteiligung mehrerer Parteien. Ein wesentlicher Aspekt des programmatischen Datenaustauschs liegt in

[11] Warenbewegungen lassen sich mithilfe von RFID-Systemen (*radio-frequency identification systems*) verfolgen, die die Identifizierung und Lokalisierung von Geräten über elektromagnetische Wellen ermöglichen. Waren, die mit RFID-Sendern (Transponder) ausgestattet sind, können kontaktfrei Daten mit RFID-Lesegeräten (Reader) austauschen, die sich innerhalb eines Empfangsbereiches befinden. Eine besonders einfache Anwendung dieser Technologie ist etwa die Diebstahlsicherung in Kaufhäusern.

der Stetigkeit der Verarbeitung: Sind die Schnittstellen implementiert und die Systeme wechselseitig autorisiert, entstehen Datenverbindungen, über die Daten in kürzester Zeit wie durch ein komplexes Röhrensystem fließen, ohne dass weitere menschliche Aktivität erforderlich ist.

Programmatische Schnittstellen und Blockchain: technisch unterstützte Transparenz und Kontrollmöglichkeiten bei der Datenverarbeitung

Der automatisierte Datenaustausch über die hier beschriebene Dateninfrastruktur steht nicht im Widerspruch zur Gewährleistung einer andauernden individuellen Bestimmungsmacht über Daten. Denn dieselben technischen Instrumentarien, insbesondere die Nutzung von programmatischen Schnittstellen und die automatische Protokollierung der Herkunft von Daten, eröffnen ein weitgehend ungenutztes Potenzial: Der Datengeber kann durch geeignete Software-Werkzeuge unterstützt werden, die die zur Verfügung gestellten Daten fortdauernd nach seinen Vorstellungen verwalten (etwa bezüglich Speicherung, Löschung, Verrauschung, Anonymisierung). Jene Werkzeuge agieren quasi als Stellvertreter oder Agenten des Datengebers, indem sie individuelle oder von Interessenvertretern gestaltete Regeln und Vorgaben zum Umgang mit Daten umsetzen. Da Datenverarbeiter und -nutzer heute vielfach auf ebensolche Modelle zum Datenaustausch zurückgreifen, ließe sich so mit begrenztem technischen Aufwand ein Zugewinn an Kontrolle, Transparenz und Nachvollziehbarkeit erzielen. Die Blockchain-Technologie beispielsweise, die vor allem im Zusammenhang mit der Digitalwährung Bitcoin[12] bekannt wurde, ermöglicht den dezentral organisierten Austausch von Informationen zwischen Parteien bei gleichzeitiger transparenter und manipulationssicherer Dokumentation aller Transaktionsschritte. In einer Blockchain (engl. für Blockkette) werden Transaktionen als Datenblöcke mithilfe kryptografischer Verfahren zu einer ständig erweiterbaren Liste verkettet, die wie eine dezentrale Buchführung funktioniert. Jeder Block enthält dabei neben einem Zeitstempel und den Daten der Transaktion einen kryptografisch sicheren Schlüssel zu dem vorhergehenden Block. Da jeder Block auf früheren Blöcken aufbaut, können zuvor aufgezeichnete Inhalte und Transaktionen nicht im Nachhinein manipuliert werden. Die Blockchain-Technologie erlaubt somit auch für den Gesundheitsbereich eine besonders sichere Dokumentation von Schritten der Datenverarbeitung und -weitergabe sowie von relevanten Metadaten, etwa zur Datenherkunft und -qualität oder zur Einwilligung in bestimmte Datennutzungen. Die Alphabet-Tochter DeepMind plant beispielsweise derzeit in Großbritannien den Einsatz von Blockchain-basierten Verfahren zur Gewährleistung der Integrität von (Meta-)Daten, mit dem Ziel, eine größere Transparenz in der Datennutzung sicherzustellen.[13]

[12] Vgl. Nakamoto 2009.
[13] Siehe https://deepmind.com/blog/trust-confidence-verifiable-data-audit [17.10.2017].

Die effiziente Erfassung, Speicherung und Verarbeitung von Daten benötigt eine leistungsfähige Rechenmaschinerie, die meist in Datenzentren bereitgestellt wird. Große Datenmengen im Multi-Petabyte-Bereich stellen hohe Anforderungen an die Kapazitäten der Datenspeicher, des Netzwerks, der Schnittstellen und der zugrunde liegenden Infrastruktur. In den letzten Jahren sind die technischen Voraussetzungen geschaffen worden, den weltweiten Datenaustausch schnell und effizient umzusetzen. Schnelle parallele Datenverarbeitungsmethoden durch Rechen-Cluster und Supercomputer ermöglichen es heute, große Datenmengen in geringer Zeit zu analysieren. Im kommerziellen und wissenschaftlichen Bereich bieten spezielle technologische Lösungen, wie MapReduce durch Apache Hadoop oder Apache Spark (siehe unten), die Möglichkeit, einfach parallelisierbare Algorithmen auf hohen Datendurchsatz (*high throughput computing*) anzuwenden. Parallelisierbare Algorithmen erlauben die Abarbeitung vieler Handlungsschritte zur gleichen Zeit – das führt zu einer deutlichen Beschleunigung in der Ausführung der Programme.

Aufgrund der großen globalen Nachfrage nach Big-Data-fähiger Rechenkapazität und des daraus resultierenden Kostendrucks ist es zu einer weitgehenden Kommerzialisierung von Dateninfrastrukturen gekommen. Die Systemarchitektur eines Datencenters bezieht heute üblicherweise eine Vielzahl von kostengünstigen Servern ein, die erst lokal und dann weiter mit anderen Server-Gruppen vernetzt werden. Die Daten werden zumeist über Server verteilt und dann unter Minimierung des Datentransfers verarbeitet. Dies ist eine Konsequenz der technischen Charakteristika heutiger Hardware, mit der Rechenleistung oft billiger und daher leichter verfügbar ist als Netzwerkkapazität oder das Einlesen von riesigen Datensätzen. Daten zu bewegen, ist oft teurer, als Daten lokal zu verarbeiten. Die für Big-Data-Anwendungen charakteristische verteilte, sich über viele miteinander vernetzte Computer erstreckende Berechnung ist hochkomplex, beispielsweise in Bezug auf Fehlertoleranz und Datensicherheit. Sie erfordert die Entwicklung einer angemessen flexiblen Softwareinfrastruktur, die über klassische Datenbanktechnologien hinausgeht. Viele der erforderlichen Techniken und Komponenten wurden ursprünglich von Internetfirmen als interne Werkzeuge entwickelt.[14] Seit etwa 2005 erfolgt solche Softwareentwicklung zunehmend auch außerhalb der großen Internetunternehmen im öffentlichen oder kommerziellen Raum. Maßgeblichen Einfluss hat hier die Apache Foundation mit ihren vielen Open-Source-Projekten (zum Beispiel Hadoop, Cassandra, Spark oder Kafka – insgesamt über 120 Millionen Zeilen Programmcode). Daneben gibt es auch Angebote einer Vielzahl spezialisierter Start-ups. Durch die Verfügbarkeit solcher neuen Big-Data-Technologien sind die technischen Hürden für die Entwicklung von entsprechenden Dienstleistungen signifikant

[14] Das Programmiergerüst (Framework) Apache Hadoop wurde beispielsweise als Open-Source-Implementierung von Googles Programmiermodell MapReduce entwickelt, und das Datenbankverwaltungssystem Apache Cassandra entstand ursprünglich bei Facebook.

gesunken, was sich auch als ökonomischer Anreiz auswirkt.[15] Mittlerweile ist erkennbar, wie diese Verfahren nicht nur innerhalb des Technologiesektors zum Einsatz kommen, sondern auch im Bereich der Wissenschaft und im öffentlichen Sektor inklusive des Gesundheitswesens.[16]

Wegen der zunehmenden Standardisierung von Hardware, Software und Protokollen kann man die erforderlichen Rechenleistungen weitgehend als abstrakte Dienstleistung anbieten. Die Verlagerung von vor Ort administrierten Rechnern in die wenig greifbare Virtualität der Datenzentren wird als Cloud-Computing bezeichnet. Cloud-Computing ist inzwischen zu einem ökonomisch hart umkämpften und sich sehr dynamisch entwickelnden Markt geworden, der 2017 insbesondere von Amazon mit seinen Web Services dominiert wird, in den aber auch andere große IT- und Internetfirmen wie Microsoft, Google oder IBM drängen.[17] Die Speicherung und Verwaltung der Daten wird dabei ein zunehmend gewichtiger Kostenfaktor.

Auf der Nutzerseite sind es vor allem Start-ups und kleinere Firmen, die von diesen Angeboten profitieren, insofern sie Rechenleistungen nach Bedarf abrechnen können. Aber auch öffentliche Einrichtungen wie Universitäten setzen für ihre wissenschaftlichen Anwendungen häufig auf solche Angebote. Die Cloud-Anbieter sind ihrerseits bestrebt, ihren Anteil an der Wertschöpfung durch eine wachsende Zahl an zunehmend höherwertigen Diensten zu steigern. Statt selbst eine Datenbank auf einem selbst betriebenen Rechner zu verwalten, verlagern Anwender oft nicht nur die Rechneradministration in die Cloud, sondern auch die Administration der Datenbanken. Brauchen sie mehr Speicherkapazität oder Datendurchsatz, so ist beides einfach per Mausklick bei dem entsprechenden Anbieter erhältlich. Daneben ist der Aufbau von Cloud-Lösungen auch im öffentlichen Bereich auf nationaler und europäischer Ebene im Gang, in der EU etwa im Rahmen der Strategie zur Hochleistungsrechentechnik (*high performance computing*).[18]

2.3 Datenanalyse und Datenwissenschaft

Welche ethischen Herausforderungen sich aus dem Einsatz von Big Data im Gesundheitsbereich ergeben, hängt wesentlich davon ab, mit welchen Analysemethoden Daten ausgewertet und welche Aussagen und Entscheidungen auf der Grundlage solcher Auswertungen getroffen werden. Relevante Aspekte berühren sowohl klassische Methoden der Datenanalyse und ihrer

[15] Schätzungen zufolge lag der Mehrwert, der allein von der Apache Foundation geschaffen wurde, bis 2012 bei ca. zehn Milliarden US-Dollar (vgl. Greenstein/Nagle 2014, 624). Diese Zahl dürfte heute um ein Vielfaches größer sein.
[16] So setzt etwa das von öffentlichen, vor allem deutschen Forschungseinrichtungen getragene Surveillance and Outbreak Response Management System (SORMAS) auf eine Kombination aus einer In-Memory-Datenbank und SAP-Cloud-Tools, um eine Art Frühwarnsystem für die Gefahr von Ebola-Epidemien in Westafrika zu etablieren (vgl. Fähnrich et al. 2015).
[17] Vgl. Leong et al. 2017.
[18] Siehe Europäische Kommission 2012a.

Transformationen im Zeitalter von Big Data als auch neue Techniken aus dem Bereich des maschinellen Lernens und der sogenannten künstlichen Intelligenz. Data-Science bezeichnet in diesem Zusammenhang einerseits ein neues Wissenschaftsparadigma der Wissensgewinnung *aus* Daten, das die wissenschaftliche Praxis tiefgreifend zu verändern beginnt, andererseits aber auch eine neue Wissenschaft *von* den Daten, die methodische Verfahren der Statistik und Informatik kombiniert, Daten zusammenführt, organisiert, annotiert und kollaborativen Analysen zugänglich macht.[19] Angesichts des erheblichen Anwachsens der Datenerhebung kann man inzwischen von einem neuen Typus der Datenwissenschaft sprechen, der sogenannten E-Science (*electronic science* oder *enhanced science*). E-Science bedeutet, dass moderne Computertechnologien dazu beitragen, wissenschaftliche Experimente vorzubereiten, durchzuführen, Daten zu generieren, auszutauschen, zu verteilen und langfristig aufzubewahren.[20] Die neuen Möglichkeiten der Datenverarbeitung befeuern die Datenerhebung und die Hoffnung, aus diesen Daten neue Einsichten und Erkenntnisse zu gewinnen. Wissenschaftliche Theoriebildungen werden angesichts der großen Datenmenge und -vielfalt dabei zunehmend durch automatisierte Analysen, basierend auf klassischen statistischen Verfahren, aber auch durch maschinelles Lernen ergänzt (siehe dazu Abschnitt 2.3.2).

2.3.1 Statistische Modellierung und Validierung von Zusammenhängen und Wirkmechanismen

Wesentlich für die Beurteilung von datenbasierten Aussagen, Vorhersagen oder Schlussfolgerungen sind die Objektivität, Reliabilität und Validität der verwendeten Daten und Analyseverfahren. Von diesen hängt die Berechtigung expliziter oder impliziter Wahrheitsansprüche ebenso ab wie die Angemessenheit der sich daraus ergebenden Handlungsempfehlungen. Testtheorie und Statistik als Lehre vom Umgang mit Daten sind wichtige Hilfsmittel in vielen Wissenschaften einschließlich der medizinischen Forschung. Der Bedarf an entsprechender methodischer Kompetenz ist in Zeiten von Big Data angesichts gesteigerter Komplexität und Heterogenität sowie zunehmender Möglichkeiten der Verknüpfung von Daten offenkundig deutlich größer als zuvor.

Ein großer Teil heutiger Big-Data-Anwendungen zielt darauf ab, Muster von potenziellen Wirkmechanismen zu identifizieren, auf deren Grundlage empirische Zusammenhänge erklärt werden können. Derartige Zusammenhänge oder Assoziationen, etwa zwischen dem Auftreten bestimmter Symptome und dem Vorliegen einer Krankheit, werden gemeinhin auch als Korrelationen bezeichnet. Dabei stellt die Entdeckung und Quantifizierung von Zusammenhängen nur einen ersten Schritt dar, auf dessen Grundlage Annahmen über kausale Relationen innerhalb der ermittelten Zusammenhänge möglich werden. Eine besondere Stärke von Big-Data-

[19] Vgl. etwa Kitchin 2014 oder Dijck 2016.
[20] Vgl. auch Hey/Tansley/Tolle 2009.

Anwendungen liegt darin, diese Annahmen durch weitere Analysen wiederum empirisch zu überprüfen. Sofern (mutmaßlich) kausal wirkende Variablen beeinflusst werden können, bilden die Ableitung von Interventionsmöglichkeiten und die Einschätzung möglicher Interventionseffekte eine weitere wichtige Anwendung von Big Data (siehe unten). Oft ist das Interesse an der Erhebung von Daten auch von dem Ziel bestimmt, Aussagen über Größen zu treffen, die als solche nicht direkt beobachtbar sind, sei es aus prinzipiellen oder aus pragmatischen Gründen. Der erste Fall betrifft beispielsweise Aussagen über Zukünftiges, der zweite Fall Situationen, in denen Messungen mit hohem Aufwand, hohen Risiken oder Nebenwirkungen verbunden sind oder sich aus ethischen bzw. rechtlichen Gründen verbieten.

Die Anwendung von Statistik beschränkt sich nicht auf die Identifikation empirischer Zusammenhänge zwischen einzelnen Merkmalen, von denen auf entsprechende Zusammenhänge in einer zugehörigen Grundgesamtheit geschlossen wird. Eine wesentliche Aufgabe besteht darin, Zusammenhänge auf der Grundlage komplexerer Modelle zu beschreiben und damit potenzielle Wirkmechanismen abzubilden. Vor dem Hintergrund derartiger Modelle ist es nunmehr möglich, Zusammenhänge zwischen Variablen nicht nur zu quantifizieren, sondern auch in ihrer Wirkrichtung und mit Blick auf mögliche kausale Einflussfaktoren näher zu beschreiben. Dabei steigen mit der Größe der Stichprobe die Möglichkeiten, zusätzliche Variablen, die empirische Zusammenhänge bedingen oder beeinflussen können, zu berücksichtigen. Da die Aussagekraft statistischer Tests direkt von der Größe der Stichprobe abhängt, für die Daten erhoben wurden, darf sie aber nicht mit praktischer Bedeutsamkeit eventuell entdeckter statistischer Zusammenhänge verwechselt werden. Anders als die Korrelation ist die statistische Signifikanz kein Maß für die Stärke eines Zusammenhangs. Inwieweit sich eine resultierende statistische Modellierung von Zusammenhängen für die Vorhersage oder die Planung effektiver Interventionen als nützlich erweist, hängt darüber hinaus von der Anzahl der zusätzlichen Variablen ab, die in die Modellierung von Zusammenhangsmustern einbezogen werden können. Weitere Voraussetzungen sind auch die Repräsentativität der Stichprobe, in der die Daten erhoben wurden, für die infrage stehende Grundgesamtheit und die Güte der erhobenen Daten. Nur unter folgenden Voraussetzungen können sich also statistische Modellierungen auch in praktischen Anwendungskontexten bewähren: erstens, wenn die interessierenden Daten objektiv gemessen werden können, also unterschiedliche Beurteiler auf gleicher Informationsgrundlage zu gleichen Einschätzungen kommen; zweitens, wenn sie hinreichend reliabel sind, also bei Messwiederholung sich vergleichbare Werte ergeben oder alternative Indikatoren zu vergleichbaren Einschätzungen von Merkmalsausprägungen kommen; und drittens, wenn die Validität der Daten vorausgesetzt werden kann, also zumindest ein wesentlicher Aspekt des interessierenden Merkmals gemessen wird.

Big-Data-basierte Analysen eröffnen Chancen in doppelter Hinsicht: Zum einen werden große Datensätze mit vielen Fallbeispielen geschaffen, was präzisere Aussagen erlaubt, zum anderen

werden verschiedene Datenquellen und die damit verbundenen vielfältigen Variablen miteinander so verknüpft, dass sich bislang unbekannte Korrelationen auffinden lassen. Ernährungs-Apps etwa sammeln Daten über Essgewohnheiten, die mit Patientendaten (zum Beispiel Puls, Blutdruck, Blutzucker, Gewicht) zu Datensätzen verknüpft werden können, die in der traditionellen medizinischen Forschung typischerweise nur mit hohem Aufwand und in kleiner Fallzahl erreichbar sind. Die durch Big-Data-Analysen neu aufgedeckten Zusammenhänge können ihrerseits sowohl im Kontext von Vorhersagen als auch im Kontext von Interventionen genutzt werden und bieten damit die Chance, diagnostische, therapeutische oder präventive Maßnahmen empirisch zu fundieren. Der Nachweis, dass früher gemessene Daten (etwa die tägliche Anzahl von Schritten) für die Prognose des späteren Gesundheitszustands (etwa der Herzgesundheit) Bedeutsames beitragen, kann etwa Interventionen nahelegen, mit denen versucht wird, die früher gemessenen Variablen – als mutmaßliche Einflussfaktoren für die spätere Entwicklung – gezielt zu verändern. Die Verfügbarkeit von detaillierten Informationen über Zusammenhänge zwischen verschiedenen Variablen, für die Messungen zu verschiedenen Zeitpunkten vorliegen, leistet damit einen Beitrag zur Entwicklung und Begründung von Präventionsmaßnahmen wie auch therapeutischen Bemühungen.

Die Verwirklichung der hier angesprochenen Stärken und Chancen von Big Data ergibt sich allerdings nicht allein aus der Vielzahl und Vielfalt verfügbarer Informationen. Vielmehr stellen sich gerade im Zusammenhang mit Big Data hohe Anforderungen an die Integration von Daten im Kontext statistischer Modellierungen. Dies hat zunächst damit zu tun, dass mit der Anzahl an Variablen und Messungen die Wahrscheinlichkeit von zumindest in Teilen widersprüchlicher Information zunimmt. Entsprechend müssen Qualität und Bedeutsamkeit von Daten (Objektivität, Reliabilität, Validität) reflektiert, gegebenenfalls Gewichtungen vorgenommen sowie Chancen und Risiken, die sich aus der Berücksichtigung alternativer Variablen in statistischen Modellen ergeben können, abgewogen werden. Da Big-Data-Anwendungen nicht selten auch auf sehr unterschiedliche Erhebungsmethoden zurückgreifen, sind die Besonderheiten dieser Erhebungsmethoden in den Blick zu nehmen. Dabei ist zu fragen, inwieweit unterschiedliche Messungen als äquivalent betrachtet werden können oder inwieweit sich aus der verwendeten Methode charakteristische Auswirkungen auf den interessierenden Gegenstand ergeben, der – wie angedeutet – in vielen Fällen ja gerade nicht direkt beobachtet werden kann. Dies führt zu der Frage, welche Variablen im Kontext statistischer Modellierungen miteinander verknüpft werden dürfen. Ein weiteres Problem, das sich aus der Vielzahl der verfügbaren Informationen ergibt, besteht darin, dass sich auf der Grundlage alternativer Modelle prinzipiell sehr unterschiedliche Vorannahmen treffen lassen. Alternative Modelle gehen zum Beispiel daraus hervor, dass bestimmte Randbedingungen und Parameter des Modells unterschiedlich definiert werden können. Aus diesem Grunde stellt sich im Kontext der Modellentwicklung die Aufgabe,

auch nach möglichen Falsifikationen und nicht lediglich nach bestätigenden Informationen zu suchen.

Die unabhängige Überprüfung und Verifizierung von Ergebnissen der Datenanalyse ist nicht nur im Zusammenhang von Big Data, sondern insgesamt im Kontext wissenschaftlicher Untersuchungen und des Erkenntnisgewinns ein zentrales Element. Solche Überprüfungen können an unabhängigen Datensätzen (zum Beispiel bei genetischen Studien in zwei unabhängigen Stichproben im Sinne einer Replikation), mittels empirischer Ansätze, die eine datenbasiert gewonnene Voraussage experimentell stützen bzw. widerlegen, oder schließlich durch die Verwendung alternativer Modelle erfolgen (da eine wissenschaftlich belegbare Aussage relativ robust gegenüber verschiedenen Modellen sein sollte). In der Kreuzvalidierung etwa verwendet man bei der Modellschätzung einen zufällig bestimmten Teil der Daten zunächst nicht, sondern setzt diesen später zur Überprüfung der Modellqualität ein. Eine andere Strategie zur Überprüfung besteht in einer organisatorischen oder/und institutionellen Trennung der Modellgenerierung von der Modellvalidierung. Das ist ein wichtiger Aspekt einer externen Qualitätssicherung und Überprüfung; für wissenschaftliche Studien im Bereich genetischer Untersuchungen an großen Kohorten wird er von führenden Fachzeitschriften oft als Publikationsbedingung gefordert. Methodische Kompetenz ist somit eine notwendige, aber nicht hinreichende Bedingung für hochwertige Big-Data-Analysen; es bedarf auch der Bereitschaft, die Algorithmen ergebnisoffen zu gestalten. Gerade im Kontext von Big Data kann eine selektive Auswahl und Verknüpfung von Daten manipulativ werden, und dies auch noch auf extern schwer überprüfbare Weise. Der mögliche Nutzen von Big Data steht und fällt mit der Expertise und Integrität der Personen oder Institutionen, die Daten generieren, auswählen, verknüpfen und interpretieren.

Das Zutreffen von Modellvorhersagen alleine erlaubt noch keine Aussage über die Richtigkeit des verwendeten Modells und auch nicht über einen Zusammenhang im Sinne von Ursache und Wirkung. Es geht meist nicht unmittelbar um die Verifikation oder Falsifikation von Kausalaussagen, sondern um die Verlässlichkeit von Vorhersagen. Statistik bietet zudem ihrem Wesen nach keine Zweifelsfreiheit im Einzelfall, sondern kann nur prädiktiven Wert haben, der mit einer gewissen Fehlerwahrscheinlichkeit einhergeht. Wahrscheinlichkeit ist hier im Sinne von relativer Häufigkeit zu verstehen und bezieht sich auf wiederholte Beobachtung (derselben Person oder verschiedener Personen). Die Vorhersage der Wahrscheinlichkeit eines bestimmten Resultats, etwa eines spezifischen Krankheitsverlaufs, ist somit eine Aussage darüber, was gemäß Erfahrungswerten gemittelt über eine Population gleichartiger Einheiten, etwa Patienten, zu erwarten ist. Aufgrund der Verwendung von neuen Verfahren des maschinellen Lernens wie des Deep Learnings (siehe Abschnitt 2.3.2) ist es jedoch zunehmend möglich, Aussagen mit sehr hoher Wahrscheinlichkeit zu erzielen, die eine große Relevanz auch für den einzelnen Patienten haben können.

Aus Korrelationen zwischen Variablen kann nicht ohne Weiteres auf kausale Effekte oder Mechanismen geschlossen werden. Der zu einem bestimmten Zeitpunkt zwischen zwei Variablen A und B beobachtete Zusammenhang kann darauf zurückgehen, (a) dass A B kausal beeinflusst, (b) dass B A kausal beeinflusst, (c) dass A und B einander beeinflussen und (d) dass A und B eine gemeinsame Ursache haben. Im Übrigen kann der beobachtete Zusammenhang zwischen A und B letztlich aus Besonderheiten der betrachteten Stichprobe resultieren. Diese verschiedenen Interpretationsmöglichkeiten zu unterscheiden, ist mitunter kompliziert und erfordert weitere Analyseschritte.

Der Wert von Daten bemisst sich oft am Wert der statistischen Modelle, die sich aus den Daten schätzen lassen. Statistische Modelle sind allerdings in einem wesentlichen Punkt begrenzt: Sie erlauben *als solche* keinen Rückschluss darauf, wie sich Zusammenhänge unter veränderten Bedingungen darstellen. Solche Veränderungen können sich zum Beispiel dadurch ergeben, dass sich die zugrunde liegende Grundgesamtheit ändert. Nimmt man etwa ein Modell, das auf Grundlage von Daten entwickelt wurde, die in einem spezifischen Umfeld gesammelt wurden – zum Beispiel Patienten einer bestimmten Geografie oder Nutzer einer bestimmten App –, so stellt sich die Frage, wie sich die Korrelationen und damit auch die Modellvorhersagen ändern, wenn man eine veränderte Population betrachtet – zum Beispiel Patienten einer anderen Geografie oder Nicht-Nutzer der betreffenden App.

In der traditionellen statistischen Analyse wird viel Wert darauf gelegt, die Bedingungen der Datenerhebung stringent zu kontrollieren, etwa durch Vermeidung von ungewollten Verzerrungen bei der Stichprobenerhebung oder durch Ausschluss von erwartbaren Störfaktoren. Im Bereich von Big Data ist eine solche Kontrolle oder auch nur das Wissen über den Datenerhebungsmechanismus oft nicht möglich bzw. nicht oder nur unzureichend vorhanden. Daraus ergeben sich grundlegende methodische Herausforderungen an die Modellierung von Daten sowie auch ein Gebot zur Vorsicht bei der Interpretation von Ergebnissen und daraus abgeleiteten (Allokations-)Entscheidungen mit Bezug auf Individuen.

Wichtige und zum Teil große Veränderungen entstehen durch nicht bekannte oder erfasste Faktoren, zum Beispiel infolge von gezielten Interventionen, etwa die Behandlung durch einen Arzt oder die Änderung von Lebensgewohnheiten basierend auf Gesundheitstipps aus einer Gesundheits-App. Hier geht es nicht mehr nur um Beobachtungsdaten und Korrelationen, sondern um Fragen von Ursache und Wirkung. Es geht also darum, wie solche Interventionen Korrelationen und Vorhersagen verändern. Die Korrelation zwischen Körpertemperatur und Beschwerden kann zum Beispiel nach Gabe eines fiebersenkenden Medikamentes unverändert bleiben, wenn das Fieber seinerseits kausal für Fieberkrämpfe ist und das Medikament nicht nur das Fieber, sondern auch das Risiko von Krämpfen reduziert. Ist das Fieber hingegen nur

Symptom einer anderen zugrunde liegenden Krankheit, deren Beschwerden durch die Temperatursenkung nicht beeinflusst werden, so wird die Korrelation zwischen Körpertemperatur und Beschwerde verändert.

Um Fragen nach der Wirksamkeit von Interventionen beantworten zu können, bedarf es daher kausaler Modelle[21] und kausaler Argumente, die ihrem Wesen nach anderer Natur sind als Erkenntnisse, die sich aus bloßen Korrelationen ergeben. Weder die Stärke einer Korrelation noch die Größe der Datenmenge erlaubt es, eine solche Diskussion zu umgehen. Es wäre also ein Missverständnis zu glauben, dass mehr Daten auch *automatisch* zu mehr Wissen über kausale Effekte führen. Es wäre zudem ein Kategorienfehler, Korrelationsaussagen mit Kausalaussagen zu verwechseln.[22] Der statistische Nachweis eines empirischen Zusammenhangs und die Bestimmung der Effektstärke sind unabhängig von einem tiefer gehenden Verständnis der Wirkmechanismen, die den Effekt herbeiführen.[23] Begründete Aussagen über Ursache-Wirkung-Zusammenhänge („Kausaleffekte“) können auch dann getroffen werden, wenn die Wirkmechanismen nicht oder nur unvollständig verstanden werden. Wegen der Komplexität biologischer Systeme ist das in den Lebenswissenschaften oft der Fall. Der medizinische Fortschritt profitiert in vielerlei Hinsicht von einem verbesserten Verständnis kausaler Zusammenhänge, etwa von Stoffwechselwegen in Zellen oder von Regelkreisen in Organismen, ist aber nicht zwingend an derartiges Wissen gebunden. So weiß man um die pharmakologische Wirksamkeit von Acetylsalicylsäure seit mehr als 2000 Jahren, ein erster molekularer Wirkmechanismus (von mehreren!) ist aber erst seit den 1970ern bekannt. Selbst wenn die Wissenschaft in ihren Erklärungen irrte, so würde das die Validität eines Kausaleffekts nicht widerlegen.[24]

Diese Unterscheidung zwischen Kausaleffekten und Kausalmechanismen oder zwischen einem Wissen um die Wirksamkeit und dem Wie des Wirkens bleibt auch unter Big-Data-Bedingungen relevant. Big-Data-Analysen können Korrelationen ermitteln, aus denen sich eine (interventionsfreie) Vorhersagekraft ergibt. Falls Vorhersagen von Interventionen gefordert sind, gilt es mittels zusätzlicher Argumente und Annahmen oder mittels Gewinnung zusätzlicher Daten, zum Beispiel aus Langzeit- oder experimentellen Studien, relevante Kausaleffekte zu identifizieren und in ihrer Stärke abzuschätzen. Korrelationsaussagen können darüber hinaus auch

[21] Siehe Pearl 2010.
[22] Allerdings ist zu berücksichtigen, dass es bisher wissenschaftsphilosophisch keine kanonische Rekonstruktion des Kausalitätsbegriffs gibt; vermutlich kommt man bei der Rekonstruktion von Kausalaussagen um pragmatische Interpretationen nicht herum (etwa dadurch, dass man die Verlässlichkeit von Kausalaussagen an experimentelle Interventionshandlungen, irreale Konditionalsätze oder Rahmenbedingungen des rationalen Wettverhaltens bindet). Vgl. unter anderem Stegmüller 1983, 501 ff.
[23] Jedoch geht ein (Vor-)Verständnis von Mechanismen oft in explizite oder implizite Annahmen ein, die ein kausales Modell konstituieren.
[24] Dabei steht außer Frage, dass der Wert von vor-wissenschaftlichen, operativ fundierten Einsichten in einem hohen Grad an Verallgemeinerbarkeit bestehen kann, wie schon in der Antike durch die Konzeption der *techne* bzw. *ars* (Kunst) in Abgrenzung zur *episteme* bzw. *scientia* (Wissenschaft) hervorgehoben wurde. In diesem Sinne sind die medizinischen Disziplinen zunächst Heil„künste“. Siehe auch Gethmann 1996.

helfen, bestimmte Zusammenhänge auszuschließen. Schließlich wird man langfristig versuchen, auf unabhängige Art und Weise die zugrunde liegenden Wirkmechanismen zu ergründen.

Big Data eröffnet in der medizinischen Forschung vor allem das Potenzial, durch die neuen Möglichkeiten des Zugriffs auf deutlich größere und vielfältigere Datenmengen, Korrelationen zwischen wesentlich mehr Faktoren schneller und besser zu entdecken und dabei auch neue Hypothesen über Wirkzusammenhänge zu entwickeln. Da die Aussagekraft statistischer Tests, wie bereits erwähnt, unmittelbar von der Stichprobengröße abhängt, ist hier allerdings immer auch die Frage nach der praktischen Bedeutsamkeit von Zusammenhängen zu stellen. Das betrifft insbesondere das Zusammenspiel der Wirkungsweise schwacher und/oder interagierender Einzelfaktoren, die in bestimmter Zusammensetzung einen Effekt haben. So kann es beispielsweise von Interesse sein zu untersuchen, welche genetischen Faktoren unter welchen Umwelt- und Lebensbedingungen mit einem erhöhten Risiko für Herz-Kreislauf- oder neurodegenerativen Erkrankungen assoziiert sind. Diese Art von Fragen wird in großen Kohortenstudien mit Tausenden und mitunter Zehntausenden Probanden untersucht, um daraus Anzeichen für die Früherkennung der Erkrankungen zu gewinnen, Prophylaxe zu betreiben und/oder potenzielle Ansatzpunkte für Therapien zu entwickeln.

2.3.2 Maschinelles Lernen und maschinelle Wahrnehmung

Das Gebiet des maschinellen Lernens ist ein rapide gewachsenes Teilgebiet der Informatik, das sich mit Modellen und Verfahren der Datenanalyse beschäftigt und mit einer Vielzahl von Anwendungsgebieten verzahnt ist. Seine historischen Wurzeln hat das maschinelle Lernen in Disziplinen wie Statistik, Physik, Mathematik, Mustererkennung und den Neurowissenschaften. Meist gibt es ein bestimmtes statistisches Modell, das mit Daten „trainiert" wird. Im typischen Fall der sogenannten automatischen Klassifikation „erlernen" Systeme anhand von spezifischen Trainingsdatensätzen Berechnungsvorschriften, die Daten in bestimmter Weise klassifizieren oder kategorisieren. Immer geht es darum, aus Rohdaten zunächst relevante Merkmale zu extrahieren und diese dann im Hinblick auf eine bestimmte Fragestellung miteinander zu verrechnen. Neben der reinen Klassifikation gibt es auch Techniken, die Wahrscheinlichkeiten für jede Alternative berechnen, oder solche, die quantitative Vorhersagen treffen. Dieses abstrakte Verfahren hat eine ungeheure und ständig zunehmende Anwendungsbreite, auch im medizinischen Bereich[25]: So erlernen Maschinen beispielsweise, MRT-Aufnahmen in Bezug auf das Vorhandensein und die Art von Tumoren zu klassifizieren oder entwickeln hochkomplexe Ent-

[25] Vgl. Litjens et al. 2017.

scheidungsregeln und Strategien, die eine automatisierte Bildverarbeitung in der Radiologie ermöglichen.[26] Auch aus unstrukturierten Daten wie Sprachaufnahmen von Patienten ist es möglich, Hinweise auf vielfältige körperliche und psychische Auffälligkeiten und Störungen zu gewinnen, von Müdigkeit und Drogenkonsum über Depressionen, posttraumatische Belastungsstörungen und Psychosen bis hin zu neurodegenerativen Erkrankungen.[27]

Auch beim maschinellen Lernen erlaubt eine größere Zahl von Fallbeispielen im Trainingsdatensatz eine höhere Vorhersagegenauigkeit des Modells. Dabei spielen die Erhebung von Daten und die Generierung und Zuweisung von Etiketten (Labels oder Tags), die das gewünschte Klassifikationsergebnis codieren, eine wichtige Rolle. Facebook etwa verfügt (auch via Instagram) über eine der weltweit größten Sammlungen von Fotos, auf denen Menschen abgebildet sind[28], die sich größtenteils über Annotationen oder Tags identifizieren lassen. Es ist daher nicht verwunderlich, dass Facebook eine bislang beispiellos leistungsfähige Gesichtserkennungssoftware entwickeln konnte[29], die beispielsweise auch von Firmen bei der Bewerberauswahl genutzt wird. Auch die britische Einzelhandelskette Tesco nutzt eine Gesichtserkennungssoftware, die die Gesichter wartender Kunden an der Kasse scannt, um die Werbung zu optimieren.[30] Auch im gesundheitsrelevanten Bereich werden diagnostische Algorithmen, die mithilfe maschinellen Lernens und großer Bilddatenbanken entwickelt wurden, zunehmend bedeutsam, beispielsweise bei der Diagnose von Hautkrebs auf Fotos[31] oder der Vorhersage von Depressionen anhand der Bildmerkmale und Metadaten von Instagram-Bildern.[32]

Neben der Zahl der Fallbeispiele bzw. Trainingsdaten ist auch die Beobachtungsbreite und -tiefe von essenzieller Bedeutung für die Qualität der Vorhersagen. Je mehr Messungen, Merkmale und Attribute in die Vorhersage einfließen und je stärker der statistische Zusammenhang mit der vorherzusagenden Zielgröße ist, desto größer sind die Chancen, dass ein Modell Fehler vermeidet, die etwa durch zu große Vereinfachung entstehen. Diese statistischen Grundwahrheiten erklären das Bestreben, in der Praxis möglichst viele Datenquellen miteinander zu verknüpfen. Die Bedeutung solcher datengetriebenen Vorhersagetechniken zur Wertschöpfung im Internet kann nicht hoch genug eingeschätzt werden. Neu ist dabei weniger ihr prinzipieller

[26] Vgl. Müller/Hanbury 2016.
[27] Zur Müdigkeit siehe etwa Krajewski et al. 2014; zum Drogenkonsum siehe Bedi et al. 2014; zu Depressionen siehe Scherer et al. 2013; zu posttraumatischen Belastungsstörungen und Psychosen siehe Bedi et al. 2015 sowie Mota et al. 2012; und zur Parkinsonkrankheit siehe Tsanas et al. 2012.
[28] Bereits im Jahr 2014 verfügte Facebook insgesamt über 250 Milliarden Fotos (vgl. http://www.businessinsider.com/facebook-350-million-photos-each-day-2013-9?IR=T [17.10.2017]). Zudem werden bei Facebook und Instagram zusammengenommen fast 450 Millionen neue Bilder pro Tag hochgeladen. (vgl. https://www.omnicoreagency.com/facebook-statistics [17.10.2017] und https://www.omnicoreagency.com/instagram-statistics [17.10.2017]).
[29] Vgl. Taigman et al. 2014, Taigman 2014 sowie Bennett 2017.
[30] Siehe https://www.theguardian.com/business/2013/nov/03/privacy-tesco-scan-customers-faces [17.10.2017].
[31] Vgl. Esteva et al. 2017.
[32] Vgl. Teramoto et al. 2017.

Einsatz als vor allem ihre erst durch die Entwicklung leistungsfähiger Rechner- und Datentransferstrukturen ermöglichte Nutzung zur Optimierung von Mikro-Entscheidungen in Echtzeit sowie die Größe des verarbeiteten Datenvolumens. Während in vielen Studien bislang mit Datensätzen zwischen unter hundert und einigen Tausend Datenpunkten gearbeitet wurde, geht es hier um Datensätze mit Hunderten Milliarden von Datenpunkten, die sekündlich mit der Nutzung von Dienstleistungen anwachsen.

Seit etwa 2010 hat es in einem Teilbereich des maschinellen Lernens, dem sogenannten Deep Learning, deutliche Fortschritte gegeben. Hierbei werden Prozesse entlang hierarchisch organisierter Schichten nachgebildet, ähnlich dem Modus, in dem im menschlichen Gehirn neuronale Netze operieren. Jede Schicht nutzt dabei die Ergebnisse der vorherigen Schicht und verarbeitet sie weiter zu neuen Ergebnissen. Angesichts größerer Datenmengen und erhöhter Rechenleistungen enthalten die verwendeten Netzwerke heute mehr aufeinander aufbauende Zwischenschichten als jemals zuvor. Der wesentliche Unterschied zu traditionellen Verfahren liegt darin, dass man im Deep Learning auch den Prozess der Merkmalsextraktion weitestgehend automatisiert hat und Modelle oftmals direkt mit Rohdaten trainieren kann. Dadurch lässt sich die Abhängigkeit von komplexen Vorverarbeitungsschritten reduzieren, in die auch Vorurteile und bloße Intuitionen von Entwicklern einfließen können.[33] Beispielsweise kann man bei der Bilderkennung direkt auf Basis der Pixel beginnen oder Sprache direkt auf Basis von Sätzen als Folge von Wörtern oder gar Buchstaben verarbeiten, um so zum Beispiel natürliche Sprache[34] zu erkennen.

Diese Effekte sind durch Modelle erzielbar, die einfache Berechnungselemente rekursiv kombinieren. Es werden sukzessive bessere Datenrepräsentationen gelernt, wobei die jeweils folgenden Repräsentationen auf den vorherigen aufbauen. Beim Computersehen etwa lernen solche Modelle zuerst lokale Kontrastfilter, dann einfache visuelle Merkmale wie Konturen oder Ecken, schließlich visuelle Komponenten (das Rad eines Autos, das Auge eines Menschen usw.) bis hin zu einer kompositionellen Repräsentation der vorfindlichen Gegenstände und Szenen. Ähnlich werden auch Spracherkennung, Signalverarbeitung, Robotik und viele weitere Anwendungsdomänen bereits von Deep Learning geprägt.

Deep Learning wird häufig mit dem Begriff der künstlichen Intelligenz in Verbindung gebracht. Zum einen hat der Durchbruch in der Verarbeitung sensorischer (insbesondere visueller und auditiver) Daten dazu geführt, dass Maschinen auf ihre Art „sehen“ und „hören“ können. Man spricht hier auch von einer Revolution der maschinellen Wahrnehmung. Das ist ein großer Schritt hin zur Entwicklung von Robotern, die sich in unserer Welt „autonom“ bewegen und

[33] Allerdings können Vorurteile durch die Verwendung falscher oder unerwünschter Merkmale lernender Algorithmen auch verstärkt werden. Vgl. Spielkamp 2017.
[34] Vgl. LeCun/Bengio/Hinton 2015.

mit Menschen auf natürlich wirkende Weise interagieren können, also zum Beispiel Roboter im Bereich der häuslichen Pflege. Auch darüber hinaus werden solche Systeme ein breites Einsatzgebiet finden, so etwa bei der Auswertung bildgebender Analyseverfahren in der Medizin. Die durch Ausbildung und Erfahrung erworbenen visuellen Fähigkeiten von Radiologen, Chirurgen, Laborassistenten usw. können durch automatische Verfahren komplementiert oder gar ersetzt werden. Es werden bereits heute sehr gute Ergebnisse bei der Unterscheidung von verschiedenen traumatischen Schädigungen des Gehirns in kernspintomografischen Aufnahmen[35] oder bei der Klassifizierung von Lungenkrebs[36] erzielt. Ähnlich verhält es sich mit auditorischen Fähigkeiten, die sprachbasierte Schnittstellen von hoher Qualität (zum Beispiel Apples Siri, Amazons Alexa oder der Google Assistant) oder sogar die bereits oben erwähnte Diagnose von Gesundheitsstörungen mittels automatisierter Verfahren der Stimmanalyse ermöglichen.

Eine Schlüsselfrage wird sein, in welchem Umfang solche Techniken zur Entwicklung von entscheidungsfähigen und -befugten maschinellen Agenten führen, die beispielsweise auch bei der Therapiegestaltung oder bei gesundheitspolitischen Entscheidungsprozessen, zum Beispiel über die Aufnahme neuer Behandlungsverfahren in die Leistungskataloge der Krankenversicherungen, beteiligt werden könnten. Ein Großteil des Internets ist bereits automatisiert; hier wird mit intelligenten Systemen interagiert, die aber oft im Verborgenen bleiben. Die oben skizzierte Revolution der maschinellen Wahrnehmung wird aber einen zunehmenden Einsatz dieser Techniken in der physischen Welt erlauben. Damit erhält die maschinelle Durchdringung des Alltags (*pervasive computing*) eine völlig neue Dimension, bei der es nicht nur um das passive Datensammeln mit Sensoren gehen wird, sondern auch um das aktive Eingreifen in das Geschehen mittels perzeptiver Systeme.

Neben den merkmalsbasierten Verfahren des maschinellen Lernens gibt es auch wichtige Varianten, die ohne Merkmalsextraktion auskommen. Das Paradebeispiel sind sogenannte Empfehlungssysteme. Ein Schlüsselbegriff ist hier die Relevanzoptimierung: die Auswahl des Wesentlichen aus der Überfülle des Angebots an Inhalten, Informationen und Produkten. Sie erfordert selbstlernende Systeme, die auf der Populationsebene, das heißt generalisierbar für eine große Gruppe von Menschen, maßgebliche Faktoren identifizieren und dann einzelne Personen und Inhalte in diesem Koordinatensystem verorten. Aus den Netflix-Nutzungsdaten etwa lassen sich Faktoren ableiten, die gewissen Seh-/Konsuminteressen der Nutzer entsprechen. Dadurch kann jeder Einzelne in dem durch diese Faktoren aufgespannten Koordinatensystem als spezifisch gewichtete Kombination eben solcher Interessen aufgefasst werden. Das Individuum wird also verortet in einem aus Daten extrahierten Merkmalsraum, der in der Regel nur begrenzt interpretierbar ist (weil er nicht mittels vorgegebener Dimensionen konzipiert wurde).

[35] Vgl. Kamnitsas et al. 2017.
[36] Vgl. Teramoto et al. 2017.

Um das Individuum einzuordnen, benötigt man nur seine Daten. Um aber die erwähnten Faktoren statistisch zu identifizieren, braucht es einen großen kollektiven Datensatz mit möglichst vielen Personen. Im Gesundheitsbereich wurden solche datenbasierten Empfehlungssysteme bereits entwickelt.[37]

Viele der gesammelten Daten werden direkt wieder in den Datenkreislauf der Dienstleistungen zurückgeführt, indem Dienste optimiert, adaptiert oder personalisiert werden, oft zum Vorteil des Nutzers. Die Qualität von Suchresultaten ist wesentlich abhängig von Nutzungsdaten: Welche Anfragen werden gestellt und wie verändert? Welche Resultate werden geklickt, welche ignoriert? Welche Rolle spielt der Kontext des Nutzers oder der aktuelle Aufenthaltsort? Stünden solche Daten nicht in gegebener Breite und großem Umfang zur Verfügung, käme es zu einer spürbaren Verschlechterung der Relevanz von Suchergebnissen und individualisierten Empfehlungen. Der reduzierte Aufwand bei der Suche nach individuell relevanten Informationen und Inhalten geht allerdings notgedrungen mit der Preisgabe persönlicher Daten einher.

Durch die bereits genannte Entwicklung der technischen Möglichkeiten, die Datenanalysen zunehmend ohne wahrnehmbare Verzögerung erlauben, können Computer immer überzeugender in Dialogform auf Nutzer reagieren. Dadurch kommt es potenziell zur Ausbildung eines „Gegenübers", das mit dem Nutzer in Echtzeit interagiert und mit dem dieser wie von Mensch zu Mensch kommunizieren kann. Die zentrale Komponente solcher Systeme ist die Spracherkennung, die die Barriere zwischen Mensch und Maschine aufbricht. Softwaregestützte Systeme sind schon seit mehreren Jahren im weitverbreiteten Gebrauch – etwa in Form von über das Internet angesteuerten Chatbots[38], wie sie bei der Fahrplanauskunft von Verkehrsunternehmen eingesetzt werden, aber auch bei auf Smartphones oder PCs vorinstallierten Systemen wie Apples Siri oder Microsofts Cortana. Seit 2016 werden von großen Anbietern auch auf eigenständiger Hardware basierende Bot-Systeme vertrieben, wie etwa Amazons Echo (Alexa) oder Google Home (Google Assistant). Sie dienen als Vertriebssysteme, auch in gesundheitsrelevanten Bereichen.[39] Solche Geräte können zudem die Funktion eines „intelligenten persönlichen Assistenten" (IPA) einnehmen. In solchen Anwendungen liegt zum einen ein großes Potenzial für eine barrierearme Lebensgestaltung, insbesondere für Menschen mit Einschränkungen in Mobilität oder Kognition. Zum anderen besteht aber auch die Gefahr eines Missbrauchs, zumal Chatbots inzwischen schon so weit entwickelt sind, dass für einen damit interagierenden Menschen kaum noch eine Unterscheidung möglich ist, ob mit einem menschlichen Partner oder

[37] Vgl. etwa Bocanegra et al. 2017 und Gräßer et al. 2017.
[38] Chatbots sind automatisierte Computerprogramme, die es Anwendern via Sprach- und/oder Textein- und -ausgabe ermöglichen, menschliche Dialoge zu simulieren.
[39] Alexa wird über Amazons Lautsprecher Echo bereits als Vertriebskanal für Pflege- und Krankenzusatzversicherungen eingesetzt (vgl. https://www.versicherungsbote.de/id/4854127/Amazon-Versicherung-DFV/ [17.10.2017]).

einer Software kommuniziert wird. Damit werden Täuschungen und gegebenenfalls Manipulationen persönlicher Entscheidungen möglich.

2.3.3 Stratifizierung und Individualisierung

Gerade Verfahren des maschinellen Lernens sind potenziell dafür geeignet, Unterschiede innerhalb von Personengruppen zu identifizieren, die bisher als homogen galten. Musste man etwa in der Vergangenheit Patienten in wenige Hauptgruppen stratifizieren, weil man, basierend auf relativ wenigen Daten, nur Grobeffekten nachgehen konnte, so verspricht Big Data in der Medizin optimierte Diagnostik, Prognose und Therapie, die die spezifischen Bedingungen des Individuums stärker berücksichtigen (sogenannte Präzisions- bzw. „personalisierte" Medizin). Dies führt zu einer neuen Sicht auf fundamentale Begriffe wie etwa den der Krankheit: Die „gleiche" Krankheit ist bei verschiedenen Patienten eben nicht immer gleich und jedenfalls nicht gleich, sondern individuell zu behandeln.

Stratifizierungen sind ihrem Wesen nach immer Reduktionen komplexer individueller Merkmalsprofile auf bestimmte Einzelaspekte. Je weniger Gruppen durch diese Verfahren am Ende gebildet werden, desto stärker ist die Reduktion der Individualität und umso größer die Gefahr einer fehlerhaften Zuordnung. Allerdings ist gerade in medizinisch-wissenschaftlichen Zusammenhängen oft nur eine zweigeteilte Gruppenbildung möglich, beispielsweise bei der Entscheidung für oder gegen die Aufnahme in ein Studienkollektiv für ein neues Medikament. In anderen Zusammenhängen, beispielsweise im Versicherungswesen, können auch graduelle Stratifizierungen vorgenommen werden, etwa hinsichtlich der Einstufung in Prämiengruppen aufgrund des Lebensalters oder bestimmter gesundheitsrelevanter Daten.

Die Bildung solcher Gruppen und die auf ihrer Basis durch Algorithmen gebildeten Prognosen können für den Nutzer solcher Analysen und die von ihnen betroffenen Personen hilfreich sein. Dabei kann ihre mögliche Fehlerspanne mit Bezug auf die jeweiligen Zwecke angemessen sein. Algorithmen können aber auch zu problematischen fehlerhaften Zuordnungen führen; komplexe Big-Data-Algorithmen können dabei die Ermittlung und Beseitigung der Fehlerquellen erschweren oder gar unmöglich machen, je nachdem, wie zum Beispiel im Algorithmus das Vorkommen und der Umgang mit bestimmten Fehlerquellen berücksichtigt oder bestimmte Randbedingungen für ein Modell festgelegt sind. So könnten im Zusammenhang mit seltenen Krankheiten, deren Träger in ihren Eigenschaften von den gebildeten Standardkollektiven abweichen, bestimmte Therapieentscheidungen problematisch werden. Beispielsweise ist zwar für die weitaus meisten Menschen, bei denen mit automatisierter hämatologischer Diagnostik eine Anämie festgestellt wird, eine Therapie mit Eisenpräparaten hilfreich oder zumindest unschädlich – bei der etwa jeden dreihundertsten Europäer betreffenden hereditären Hämochromatose

würde eine solche Behandlung aber das Krankheitsbild massiv verschlimmern. Solche Informationen müssen von Beginn an in den Algorithmus einfließen oder in Nachhinein berücksichtigt und korrigiert werden.

2.4 Personen- und Gesundheitsbezug

Aus der Fülle der Daten, die mit den beschriebenen Big-Data-Methoden erhoben und verarbeitet werden können, sind für den Gesundheitsbereich vor allem solche Daten von Interesse, die das Potenzial haben, neue Erkenntnisse in der biomedizinischen Forschung, Diagnostik, Prädiktion und Therapie zu ermöglichen. Solche personenbezogenen Gesundheitsdaten gelten als besonders sensible Daten, weil sie tiefe Einblicke in einen sehr intimen Bereich ermöglichen (siehe Kapitel 3 und 4). Ein Schlüsselmerkmal der aktuellen technischen Entwicklungen liegt darin, dass personenbezogene Daten aus einer immer größeren Zahl von Quellen gesammelt und verknüpft werden können und im Verlauf des Auswertungsprozesses dabei auch solche Daten Gesundheitsrelevanz erlangen können, von denen man dies auf den ersten Blick nicht erwarten würde (siehe Abschnitt 2.4.3).

2.4.1 Personenbezug

Fast alle im Internet oder mittels persönlicher Geräte gesammelten Daten haben einen expliziten oder impliziten Personenbezug. Dies liegt einerseits in der Natur der Datenerhebung und ist andererseits auch wesentlich für die Wertschöpfung aus Daten. Erst der Bezug auf eine Person und ihre Daten sowie auf das damit verbundene Verknüpfungspotenzial von Daten stiftet in der Regel einen monetarisierbaren oder klinisch operationalisierbaren Wert. Dabei entscheidet die Stärke des Personenbezugs sowohl über den Umfang dieses Wertes als auch über die Zuverlässigkeit von Anonymisierungen.

Vollständige Identifikation erfolgt in der Regel über den Namen einer Person, insbesondere sofern durch Postanschrift, Telefonnummer, E-Mail-Adresse, Kreditkartennummer usw. Eindeutigkeit in Bezug auf das Individuum hergestellt werden kann. Üblicherweise gilt die Verwendung von Pseudonymen als eine Alternative zum Klarnamen, wenn die Identität einer Person verborgen bleiben soll. Im Web werden Pseudonyme oft als Nutzernamen oder Nutzerkennungen codiert. Bemerkenswert ist insbesondere die Verwendung von Einmalanmeldediensten, sogenannten *single sign-ons* (SSO). Aus Sicht der Nutzer geht es oft darum, eine Vielzahl von Diensten, Apps und Websites zu nutzen, ohne für jedes ein eigenes Login anlegen zu müssen. Stattdessen erlauben viele Dienste die (Wieder-)Verwendung der Anmeldedaten von Facebook (Facebook Connect), Twitter, Google und anderen. Dies führt zu einer erheblichen Verbreitung von Daten, bei denen die betreffende Website Zugriff auf Profildaten des Nutzers bekommt, der SSO-Dienstleister aber umgekehrt Daten über die Webaktivitäten des Nutzers sammeln kann. Waren soziale Log-ins zum Austausch zwischen Bekannten und Freunden gedacht,

so werden sie auf diese Weise für Konzerne zum Instrument des Datensammelns, das die Fragmentierung der persönlichen Daten im Internet reduziert, um einen umfassenderen Zugang zu Verhalten und Präferenzen der Nutzer zu erlangen. Der bedingte Schutz der Privatheit, den Datenfragmentierung für den Einzelnen bietet, wird damit tendenziell ausgehöhlt.

Bei vielen Daten ist der Bezug auf Personen nur indirekt, insofern der zugrunde liegende Identifikator (ID) statt direkt mit der Person, mit einem Gerät (Geräte-ID) oder einem Programm (etwa einem Internetbrowser) verbunden ist. Relevant sind hier vor allem Gerätekennungen von Smartphones, weil sie in der Regel von einer einzigen Person benutzt werden. Die Verwaltung von Gerätekennungen obliegt den Herstellern oder Plattformbetreibern, die auch in ihren Richtlinien die Weiterverwendung regulieren. Weit verbreitet ist die Verwendung von HTTP-Cookies, die Daten in einem Internetbrowser speichern und diese dadurch eindeutig markieren können (sogenannte Tracking-Cookies). Durch die Technik des *cookie syncing* können Identitäten von verschiedenen Dienstleistern zudem vollautomatisch – und oft ohne Kenntnis des Nutzers – abgeglichen werden, was eine Grundbedingung dafür ist, umfassendere und persistentere Nutzerprofile aufbauen zu können. Diese Technik ist vor allem in der Online-Werbebranche weitverbreitet (siehe Abschnitt 2.5.4). Techniken wie das *canvas fingerprinting*[40] können Nutzer ganz ohne Verwendung von Cookies identifizieren, indem Browser anhand von Merkmalen der installierten Plug-ins, Fonts, Versionen usw. identifiziert werden. Heute ist eine unübersichtliche Lage entstanden, in der immer wieder neue Wege gefunden werden, um Geräte oder Browser mit einem digitalen Fingerabdruck zu versehen und so wiedererkennbar zu machen.

Solche Identifikatoren werden zunehmend auch in gesundheitsbezogenen Situationen relevant, etwa durch den Einsatz von mobilen Geräten und Apps, die Aktivitäten und Vitalfunktionen individueller Nutzer aufzeichnen (siehe Abschnitt 2.5.5). Dies erfolgt im Rahmen von Internetsuchen oder Online-Einkäufen zu bestimmten Gesundheitsthemen, oder sogar mit dem Ziel einer präzisen Dokumentation der Inanspruchnahme und Abrechnung von Gesundheitsleistungen.[41] Die Frage der Identifikatoren ist deswegen so wichtig, weil sie einerseits über die Wahrung oder Verletzung der Privatsphäre entscheidet, und zum anderen, weil sie den Schlüssel zur Verknüpfung und Anreicherung von Daten bildet. Kann man etwa Identifikatoren einander zuordnen, so können anonyme Daten deanonymisiert werden, wenn einer der Identifikatoren eindeutig eine Person identifiziert. Ebenso kann durch Zuordnung von Identifikatoren eine Beobachtungspräzision erreicht werden, die dazu führt, dass es nur noch ein Individuum gibt, auf das alle Daten passen. Bereits vier Raum-Zeit-Koordinaten mit geringer Auflösung

[40] Vgl. Acar et al. 2014.
[41] Vgl. https://www.google.com/patents/US9740823 [17.10.2017].

reichen beispielsweise aus, um mit 95 Prozent Wahrscheinlichkeit ein Individuum zu identifizieren.[42] Ähnliches gilt für Suchanfragen und für Anfragen bei digitalen Kartendiensten. Google Maps etwa kann mit hoher Wahrscheinlichkeit auf die Heimatadresse des Nutzers schließen, weil sie oft ein Endpunkt der Routenplanung ist. Werden diese Daten mit Suchanfragen kombiniert, so ist leicht, von der Straßenadresse weiter auf das Individuum zu schließen.

Neben diesen impliziten Identifikatoren gibt es auch Daten, die explizit als eindeutige Kennzeichen der Identifizierung eingesetzt werden: genetische Daten, diverse biometrische Daten (wie Fingerabdrücke und Iris-Scans), aber auch Fotos oder Videos von Gesichtern, die zum Beispiel mit öffentlichen Datenquellen wie Facebook-Profilfotos abgeglichen werden können. Die Kontroverse, die die bei Facebook anhand von Fotos angebotene automatische Gesichtserkennung und Personenidentifizierung ausgelöst hat[43], zeigt eindringlich die gesellschaftliche Relevanz solcher Fragen der Reidentifizierbarkeit.

2.4.2 Gesundheitsbezug

Gesundheitsrelevante Daten fallen in verschiedenen, einander teilweise überschneidenden Kontexten an, von der medizinischen Praxis und gesundheitsbezogenen Forschung über Behörden und Versicherer bis hin zur aktiven und unbeabsichtigten Datengenerierung durch Bürger bzw. Patienten.[44]

In Forschung und Medizin (siehe Abschnitt 2.5.1 und 2.5.2) fallen durch den Einsatz zunehmend leistungsfähiger bildgebender und molekularbiologischer Verfahren besonders große Datenmengen an. So liegt der typische Umfang der aus digitalem Röntgen, Ultraschall, CT- oder MRT-Scans pro Untersuchung gewonnenen Bilddaten im Bereich von einigen Megapixeln bzw. Megavoxeln. Auch Hochdurchsatzverfahren wie die sogenannten Omik-Technologien führen zu einem erheblichen Anwachsen des Datenvolumens. Zu den Omik-Technologien zählen zum Beispiel die Genomik (Erforschung des Aufbaus von Genomen und der Wechselwirkungen zwischen Genen), die Proteomik (Erforschung von Eiweißen), die Metabolomik (Erforschung des Stoffwechsels) sowie auch die Nutriomik (Erforschung der Interaktion von Nährstoffen mit dem Organismus) und die Nutrigenetik (Erforschung der Interaktion zwischen Ernährung und Genetik).[45]

Im klinischen Kontext werden zudem bereits seit Jahrzehnten systematisch von allen im Gesundheitssystem behandelten Menschen Anamnesen, Labor- und Bildbefunde sowie Diagnoseschlüssel für die Administration von Kliniken und Arztpraxen sowie das Abrechnungswesen

[42] Vgl. Montjoye et al. 2013.
[43] Vgl. hier zum Beispiel http://www.zeit.de/digital/datenschutz/2012-09/facebook-gesichtserkennung-dpc [17.10.2017].
[44] Vgl. Langkafel 2014, 14.
[45] Vgl. Wirth 2015.

gesammelt. Diese Daten sind grundsätzlich durch die ärztliche Schweigepflicht besonders geschützt, nichtsdestoweniger aber eine auch monetär höchst wertvolle Ressource für Big-Data-Anwendungen. Das E-Health-Gesetz[46] sieht einen straffen Zeitplan für die weitere Digitalisierung von Gesundheitsdaten vor. Es schreibt unter anderem die für 2019 zunächst freiwillige Einführung der elektronischen Patientenakte vor. Schon jetzt bieten private Unternehmen die Möglichkeit, mobile Gesundheitsakten[47] anzulegen, die Patienten den Datenaustausch mit Ärzten und anderen Gesundheitsdienstleistern erleichtern sollen. Hinzu kommen vielfältige weitere gesundheitsbezogene Daten, die Bürger bzw. Patienten über Sensoren in mobilen Endgeräten und Apps aufzeichnen und online verwalten und teilen (siehe Abschnitt 2.5.5).

2.4.3 Dekontextualisierung und Rekontextualisierung

Die skizzierten technischen und gesellschaftlichen Entwicklungen haben weitreichende Bedeutung für den Gesundheitsbereich. Traditionell weisen individuelle Gesundheitsdaten aufgrund ihrer intimen Einbettung in die persönliche Lebenssphäre und angesichts der Gefahr negativer Auswirkungen ihrer Veröffentlichung in sozialen Kontexten wie der Arbeitswelt (siehe Abschnitt 2.5.3) eine besondere Sensibilität auf, die auch im Zusammenhang mit der rechtlichen und ethischen Evaluation von Big Data (siehe Kapitel 3 und 4) besonders zu berücksichtigen ist. Big-Data-Technologien ermöglichen jedoch darüber hinaus eine umfassende Dekontextualisierung und Rekontextualisierung von Daten, die zu unterschiedlichen Zwecken erfasst, analysiert und neu verknüpft werden. Dies führt zu einer Entgrenzung des gesundheitsrelevanten Bereichs bzw. lässt diese absehbar erscheinen, da sich die klare Abgrenzung gesundheitsrelevanter von nicht gesundheitsrelevanten Daten angesichts der zunehmenden Verknüpfung aller Lebensbereiche, in denen faktisch oder potenziell Daten erhoben werden können, immer schwieriger gestaltet.

Individuelle Datensätze aus disparaten Quellen werden zunächst in depersonalisierter und anonymisierter Form verarbeitet; sie werden sozusagen in eine algorithmische „Blackbox“[48] eingespeist, in der Datenverarbeitungsprozesse stattfinden, deren Zusammenhänge und Regeln sich aber von außen mitunter kaum noch nachvollziehen lassen.[49] Diese Vorgänge können unterschiedliche Ergebnisse hervorbringen: Die ursprünglichen Daten werden entweder nach bestimmten Kriterien gruppiert, wodurch neue Cluster und Kohorteneinteilungen zustande kommen, oder sie werden in einen anderen Kontext übertragen. Diese Vorgänge ermöglichen im

[46] Gesetz für sichere digitale Kommunikation und Anwendungen im Gesundheitswesen sowie zur Änderung weiterer Gesetze vom 21. Dezember 2015 (BGBl. I, 2408). Das Gesetz fordert die Beschleunigung der Einführung und Ausweitung digitaler Anwendungen im Gesundheitswesen, wie beispielsweise die Einführung der elektronischen Gesundheitskarte oder die schrittweise Ablösung bislang papierbasierter Prozesse beim Formularwesen durch IT-unterstützte Verfahren.
[47] Zum Beispiel lifetime.eu oder medis-app.de.
[48] Zum Begriff der Blackbox-Medizin siehe auch Nicholson Price II, 2015.
[49] Siehe hierzu Bleicher 2017.

Gesundheitsbereich zum Beispiel, dass in der Klinik erhobene Laborwerte mit in Forschungslaboren durchgeführten Gesamtgenomanalysen verknüpft werden und so verbesserte Therapiemöglichkeiten für individuelle Patienten ausgewählt werden können. Auf diese Weise eröffnen sich neue Möglichkeiten, die Diagnosestellung zu präzisieren und Erfolg versprechende therapeutische Strategien zu etablieren.

Dekontextualisierung und Rekontextualisierung werfen aber auch Probleme hinsichtlich des Datenschutzes und der informationellen Selbstbestimmung auf. Sie erleichtern eine Deanonymisierung von Daten bzw. die Reidentifizierung einzelner Nutzer. Zudem können – unter Vernachlässigung des probabilistischen Charakters der Ergebnisse – durch pauschale Eingruppierungen falsche oder verfrühte Rückschlüsse bezüglich bestimmter Merkmale, Verhaltensweisen oder gar des Lebenswandels der betreffenden Person gezogen werden. Wer zum Beispiel das Bonusprogramm seiner Krankenversicherung nicht nutzt und im Supermarkt beim Weinkauf regelmäßig eine elektronische Kundenkarte verwendet, mag ungerechtfertigt einer zu hohen Risikogruppe zugeordnet werden, und dies, obwohl er sich in anderer Hinsicht sehr wohl gesundheitsförderlich verhält, das aber nicht digital dokumentiert, etwa durch regelmäßigen Sport ohne Fitness-Tracker oder Verwendung gesunder Lebensmittel vom Wochenmarkt, wo ohne Kundenkarte eingekauft wird.

Nachdem alle Daten, die in irgendeiner Form erhoben werden (beispielsweise über Arbeits- und Ruhezeiten, Wohn- und Aufenthaltsorte, Freizeitaktivitäten, Konsumverhalten etc.), in Relation zur persönlichen Gesundheit interpretiert werden *können*, ist es prinzipiell möglich, all diese Daten auch als gesundheitsrelevant einzuschätzen. Ob bestimmte Daten als sensibel oder gesundheitsrelevant zu betrachten sind, lässt sich angesichts dieser Entwicklungen somit oft nicht mehr zum Zeitpunkt ihrer Erhebung bestimmen, sondern hängt zunehmend vom Kontext ab, in dem sie verwendet werden. Das wirft die Frage auf, wie eine dynamische und kontextabhängige Beurteilung von Chancen und Risiken der Datenverwendung aus rechtlicher und ethischer Perspektive gelingen kann und wie sich vor diesem Hintergrund Lösungsansätze für eine praktikable Bewältigung der mit diesen Chancen und Risiken verbundenen Herausforderungen formulieren lassen.

Bereitschaft zur Datenweitergabe in gesundheitsrelevanten Kontexten

Verschiedenen Studien zufolge hängt die Bereitschaft, Daten weiterzugeben, stark vom Verwendungskontext ab. Ein entscheidendes Kriterium für die Bereitschaft zur Datenweitergabe ist zunächst, zu welchem Zweck die Erhebung und Auswertung von Daten erfolgt. Laut einer europaweiten Studie des Vodafone-Instituts steht die Mehrheit der Europäer – insbesondere in

Deutschland – der massenhaften Erhebung und Auswertung von Daten kritisch gegenüber.[50] Szenarien, in denen insbesondere anonymisierte Gesundheitsdaten zum Zwecke der Verbesserung der Diagnose und Therapie von Krankheiten gesammelt und analysiert oder zu Forschungszwecken weitergegeben werden, stoßen demnach jedoch viel eher auf Zustimmung als die Nutzung von Daten, die aus Gesundheits-Apps stammen und die bei der Planung von Krankheitsprävention helfen sollen oder gar von Versicherern zur Anpassung von Tarifen eingesetzt werden. Laut einer Studie des Meinungsforschungsinstituts YouGov ist jedoch selbst dazu immerhin ein Drittel der Nutzer bereit, gesundheits- und fitnessbezogene Daten zu messen und mit einer Krankenversicherung zu teilen, sofern sich daraus finanzielle Vorteile ergeben.[51]

Nach einer nicht repräsentativen Online-Befragung, die von der Wirtschaftsprüfungsgesellschaft PricewaterhouseCoopers in Auftrag gegeben wurde, sind vor allem jüngere Befragte bereit, private und gegebenenfalls auch familiäre Daten zur Verbesserung von Therapien bereitzustellen.[52] Relevant ist außerdem, an wen welche Gesundheitsdaten weitergegeben werden. Der größte Zuspruch gilt der Weitergabe an Ärzte und Kliniken, gefolgt von akademischen Forschungseinrichtungen, Krankenversicherungen, forschenden Pharmaunternehmen, Apothekern, Behörden und neutralen Agenturen wie zum Beispiel Verbraucherschutzzentralen. Laut einer repräsentativen Erhebung der Bitkom wollen 60 Prozent der Befragten eine elektronische Patientenakte nutzen, in der Daten, die in Arztpraxen, Kliniken oder anderen Gesundheitseinrichtungen anfallen, elektronisch gespeichert werden.[53] Hierbei möchten drei Viertel selbst darüber bestimmen, welche Ärzte Zugriff auf die digitalen Daten in ihrer E-Akte haben. Eine repräsentative Studie der Krankenkasse pronova BKK schließlich ergab, dass zwei von drei Bundesbürgern persönliche Daten für Beratungszwecke ihrer Krankenkasse freigeben würden, vorausgesetzt, dass sie anonymisiert und unter Einhaltung aller geltenden Datenschutzregeln verarbeitet werden.[54] Die größten Bedenken der Befragten gelten studienübergreifend den Kontrollverlusten, die sich bei der Weitergabe von Daten an Dritte ergeben, insbesondere wenn diese zu kommerziellen Zwecken erfolgt und aufgrund der weiteren Datennutzung eventuelle Nachteile für den Datengeber drohen.

[50] Vgl. Vodafone Institute for Society and Communications 2016, 15.
[51] Siehe https://yougov.de/loesungen/ueber-yougov/presse/presse-2015/pressemitteilung-self-tracking-rund-jeder-dritte-wurde-gesundheitsbezogene-daten-an-krankenversicherer-weitergeben [17.10.2017].
[52] Siehe https://www.pwc.de/de/gesundheitswesen-und-pharma/assets/personalisierte-medizin-studie-2016.pdf [17.10.2017].
[53] Siehe https://www.bitkom.org/Presse/Anhaenge-an-PIs/2017/03-Maerz/Verbraucherstudie-Telemedizin-2017-170327.pdf [17.10.2017].
[54] Siehe https://www.pronovabkk.de/downloads/14e3337132d6b4b5/Studie_Gesundheitsversorgung_2017.pdf [17.10.2017].

2.5 Akteure und Handlungskontexte

An der Erhebung, Verarbeitung und Nutzung von Datenmassen sind verschiedene Akteure in diversen, einander zum Teil überschneidenden Funktionen und Rollen beteiligt. Die mit den jeweiligen Akteuren verbundenen Strukturen und Motivationen bringen nicht nur unterschiedliche Rahmenbedingungen bei der Datenakquisition und -auswertung mit sich, sondern können sich auch unterschiedlich auf die Sicherheit, Qualität und Überprüfbarkeit der erhobenen Daten und des jeweiligen Umgangs mit ihnen auswirken. Uneindeutigkeiten ergeben sich nicht nur aus der Vielzahl unterschiedlicher Akteure und Handlungskontexte, sondern auch aus den damit jeweils verbundenen Zielen, Einschätzungen von Chancen und Risiken und den jeweils zur Verfügung stehenden Ressourcen. Diese mehrdimensionale Vielfalt lässt erwarten, dass sich für den Einsatz von Big Data im Gesundheitsbereich keine einfache und in sich kohärente normative Gesamtbeurteilung finden lassen wird.

Deshalb werden im Folgenden ausgewählte Anwendungskontexte von Big Data exemplarisch auf ihre jeweiligen Chancen und Risiken untersucht: erstens die biomedizinische Forschung, zweitens die Gesundheitsversorgung, drittens die Datennutzung durch Versicherer und Arbeitgeber, viertens die kommerzielle Verwertung gesundheitsrelevanter Daten und fünftens ihre Erhebung durch Betroffene selbst. Auch wenn die Kontexte nachfolgend separat behandelt werden, sind sie miteinander verbunden und gehen teils fließend ineinander über. Die oben beschriebenen Big-Data-getriebenen Entgrenzungstendenzen zeigen sich auch darin, dass Forschung und Versorgung oder kommerzielle und nicht kommerzielle Daten-Nutzung einander immer stärker annähern und zunehmend schwer klar zu trennen sind.

2.5.1 Big Data in der biomedizinischen Forschung

Die Auswertung großer Mengen gesundheitsrelevanter Daten spielt in der Wissenschaft schon seit Längerem eine wichtige Rolle, da sie erheblichen Wissenszuwachs bei zunehmender Effizienz verspricht. Sie soll nicht nur zu einem besseren Verständnis biomedizinisch relevanter Zusammenhänge und Prozesse führen, sondern auch präventive, diagnostische und therapeutische Maßnahmen in der medizinischen Praxis verbessern. Die biomedizinische Forschung stößt dabei in neue Dimensionen von Daten und Datenintegration vor, insbesondere in der Genomforschung, aber auch in der Hirnforschung. Im Rahmen von Big Data nutzen und erheben Forscher für ihre wissenschaftlichen Projekte Daten in immer größerem Umfang, analysieren diese mithilfe zunehmend avancierter Techniken und verknüpfen immer mehr unterschiedliche Quellen und Arten von Daten. Das führt zu einer steigenden Komplexität bei der Auswertung.

Zu den zentralen Akteuren im wissenschaftlichen Bereich gehören Forschungsinstitutionen wie Hochschulen, außeruniversitäre öffentliche und private Forschungseinrichtungen sowie deren Forscher und Assistenzpersonal, aber auch die Probanden und Patienten, die ihre Daten für die

Forschung zur Verfügung stellen. Die Arbeit mit Datenmassen erfolgt im Forschungsbereich in der Regel nach hohen und gut kontrollierbaren Standards der Datenerhebung, -verwendung und -sicherheit, insbesondere in Forschungsnetzwerken auf nationaler, europäischer und internationaler Ebene, die speziell auf den Umgang mit großen Datenmengen ausgerichtet sind (zum Beispiel Krebsregister, Deutsches Forschungsnetz, PRACE-Netzwerke). Geografisch verteilte Wissenschaftsorganisationen machen sich die neuen technischen und infrastrukturellen Möglichkeiten von Big Data zunutze und vernetzen sich zum Zweck des Datenaustauschs und der gemeinsamen Analyse und Auswertung.[55] Darauf aufbauend erzeugen wissenschaftliche Organisationen kollaborative Dateninfrastrukturen zur langfristigen Datenhaltung, -identifikation und -replikation sowie zur semantischen Annotation und Metadatensuche.[56] Zunehmend kommen dabei auch Verfahren, Software und Standards aus dem kommerziellen Bereich zum Einsatz (siehe Abschnitt 2.5.4).

Die integrative Berücksichtigung vielfältiger Daten gilt in der biomedizinischen Forschung auch deshalb als besonders aussichtsreich, da Krankheiten durch die Kombination und Interaktion verschiedener Faktoren wie zum Beispiel genetischer Veranlagungen, Umwelteinflüsse oder der persönlichen Lebensführung bestimmt werden. Aufgrund dieser Komplexität können die Ausbruchswahrscheinlichkeit, die Ausprägung, der Verlauf und das subjektive Erleben von Krankheiten erheblich zwischen Personen variieren, was die Prädiktion und Prävention, Diagnose und Therapie erschwert. Oft sind ursächliche Mechanismen noch nicht oder nur teilweise verstanden, sodass eine gezielte und damit möglicherweise effizientere Behandlung nicht möglich ist. Die Analyse großer Datenmengen eröffnet hier große Chancen, zum Beispiel um Merkmale zu identifizieren, die Patienten mit einem bestimmten Krankheitsbild gemeinsam sind.[57]

Im Gesundheitsbereich werden größere Datensammlungen gegenwärtig vor allem aus genomischen Daten erstellt. Genetische Faktoren spielen eine wichtige Rolle in vielen gesundheitsrelevanten Prozessen und die technischen Möglichkeiten für Sequenzanalysen sind in diesem Bereich bereits sehr weit fortgeschritten. Im Rahmen von genomweiten Vergleichsstudien untersuchen Forscher zum Beispiel, wie oft einzelne Varianten im Genom gemeinsam mit bestimmten Zielmerkmalen, wie etwa besonderen Krankheiten, auftreten. Mithilfe solcher Korrelationen hofft man, genetisch bedingte Erkrankungsrisiken besser einschätzen zu können. Für die

[55] Siehe etwa die Kollaboration des Indiana Biosciences Research Institute, des Pharmaunternehmens Eli Lilly and Company, der Roche Diagnostics Corporation, des Regenstrief Institute und der Indiana University School of Medicine zur Erforschung der Prävention und Behandlung von Typ-2-Diabetes (vgl. http://www.indianabiosciences.org/news/?newsId=39 [17.10.2017]).
[56] Siehe beispielsweise das Projekt „NCT DataThereHouse“ des Nationalen Centrums für Tumorerkrankungen Heidelberg (vgl. https://www.nct-heidelberg.de/das-nct/vorstellung/innovationen/nct-datatherehouse.html [17.10.2017]).
[57] Konkrete Beispiele mit direktem Einfluss auf die Behandlung von Patienten sind hier etwa die Reklassifizierung mutierter BRCA1- oder BRCA2-Gene von „krankheitsverursachend“ zu „neutral“ sowie die umgekehrte Reklassifizierung von „wahrscheinlich neutral“ zu „wahrscheinlich krankheitsverursachend“ bei bestimmten Mutationen im CTNS-Gen. Siehe Lek et al. 2016.

verlässliche Identifikation von krankheitsrelevanten genetischen Merkmalen wird allerdings eine ausreichend große Anzahl an Stichproben benötigt. Im Bereich der Alzheimer-Forschung werden zum Beispiel Kohorten mit mehreren Zehntausend Patienten untersucht.

Medizinische Forschung der Gentestfirma 23andMe

Neue Möglichkeiten für die Beschaffung ausreichend großer Datenmengen ergeben sich mitunter auch aus der Einbindung von Ressourcen jenseits des klassischen Wissenschaftsbetriebs. Die US-amerikanische Biotechnologiefirma 23andMe beispielsweise, die genetische Analysen zur Abstammung und zu gesundheitlichen Risiken direkt an Privatpersonen vermarktet, bietet ihren Kunden die Freigabe ihrer Daten für Forschungsprojekte in Kollaboration mit öffentlichen oder privaten Forschungseinrichtungen an. Zur Erforschung der Parkinsonkrankheit verglich die Firma etwa bestimmte Genregionen von 3.400 durch Patientengruppen und Kliniken vermittelten Parkinsonpatienten mit denen von fast 30.000 Firmenkunden, die nicht von der Krankheit betroffen waren, und konnte so zwei neue Genvarianten identifizieren, die zu einer erblichen Disposition für diese Krankheit beitragen.[58] 23andMe wirbt damit, dass jeder, der seine Daten für solche Forschungsprojekte freigibt, zu Hunderten von Studien beiträgt und dass dieser Beitrag den Fortschritt in der Forschung deutlich begünstigt.[59] Es gibt aber auch Kritik: Die Verwendung der gesammelten genetischen und ergänzenden gesundheitsrelevanten Daten als Firmeneigentum könne die Forschung oder die Nutzung von Forschungsergebnissen auch behindern, nämlich dann, wenn 23andMe über die künftige Verwendung bzw. Weitergabe in erster Linie nach finanziellen Gesichtspunkten entschiede.[60]

Dank der stark gesunkenen Sequenzierungskosten spielen zunehmend auch umfangreichere genetische Analysen mit entsprechend größeren Datenaufkommen eine Rolle. Die technischen Kosten einer vollständigen Genomsequenzierung liegen schon heute deutlich unter 1.000 Euro und werden weiter fallen.[61] Schätzungen zufolge werden im Jahr 2025 bereits 100 Millionen bis zwei Milliarden menschliche Genome sequenziert sein.[62] Dazu gehören sogenannte Exom-Analysen, bei denen nicht nur ausgesuchte Genvarianten, sondern sämtliche Gene der untersuchten Person analysiert werden, oder sogar Gesamtgenomanalysen, die auch jene Abschnitte der Erbsubstanz mit einbeziehen, die keine Proteine codieren. Solche Untersuchungen eröffnen

[58] Vgl. Do et al. 2011.
[59] Vgl. https://www.23andme.com/en-int/research [17.10.2017].
[60] Vgl. Seife 2013.
[61] Bei Veritas kostet die Genomsequenzierung derzeit beispielsweise 999 Dollar (vgl. https://www.veritasgenetics.com/mygenome [30.08.2017]). Illumina gibt an, dass mit ihren Geräten zukünftig Genomsequenzierungen für 100 Dollar möglich werden (vgl. https://www.illumina.com/company/news-center/press-releases/press-release-details.html?newsid=2236383 [30.08.2017]).
[62] Vgl. Stephens et al. 2015. Zum Vergleich: Nach Angaben des Präsidenten von Illumina, der Herstellerfirma der Maschinen zur Genomsequenzierung, Francis deSouza werden bis Ende 2017 insgesamt etwa 1,6 Millionen Genome sequenziert sein (vgl. Regalado 2017).

größere Chancen, biologische Regulationsmechanismen zu verstehen, die zum Beispiel zu einer Tumorentwicklung beitragen können. Denn erfasst werden nicht nur einzelne molekulare Marker, sondern auch Kopienzahlveränderungen, strukturelle Variationen, die zur Entstehung von Fusionsproteinen führen können, und Veränderungen in nicht codierenden Bereichen, die sich spezifisch im Tumor finden.

Neue Behandlungsmöglichkeiten für Hirntumore durch Gesamtgenomanalysen

Eine Studie an Kindern mit bestimmten Hirntumoren, in der das gesamte Genom von bislang etwa 400 Tumorproben sowie von Blutproben sequenziert wurde, ergab zum Beispiel, dass in fast allen Hirntumorpatienten eine Veränderung in einem molekularen Signalmechanismus auftrat, der zentral an der Wachstumskontrolle von Zellen beteiligt und bereits gut wissenschaftlich untersucht ist.[63] Daher steht jetzt eine Reihe von potenziellen neuen Medikamenten zur Behandlung dieser Tumorerkrankung zur Verfügung, die insbesondere auch in Kombination wirksam sein könnten und künftig für betroffene Kinder eine schonendere Behandlung ermöglichen könnten. Die Rohdaten wurden im European Genome-phenome Archive (EGA) gespeichert und sind auf Anfrage allen Konsortiumsmitgliedern und Patienten zugänglich. Dieses Beispiel zeigt auch, dass Big-Data-Analysen in der medizinischen Forschung schnell Auswirkung auf die medizinische Praxis haben können (siehe Abschnitt 2.5.2).

Bei vielen Erkrankungen sind die Zusammenhänge jedoch sehr komplex und Veränderungen im Genom spielen nur eine begrenzte Rolle. Um Diagnose, Therapie und Prävention tatsächlich langfristig zu verbessern, müssen umfassendere Datensammlungen angelegt werden, die verschiedene Informationen integrativ zusammenfassen. Big Data eröffnet hierfür große Chancen, da solche umfangreichen und quellübergreifenden Analysen erstmals im großen Stil möglich werden. Ergänzend zu genetischen Daten können beispielsweise weitere Omik-Daten (zum Beispiel zum Transkriptom, Proteom, Metabolom, siehe Abschnitt 2.4.2) mit Bilddaten und klinischen Daten kombiniert werden. Für diese Integrationsleistung ist neben der bloßen Menge der einbezogenen Daten auch die Qualität ihrer interpretatorischen Aufbereitung von entscheidender Bedeutung. Die integrative Analyse vielfältiger gesundheitsrelevanter Daten mithilfe systembiologischer Methoden, die dank maschinellen Lernens (siehe Abschnitt 2.3.2) immer leistungsfähiger werden, erlaubt es, biologische und medizinische Zusammenhänge besser zu verstehen und auf dieser Grundlage sehr viel genauere Diagnosen und Prognosen zu stellen.

[63] Vgl. Jones et al. 2013.

Gegenwärtig werden weltweit große Anstrengungen unternommen, mittels öffentlicher Förderung und massiver privatwirtschaftlicher Investitionen Datensammlungen mit Patienteninformationen in bisher nicht gekanntem Umfang anzulegen. Die britische UK Biobank etwa registrierte seit 2007 rund 500.000 Teilnehmer und bietet Forschern die Gelegenheit, mit den Datensätzen wissenschaftlichen Fragestellungen nachzugehen.[64] Die deutsche „NAKO Gesundheitsstudie" (ehemals Nationale Kohorte) begleitet seit 2014 etwa 200.000 Probanden, um die Ursachen der Entstehung von Volkskrankheiten wie Krebs, Demenz, Diabetes, Infektionskrankheiten und Herz-Kreislauf-Erkrankungen zu erforschen.[65] Das EGA speichert und verwaltet Omik-Daten, insbesondere Genomdaten; derzeit umfasst die Sammlung mehr als 3.500 Datensätze aus über 1.500 Studien[66]. Die Alzheimer's Disease Neuroimaging Initiative (ADNI)[67] und die „1000 Gehirne-Studie"[68] ermöglichen Forschern derweil Zugang zu umfangreichen Daten rund um das Gehirn, seine Alterung und neurodegenerative Prozesse.

Ziel der integrativen Datensammlungen ist es, durch breit angelegte Vergleiche krankheitsrelevante Veränderungen zu identifizieren. Wegen der hohen Variabilität und Komplexität der Zusammenhänge, die zur Entstehung von Krankheiten beitragen, sind für valide Analysen besonders große Datensätze von Patienten und geeigneten Kontrollgruppen unerlässlich. Die dafür meist notwendige Zusammenführung von Daten, die von mehreren Institutionen in oft unterschiedlichen Kontexten erhoben werden, bringt besondere Herausforderungen für den Einsatz von Big Data in der medizinischen Forschung mit sich, insbesondere was die Standardisierung und Qualität von Daten, den Datenschutz und den Zugang zu diesen Daten betrifft.

Während in anderen Bereichen wie zum Beispiel der Touristikbranche einheitliche Datenformate bereits seit Längerem erfolgreich genutzt werden, wurde dies bisher im Gesundheitssektor lediglich für einzelne Teilbereiche verfolgt. Ein Beispiel hierfür ist der internationale Datenstandard HL7 (Health Level Seven), der vor allem Spezifikationen für die Darstellung medizinischer Daten und Informationen liefert und damit die Kommunikation zwischen Institutionen und Bereichen des Gesundheitswesens ermöglichen soll. Tatsächlich wird dieser Datenstandard lediglich innerhalb von Krankenhäusern eingesetzt, aber nicht für den Datenaustausch zwischen dem klinischen und dem niedergelassenen Sektor genutzt, da in diesem Bereich andere Datenaustauschformate Verwendung finden. Gemeinnützige Organisationen wie das Clinical Data Interchange Standards Consortium (CDISC) setzen sich für die Einführung nationaler und internationaler Standards für den Datenaustausch im Bereich klinischer Studien

[64] Siehe http://www.ukbiobank.ac.uk/participants [17.11.2017].
[65] Siehe http://nako.de/studienteilnehmer/das-untersuchungsprogramm [17.11.2017].
[66] Siehe https://ega-archive.org [17.11.2017].
[67] Siehe http://adni.loni.usc.edu [17.11.2017].
[68] Siehe http://www.fz-juelich.de/inm/inm-1/DE/Forschung/1000_Gehirne_Studie/1000_Gehirne_Studie_node.html [17.11.2017].

ein und konnten damit bereits Teilerfolge erzielen. So werden die von CDISC entwickelten Datenstandards in den USA zum Beispiel von der Food and Drug Administration (FDA) für die Einreichung klinischer Daten bei der Arzneimittelzulassung akzeptiert.

Dennoch werden nach wie vor im Gesundheitssektor wissenschaftliche und klinische Daten überwiegend in ganz verschiedenen, zum Teil institutseigenen Programmen erfasst. Eine Vereinheitlichung der Schnittstellen und der Übertragung von Daten und assoziierten Metadaten würde den Vergleich und die neue Kombination von Daten erleichtern und somit neue Behandlungsmöglichkeiten eröffnen. Im Idealfall sollten Daten über verschiedene Einrichtungen, Kliniken, Forschungseinrichtungen hinweg miteinander kombinierbar sein. Das erfordert jedoch eine Abstimmung der verwendeten Datenformate und Protokolle. Ferner sollte die Qualität der Daten ständig geprüft und die Auswertung kontinuierlich verbessert werden. Hierfür ist eine genaue Dokumentation der Herkunft der Daten, ihrer Annotation sowie der Verarbeitungsschritte (*provenance tracking*) von essenzieller Bedeutung. In diesem Rahmen werden zum Beispiel graphenbasierte Metadatenbanken, sogenannte Knowledge-Spaces, entwickelt, die umfassende semantische Suchvorgänge ermöglichen (im Unterschied zur einfachen Suche nach Schlüsselwörtern). Das setzt die sorgfältige und aufwendige Kuration der Daten voraus – ein anspruchsvoller Prozess, in dem die Art der Daten genau beschrieben werden muss. Sind die Daten auf diese Weise umfassend beschrieben, ist eine wesentliche Voraussetzung für eine Reproduzierbarkeit von Analysen erfüllt.[69]

Gerade bei kollaborativen Großprojekten ist es darüber hinaus wichtig zu klären, wer unter welchen Bedingungen Zugang zu den Daten erhält, um einen transparenten Austausch bei gleichzeitiger Gewährleistung hoher Datenschutzstandards zu ermöglichen. Es bedarf auch klarer Vorgaben, unter welchen Umständen Probanden und Patienten Zugang zu ihren Daten haben können, wie sie modular und dynamisch in ihre Nutzung einwilligen können (siehe 4.1.2) und wie die Daten langfristig erhalten bleiben. In der Praxis fehlt es häufig noch an gut funktionierenden Regeln für den Austausch. Das liegt zum einen an Datenschutzbedenken. Viele Gesundheitsdaten, etwa Bilddaten vom Gehirn, sind für jeden Menschen einzigartig. So besteht die Möglichkeit, Datensätze Personen zuzuordnen. Diese Möglichkeit wird auch nicht durch technische Verfahren wie zum Beispiel *skull stripping* unmöglich gemacht, bei dem man eine Trennung der Bilddaten des Gehirns von denen der umgebenden Knochen, Haut und Bindegewebe und damit vom Gesicht der gescannten Person durchführt. Die Sorge um eine Identifizierbarkeit ist einer der Gründe, warum derartige Daten besonders restriktiv gehandhabt wer-

[69] Vgl. Livingston et al. 2013 sowie Davis-Turak et al. 2017.

den. Der Datenaustausch wird in der Praxis zusätzlich dadurch erschwert, dass geeignete Kontaktaufnahmemöglichkeiten und Einwilligungsmodelle für Patienten und Probanden zur Sekundärnutzung der Daten fehlen.

Ein sich aus solchen Bedenken und Schwierigkeiten ergebender sehr restriktiver Umgang mit Daten erschwert wissenschaftlich sinnvolle Bemühungen, mithilfe von Big-Data-Analysen über Institutionen, Bundesländer- und nationale Grenzen hinweg zum Beispiel auch schwach wirkende und miteinander interagierende Faktoren für bestimmte Erkrankungen zu identifizieren oder seltene Erkrankungen[70] zu untersuchen. Gerade beim grenzüberschreitenden Austausch sind hohe bürokratische Hürden zu überwinden. Das bereits erwähnte EGA etwa hat neben einem öffentlich frei zugänglichen Bereich auch einen Bereich mit beschränktem Zugang, der durch derzeit über 400 Datenzugriffsausschüsse geregelt wird.[71] Dadurch entstehen interessierten Forschern erhebliche Verwaltungskosten und Verzögerungen bei weiterführenden Analysen.

Lösungsansätze bieten neben neuen Modellen der Einwilligung in die Datennutzung (siehe Abschnitt 4.1.2) auch technische Verfahren, die einen unkomplizierten Datenaustausch bei gleichzeitiger Minimierung der Identifizierungsrisiken erlauben. So besteht zum Beispiel die Möglichkeit, Analysen von Bilddaten lokal durchzuführen, bestimmte Merkmale aus Bilddaten zu extrahieren und nur diese zu exportieren (*in situ querying*), aber sie dennoch mit denen anderer Einrichtungen zu verknüpfen.[72] Auch kann man mit Gruppendaten oder künstlich „verrauschten" Daten arbeiten, die den Rückschluss auf das Individuum stark erschweren. Gleichwohl hängt die Frage nach den geeigneten Instrumenten von den jeweiligen Anforderungen ab. So kann man unter Umständen sogar im Nachhinein Personen identifizieren, wenn und weil die Daten Rückschlüsse auf eine akute Erkrankung oder gar einen lebensbedrohlichen Zustand erlauben. Die genannten technischen Verfahren setzen zudem umfangreiche methodische Entwicklungen im Bereich Datenbanken, Annotation oder Metadatengewinnung voraus. Diese bedürfen auch institutioneller Unterstützung, wie sie zum Beispiel durch das neu gegründete Center for IT-Security, Privacy and Accountability (CICSA) der Universität des Saarlandes gewährleistet werden soll.

Nicht nur Datenschutzbedenken, sondern auch Unsicherheit und unterschiedliche Vorstellungen darüber, wer in welchem Ausmaß das Recht hat, über die generierten Daten zu verfügen, bringen Herausforderungen für den Datenaustausch mit sich. Obwohl dies rechtlich unzutreffend ist (siehe Abschnitt 3.2.2), „gehören" aus der Perspektive vieler Forscher im Labor und in der Klinik die Daten zunächst demjenigen, der die Messungen durchführt bzw. den Auftrag

[70] Vgl. Nationales Aktionsbündnis für Menschen mit Seltenen Erkrankungen 2016, 14 f.
[71] Vgl. https://www.ebi.ac.uk/ega/dacs [07.11.2017].
[72] Zum Beispiel Frackowiak/Ailamaki/Kherif 2016.

dazu erteilt hat. Folglich möchte diese Person in der Regel selbst entscheiden, wem sie Zugang zu den Daten gewährt. Auch Institutionen betrachten Daten häufig als ihr „Eigentum", unter anderem deshalb, weil die Daten unter erheblichem Einsatz finanzieller und personeller Ressourcen generiert wurden. Zudem ist die exklusive Publikation von auf den jeweiligen Daten aufbauenden Studien häufig mit Anerkennung für den Forscher bzw. die Institution, aber auch mit der Allokation von Forschungsmitteln verbunden. Bei Firmen spielen zudem kommerzielle Motive eine Rolle, die einer freizügigen Datenweitergabe entgegenstehen mögen.

Gerade bei ganz oder teilweise öffentlich finanzierten Projekten und Anwendungen können solche Bestrebungen, Daten exklusiv zu nutzen, allerdings zu Konflikten führen. So wird ein signifikanter Teil der Forschung und insbesondere der Forschungsinfrastruktur öffentlich finanziert, und viele Bürger tragen mit ihren Daten dazu bei, die „Datenschätze" überhaupt erst aufzubauen. Auch manche tragbaren Medizingeräte sammeln viele Gesundheitsdaten, deren Nutzung für die Forschung die Gerätehersteller mitunter trotz einer Finanzierung der Geräte durch das Solidarsystem verwehren.[73] Viele Forscher reklamieren daher unter Verweis auf ein gesamtgesellschaftliches Interesse, an den Erkenntnissen und Gewinnen, die aus den beschriebenen Datenströmen erwachsen, beteiligt zu werden, ein Mitspracherecht bei der Nutzung von in solchen Kontexten gewonnenen Daten.

Den genannten Herausforderungen wird mit unterschiedlichen Lösungsansätzen begegnet. Während einige Staaten viele Gesundheitsdaten schon jetzt zentralisieren (zum Beispiel Griechenland, skandinavische Länder) oder einigen Institutionen einen recht offenen Umgang mit Forschungsdaten ermöglichen (zum Beispiel UK Biobank), bleibt Deutschland in Bezug auf den Datenaustausch bislang restriktiver. Das deutsche E-Health-Gesetz sieht künftig allerdings eine Öffnung der Telematikinfrastruktur, die derzeit zur besseren und sicheren Vernetzung der Akteure im Gesundheitswesen entwickelt wird, für die Gesundheitsforschung vor. Ebenso steht eine stärkere Verzahnung von im Gesundheitssystem gesammelten Daten und solchen der biomedizinischen Forschung im Mittelpunkt einer Studie im Auftrag des Bundesgesundheitsministeriums zu Weiterentwicklung der deutschen E-Health-Strategie.[74] Das BMBF hat zudem 2017 eine Medizininformatik-Initiative[75] initiiert, deren Ziel es ist, unterschiedliche Datensätze sowie Speicher- und Analysemethoden in einer nationalen Infrastruktur zu verknüpfen. Die Medizininformatik-Initiative setzt zunächst auf Universitätskliniken und ihre Partner, die eine enge Verbindung zwischen Krankenversorgung und klinischer Forschung pflegen. Weitere

[73] In die Kritik geraten ist in diesem Zusammenhang beispielsweise der Medizinproduktehersteller Medtronic, der die von seinen weitverbreiteten Blutzuckersensoren und anderen Geräten gewonnenen Daten externen Forschern und teilweise sogar den Nutzern selbst nicht zur Verfügung stellt (vgl. Wilbanks/Topol 2016, 347).
[74] Vgl. Strategy&/PricewaterhouseCoopers 2016.
[75] Siehe http://www.medizininformatik-initiative.de [17.10.2017].

Forschungsinstitute, Hochschulen, private Kliniken und Unternehmen aus relevanten Branchen wie IT, Pharma, Biotechnologie und Medizintechnik sollen künftig ebenfalls einbezogen werden.

Daneben gibt es auch gemeinnützige Initiativen, bei denen Bürger ihre Gesundheitsdaten zur Weitergabe an Forscher unter Einhaltung von Sicherheitsstandards (und Gewährung einer verlässlichen Opt-out-Regelung) autonom und aktiv in Datenbanken einspeisen können, mit dem Ziel einer effizienten Verbesserung sowohl der globalen als auch der individuellen Gesundheitsverhältnisse.[76] Beispiele hierfür sind die US-amerikanische Open-Science-Organisation Sage Bionetworks[77], die bereits mehrere Angebote zur kollaborativen Forschung entwickelt hat, und die Schweizer Kooperative MIDATA[78], die Bürgern anbietet, über gemeinnützige Genossenschaften Repositorien zur sicheren Speicherung, Verwaltung und Weitergabe von verschiedenen gesundheitsbezogenen und anderen persönlichen Daten zu nutzen. Mitunter ist sogar ein ganz offener Zugang zu Daten möglich. Open Targets[79] etwa, eine Gemeinschaftsinitiative öffentlich geförderter Institutionen und Firmen, zielt darauf ab, eine Plattform zu etablieren, um therapeutische Zielstrukturen durch genomweite Untersuchungen zu validieren und die Daten öffentlich frei zur Verfügung zu stellen.

2.5.2 Big Data in der Gesundheitsversorgung

Der Gesundheitssektor wird durch eine Vielfalt von Akteuren mit teilweise divergierenden Interessen geprägt.[80] Dazu gehören beispielsweise die Erbringer, Kostenträger (insbesondere gesetzliche und private Krankenversicherungen) und natürlich auch die Empfänger von Gesundheitsleistungen, aber auch Behörden, Interessenverbände und forschende Akteure mit einem unmittelbaren Bezug zur medizinischen Praxis. Unter den Leistungserbringern sind mit Blick auf Big Data insbesondere öffentliche und private Kliniken mit ihren Fachabteilungen und IT-Bereichen relevant. Sie nehmen verschiedene Funktionen der Generierung, Nutzung und Administration großer Datenmengen wahr. Auch für niedergelassene Ärzte und ihre Mitarbeiter sowie für Apotheker und andere Gesundheitsdienstleister wird der Umgang mit großen Datenmengen aus unterschiedlichen Quellen zunehmend relevant (zum Beispiel Zugriff auf spezialisierte Datenbanken).[81] Als technische Grundlage eines erleichterten Datenaustausches verlangt das E-Health-Gesetz die Einrichtung systemneutraler, interoperabler Schnittstellen, die auch

[76] Vgl. Wilbanks/Topol 2016.
[77] Siehe http://sagebase.org [17.10.2017].
[78] Siehe https://www.midata.coop [17.10.2017].
[79] Siehe https://www.opentargets.org [17.10.2017].
[80] Vgl. Deutscher Ethikrat 2016, 8 f.
[81] Über den Westdeutschen Teleradiologieverbund unter dem Dach der MedEconTelemedizin GmbH etwa tauschen schon heute Kliniken und Arztpraxen pro Monat etwa 35.000 Untersuchungsergebnisse zu Patienten untereinander aus (vgl. https://www.medecon-telemedizin.de/news/westdeutscher-teleradiologieverbund-nimmt-auf-conhit-den-300-teilnehmer-ins-visier [17.10.2017]). Im Projekt FALKO.NRW, das 2016 gestartet wurde, soll der Datenaustausch weiter verbessert und die Vernetzung der derzeit ca. 300 Einrichtungen verdichtet werden (siehe https://falko.nrw).

für industrielle Anbieter und wissenschaftliche Einrichtungen zugänglich sein sollen. Ausdrücklich sollen auch Erbringer telemedizinischer Leistungen einbezogen werden (vgl. § 291d-e SGB V).

Weitere Akteure aus dem Bereich des Gesundheitswesens sind die öffentlich-rechtlichen Körperschaften der ärztlichen Selbstverwaltung, aber auch privatrechtlich organisierte Verbände wie Fachgesellschaften und Interessenvertretungen, zum Beispiel der 2012 gegründete Bundesverband Internetmedizin. Diese Akteure sind im Rahmen ihrer gesetzlichen Zuständigkeiten bzw. selbst gewählten Aufgabenstellungen ebenfalls mit der Zusammenführung, Verwaltung, Aufbereitung und Zurverfügungstellung von Daten betraut. Zunehmend übernehmen sie außerdem Aufgaben der Weiterentwicklung und datenschutzrechtlichen Begleitung internetbasierter Gesundheitsangebote. Patientenorganisationen hingegen verfolgen in erster Linie das Ziel, die Interessen der Patienten zu wahren und deren zuverlässige Versorgung zu gewährleisten. Sie haben daher ein Interesse an bestmöglicher Versorgung, aber auch an Datensicherheit, möglichst zuverlässiger Anonymisierung sowie einer zustimmungsfähigen Verwendung persönlicher Gesundheitsdaten. Hinzu kommen Aufgaben der Beratung von Patienten zu medizinischen Leistungen. In einer Grauzone zwischen unabhängiger Patientenberatung und Werbung durch Gesundheitsdienstleister bewegen sich schließlich kommerzielle Angebote etwa der Ärztebewertung.

Schließlich erheben und nutzen auch diverse Akteure aus den Bereichen der Gesundheitspolitik, Gesundheitsverwaltung und anderen mit der öffentlichen Gesundheit befassten Institutionen mitunter Daten in großen Mengen, etwa um Prognosen von langfristigen Entwicklungen bei altersassoziierten Erkrankungen zu erstellen oder Zeitpunkt und Ausmaß von Grippewellen[82] zu antizipieren. Zu den beteiligten Organisationen gehören etwa Ministerien, Gesundheitsämter oder das mit der Überwachung und Prävention von Krankheiten befasste Robert Koch-Institut, aber auch andere Einrichtungen wie die Bundeszentrale für gesundheitliche Aufklärung, der Gemeinsame Bundesausschuss oder das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Die Landesämter für Statistik, das Statistische Bundesamt oder Eurostat, das Statistische Amt der Europäischen Union, erheben, sammeln und analysieren Daten auch zu einer Vielzahl von Gesundheitsthemen und stellen die gewonnenen Informationen anderen Ämtern, politischen Entscheidern, wirtschaftlichen Akteuren und der Bevölkerung zur Verfügung.

Neben Verbesserungen in der Erforschung von Krankheiten, die man sich von einem Big-Data-basierten, präziseren Verständnis der Krankheitsentstehung und -behandlung erhofft, eröffnet

[82] Vgl. http://www.flu-prediction.com [17.10.2017].

der Einsatz von Big Data auch Chancen in der medizinischen Praxis, insbesondere durch stärker personalisierte Behandlungskonzepte sowie Effektivitäts- und Effizienzsteigerungen. Experten erwarten, dass Algorithmen die Diagnostik in vielen Bereichen drastisch verbessern und genauer sein werden als Ärzte.[83] Fortschritte im Bereich künstlicher wahrnehmender Systeme eröffnen beispielsweise neue Möglichkeiten der (teil-)automatischen Auswertung von medizinischen Bildern. Die medizinische Bildverarbeitung ist inzwischen eine wichtige Ingenieursdisziplin, die durch Fortschritte, auch in der Echtzeitfähigkeit der Verfahren, in Zukunft voraussichtlich innovative Diagnose- und Behandlungsverfahren ermöglichen wird – etwa in der computerassistierten Chirurgie, bei der auf Grundlage von umfassenden Bilddaten ein präzises Modell der zu operierenden Region erstellt wird, das bei der Planung des Eingriffs und der Navigation während der Operation zum Einsatz kommen kann. Ebenso werden neue Möglichkeiten für die Überwachung von Krankheiten beispielsweise durch Telemedizin, die Minimierung von Risiken schädlicher Medikamenteninteraktionen sowie die Vereinfachung des Informationstransfers bei Arztwechsel oder Behandlung durch mehrere Gesundheitsdienstleister geschaffen. Das ist gerade bei komplexen oder seltenen Erkrankungen von Bedeutung.[84]

Der Rückgriff auf große relevante Datenmengen ermöglicht auch eine bessere Stratifizierung von Patienten, sodass Nebenwirkungen reduziert werden und unnötige Therapieversuche unterbleiben können.[85] Damit verbindet sich die Erwartung, Kosten für das Gesundheitssystem zu senken und die Therapieentwicklung, etwa in der pharmazeutischen Industrie, zu beschleunigen. Außerdem sollen Patienten schneller von Therapiekonzepten profitieren, die bei anderen Patienten bereits erfolgreich eingesetzt wurden.[86] Eine optimierte Früherkennung kann die Prognose verschiedener Erkrankungen verbessern. Durch den Einsatz von persönlichen datensammelnden Geräten und Gesundheits-Apps kann das Gesundheitsverhalten gefördert werden, wenngleich auch negative Konsequenzen wie enge Fokussierung auf die Selbstvermessung körperlicher Zustände und unnötige Behandlungen denkbar sind (siehe Abschnitt 2.5.5). Die Sammlung und Auswertung gesundheitsbezogener Daten bietet insofern ein erhebliches präventives Potenzial, als für bestimmte Anlagenträger gruppenspezifische Risiken besser vermieden werden können. Für den Staat eröffnen die genannten Möglichkeiten von Big Data zudem Chancen auf eine bessere Bedarfsplanung des Gesundheitssystems, eine höhere Effektivität und Effizienz und damit eine Senkung der Kosten für die Haushalte.

---

[83] Vgl. Obermeyer/Emanuel 2016.
[84] Am Nationalen Centrum für Tumorerkrankungen in Heidelberg wurden etwa bereits mehr als 550 Patienten mit Tumorerkrankungen rekrutiert, bei denen im Rahmen einer gemeinsamen Auswertung von zum Beispiel Gesamtgenomuntersuchungen und Transkriptomanalysen relevante Zusammenhänge zwischen Krankheitsverläufen und genetischen Profilen aufgedeckt und damit neue und personalisierte Therapieoptionen entwickelt werden sollen (vgl. Horak et al. 2017).
[85] Vgl. Fachforum „Digitalisierung und Gesundheit“ im Hightech-Forum 2017, 14.
[86] Vgl. das Interview mit Hasso Plattner im Verlagsspezial der Frankfurter Allgemeinen Zeitung vom 17. April 2015: http://www.anna-seidinger.com/pdf/FAZ_MedizinZwischenMoeglichkeitenUndErfolg_2015.pdf [17.10.2017].

Den Chancen datenintensiver Ansätze stehen allerdings auch Risiken gegenüber. Diese liegen etwa im tatsächlichen oder wahrgenommenen Verlust der Kontrolle über die eigenen Daten und den immer weniger begrenzten Zugriff auf intime Informationen durch Leistungsanbieter („gläserner Patient“) sowie im dadurch erleichterten Missbrauch der Daten durch unberechtigte Dritte (Versicherungen, Arbeitgeber, Medien; Letztere vor allem bei Personen öffentlichen Interesses). Hinzu kommt die Sorge, dass eine verstärkte Nutzung Big-Data-gestützter Ansätze die persönliche Zuwendung zum Patienten weiter reduzieren könnte, weil Arzt und Patient mehr und mehr mit bzw. über Maschinen statt direkt miteinander kommunizieren. Ebenso besteht durch einen unkritischen oder unsachgemäßen Einsatz von Big-Data-basierten Handlungsempfehlungen die Gefahr von Fehldiagnosen.

Dort, wo entsprechende Daten dem Staat zur Verfügung stehen, könnten durch sektorenübergreifenden Datenaustausch etwa zwischen Gesundheits-, Sozial- und Arbeitsbehörden dem Bürger Risiken einer unangemessenen Überwachung entstehen. Darüber hinaus könnte ein System aus zusammengeführten umfassenden Daten über jeden Bürger und sein Verhalten sogar zur „Erziehung“ nach den Vorstellungen eines totalitären Staates genutzt werden. Ein Sozialkreditsystem entschiede dann zum Beispiel über den Zugang des Einzelnen zu öffentlichen und privaten Diensten – vom Verkehr über Bildung bis zu Krediten und Versicherungen. Erste Beispiele für solche Entwicklungen gibt es bereits.[87] Dass solche Praktiken hierzulande verfassungswidrig wären, liegt auf der Hand. Doch schon die Existenz solcher Möglichkeiten ist ein hinreichender Anlass zur Aufmerksamkeit.

2.5.3 Nutzung gesundheitsrelevanter Daten durch Versicherer und Arbeitgeber

Die Prüfung des aktuellen Gesundheitszustands sowie der Krankengeschichte gehört zur Standardpraxis vor bestimmten privatwirtschaftlichen Entscheidungen, bei denen Risikoabschätzungen über die künftige gesundheitliche Entwicklung einer Person eine besondere Rolle spielen. Das gilt zum Beispiel für Vertragsabschlüsse (Lebensversicherungen, Berufsunfähigkeitsversicherungen, private Krankenversicherungen, Kreditverträge). Entsprechend prüfen auch öffentliche Stellen die gesundheitliche Eignung von Bewerbern um den Beamtenstatus. Die Kassen der gesetzlichen Krankenversicherung (GKV) hingegen sind zwar verpflichtet, Beitrittswillige ohne vorherige Gesundheitsprüfung aufzunehmen (Kontrahierungszwang); sie haben jedoch nicht anders als Unternehmen der PKV ein Interesse an einer gewissen gesundheitlichen Überwachung ihrer Versicherten, um durch rechtzeitigen Rat und Anreize deren Gesundheitszustand möglichst gut und die eigenen Kosten möglichst niedrig zu halten. Aus dem gleichen Grund können auch Arbeitgeber, inklusive Bund, Länder und Kommunen, ein Interesse an Gesundheitsinformationen über ihre Arbeitnehmer bzw. Beamten haben.

[87] Siehe etwa Condliffe 2016 und Dorloff 2017.

Big Data eröffnet hier umfangreiche neue Zugriffs- und Auswertungsmöglichkeiten, die von den geltenden rechtlichen Bestimmungen nicht durchgehend erfasst sind (siehe Kapitel 3). Immer umfangreichere Datenmengen und -verknüpfungen ermöglichen zunehmend individualisierte Profile einzelner Personen oder Personengruppen. Die Profile beziehen sich dabei auf solche Eigenschaften, die für einen bestimmten Zweck als aussagekräftig gelten. Versicherungsverträge und private Kredite gehören zu den Grundlagen individueller Existenzsicherung. Vertragsentscheidungen werden außerhalb der Solidarsysteme von Renten- und gesetzlicher Krankenversicherung überwiegend von privatwirtschaftlich tätigen Unternehmen gefällt. Daraus erwächst für diese eine hohe Verantwortung für den Umgang mit den für die Evaluierung erforderlichen Daten.

Gendiagnostik und Risikoprädiktion

Genetische Informationen erlauben es in manchen Fällen, den Verlauf einer Erkrankung mit einer gewissen Wahrscheinlichkeit vorherzusagen. Dies ist ein wesentlicher Grund dafür, dass deren Verwendung für Arbeits- und Versicherungsverträge durch das Gendiagnostikgesetz (GenDG) besonders restriktiv geregelt ist. Im GenDG hat der Gesetzgeber Versicherern grundsätzlich untersagt, Daten aus genetischen Untersuchungen oder Analysen entgegenzunehmen oder zu verwenden. Ein wesentlicher Grund dafür ist der Schutz vor Diskriminierung. Nun können Daten aus genetischen Untersuchungen und Verhaltensdaten nicht einfach gleichgesetzt werden. Auch die prädiktive und prognostische Verwendung von Verhaltensdaten kann jedoch zu Diskriminierung führen. Der Nationale Ethikrat hat bereits 2007 in seiner Stellungnahme „Prädiktive Gesundheitsinformationen beim Abschluss von Versicherungen“ darauf hingewiesen, dass eine Regulierung allein des Umgangs mit prädiktiven und prognostischen Gesundheitsinformationen genetischer Art den ethisch relevanten Kern des Problems nur unzureichend erfasst.[88]

Allerdings haben sich in jüngster Zeit – das GenDG gilt seit 2010 – vermehrt klinische Konstellationen ergeben, nach denen die Offenbarung genetischer Eigenschaften gegenüber Kostenträgern im unmittelbaren Eigeninteresse des Patienten liegen kann. Bei Anlageträgerschaft für erblichen Brust- und Eierstockkrebs gilt dies beispielsweise für den Zugang zu spezifischen Vorsorgemaßnahmen (Kernspinmammografie, präventive Mastektomie) oder Therapien (Chemotherapie mit PARP-Inhibitoren).

Zu den Grundprinzipien von Versicherungen gegen Lebensrisiken, die der Einzelne aus eigener Kraft oft nicht bewältigen kann, gehört die Risikostreuung über die Versichertengemeinschaft.

[88] Siehe Nationaler Ethikrat 2007. Zu prädiktiven Gesundheitsinformationen bei Einstellungsuntersuchungen siehe Nationaler Ethikrat 2005.

Ein Versicherungsfall führt zur Umverteilung von Ressourcen aus dem Kollektiv der Versicherten an den einzelnen Anspruchsberechtigten. Insofern kann die Quantifizierung individueller Risikoprofile vor Eintritt des Versicherungsfalles und damit eine faire Gestaltung von Versicherungsprämien nicht nur dem Eigeninteresse des Unternehmens, sondern letztlich auch der sozialen Gerechtigkeit dienen (siehe Kapitel 4). Die Verbesserung der Kalkulationsgrundlagen fördert die Wirtschaftlichkeit der Versicherungswirtschaft und kann günstigenfalls zu höherer Kosteneffizienz im Interesse der Beitragszahler führen. Ähnliches gilt für Kreditverträge: Eine Bewilligung nur unter der Voraussetzung hinreichender Bonität schützt das Kollektiv der Kreditnehmer vor zu hohen Ausfallsrisiken und mitunter auch den einzelnen Antragsteller vor Selbstüberforderung.

Für die Entwicklung von Risiko-Scores aus einem komplexen Geflecht von Einzelfaktoren ist Big Data seinem Wesen nach ideal geeignet; in der Versicherungswirtschaft werden bereits intensive Anstrengungen zur Implementierung Big-Data-geleiteter Risikozuordnungen unternommen.[89] Dass Versicherungsunternehmen berechtigt sind, Prämien zum Beispiel für Lebensversicherungen anhand von risikorelevanten individuellen Daten, wie etwa Vorerkrankungen, zu gestalten, ist allgemein akzeptiert. Über solche objektivierbaren und gegenüber abgelehnten bzw. mit Risikoaufschlägen belegten Antragstellern üblicherweise transparenten „harten" Kriterien hinaus kommen aber für die Vertragsgestaltungen auch eine Vielzahl „weicher" Beurteilungsparameter in Betracht. Dazu gehören etwa die Stabilität des sozialen Umfeldes, Konsumgewohnheiten oder risikogeneigtes Freizeitverhalten. Besonders wichtig sind individuelle Gesundheitsrisiken, die zum einen von erheblicher realer Bedeutung, zum anderen aber schwer quantifizierbar und überprüfbar sein können, etwa der frühere Tabakkonsum oder die Befolgung von Gesundheitsvorsorgeempfehlungen. Nicht nur private Versicherer, sondern auch gesetzliche Krankenkassen beginnen zunehmend, solche Faktoren etwa mit Boni für Nichtraucher oder Teilnehmer an Präventionsprogrammen zu berücksichtigen und so eine Praxis der Individualisierung von Risiken zu entwickeln.[90]

Versicherungen betreffen naturgemäß weitgehend unvorhersehbare Ereignisse. Eine durch Big Data mögliche, zu detaillierte Voraussagbarkeit von Lebensverlauf und Lebenserwartung des Versicherten, stellt daher das Geschäftsmodell von privaten Versicherungen infrage. Umgekehrt folgen aus möglichen Wissensasymmetrien in beiden Richtungen Missbrauchsmöglichkeiten im Sinne von Diskriminierung bzw. der Selektion von „schlechten Risiken" (*adverse selection*). So könnte die Analyse kommerziell verfügbarer, Big-Data-generierter persönlicher Verhaltensprofile, eventuell sogar der Zugriff auf Genomdaten aus wissenschaftlichen Kooperationsprojekten – die nicht Gegenstand des Gendiagnostikgesetzes sind – Versicherer in die Lage

[89] Siehe zum Beispiel Hauner 2016.
[90] Vgl. https://www.krankenkassen.de/gesetzliche-krankenkassen/leistungen-gesetzliche-krankenkassen/praevention-vorsorge-krankenkassen [17.10.2017].

versetzen, sich gezielt risikoarme Antragsteller auszuwählen. Umgekehrt könnte ein potenzieller Versicherungsnehmer zunächst durch eine insgeheim im Ausland durchgeführte Genomanalyse seine künftigen Krankheitsrisiken abschätzen lassen, um davon ausgehend seinen individuellen Versicherungsschutz zu gestalten.[91]

Grundsätzlich kann eine qualitative Verbesserung von Mechanismen der Risikostratifizierung durch Big Data auch die Qualität des Risikoschutzes für das Kollektiv von Versicherten bzw. Kreditnehmern erhöhen. Problematisch können hierbei jedoch Szenarien sein, bei denen bei gleichbleibendem Budget der dem einen Versicherten gewährte Bonus gleichbedeutend mit einem Malus für den anderen ist, dem ein solcher Bonus nicht zuteil wird. Dementsprechend obliegt Versicherern, auch innerhalb des Solidarsystems, im Fall von Bonusprogrammen eine rechtliche als auch ethische Begründungslast (siehe Kapitel 4).

Die mit Big Data oftmals verbundene Intransparenz kann auch insofern Probleme verursachen, als sie die Nachvollziehbarkeit von Faktoren verdunkeln kann, die zu einer das Individuum unmittelbar treffenden quantitativen oder qualitativen Allokationsentscheidung in Form eines Risikoaufschlags bzw. zur Ablehnung eines Versicherungs- oder Kreditantrags führen. Dies dürfte es erschweren, argumentativ bzw. auf dem Rechtsweg gegen nachteilige Entscheidungen vorzugehen. Mitunter hat die fehlerhafte Interpretation eines einzelnen Parameters für den Versicherten oder Antragsteller schwerwiegende Konsequenzen. Im Versicherungswesen sind zwar bereits niederschwellig zugängliche Beschwerdemechanismen[92] eingerichtet sowie eine behördliche Aufsicht durch die Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin) sichergestellt; ob hier aber hinreichende Kompetenzen und Kapazitäten vorgehalten werden, den spezifischen Herausforderungen durch Big Data gerecht zu werden, muss derzeit bezweifelt werden.

Auch Entscheidungen von Arbeitgebern in Bewerbungsverfahren sind von erheblicher Bedeutung für den Betroffenen, die ihnen zugrunde liegenden Entscheidungsparameter aber großenteils intransparent. Wenngleich Gesundheitsprüfungen im Bewerbungsverfahren häufig untersagt sind, wird bereits heute häufig ergänzend zum offiziellen Auswahlprozess auf online verfügbare Informationen über Bewerber zugegriffen. Etwa die Hälfte der Unternehmen prüft bereits eventuell vorhandene Profile in sozialen Medien, um sich einen besseren Eindruck vom Charakter des Bewerbers zu verschaffen.[93] Die Möglichkeiten, aus solchen Daten mithilfe von Big-Data-Analysen genuin gesundheitsrelevante Informationen zu extrahieren, nehmen

[91] Vgl. Henn 2016.
[92] Zum Beispiel www.versicherungsombudsmann.de [17.10.2017].
[93] Vgl. https://www.bitkom.org/Presse/Presseinformation/Jedes-zweite-Unternehmen-ueberprueft-Bewerber-in-Sozialen-Netzwerken.html [17.10.2017]; Weitzel et al. 2016.

ebenso wie die Angebote kommerzieller Unternehmen, solche Analysen bereitzustellen, ständig zu.[94] Die rechtliche Beurteilung solcher Ansätze ist derzeit umstritten.[95]

Auch innerhalb bestehender Verträge haben Arbeitgeber und Krankenversicherungen ein Interesse an der Gesundheit ihrer Vertragspartner, da im Krankheitsfall hohe Kosten entstehen können. Die Überwachung des Arbeitnehmer- bzw. Patientenverhaltens lässt Anreize oder Sanktionen für eine gesunde bzw. ungesunde Lebensführung zu. Diese müssen nicht unbedingt über finanzielle Boni oder Mali erfolgen. Sie können auch in der Sammlung von Informationen oder in der Förderung „gesunden" Verhaltens bestehen, etwa wenn Geräte oder Apps zur Messung gesundheitsrelevanter Faktoren bereitgestellt werden oder eine vergleichsweise sportliche Lebensführung angeregt und deren Realisierung erleichtert wird. Hersteller tragbarer Messgeräte kooperieren bereits mit Arbeitgebern und Versicherern.[96]

Solche Programme eröffnen für alle Beteiligten Chancen. Verminderte Krankenstände kommen nicht nur Arbeitgebern und Versicherern zugute, die dadurch Kosten sparen, sondern auch Arbeit- und Versicherungsnehmern, deren Gesundheitszustand durch solche Präventionsprogramme verbessert wird. Auch Mitversicherte profitieren von den eingesparten finanziellen Ressourcen, die für andere Leistungen eingesetzt werden können, und Kollegen werden entlastet, da sie weniger Vertretungen übernehmen müssen.[97] Die Risiken, die mit der Sammlung und Verwertbarkeit großer Mengen gesundheitsrelevanter Daten in diesem Zusammenhang verbunden sind, dürfen gleichwohl nicht ignoriert werden. Prämienanpassungen oder Abmahnungen wegen gesundheitsschädlichen Verhaltens liegen nicht im Interesse der Datengeber. Selbst depersonalisiert erhobene Daten können relativ leicht bestimmten Individuen zugeordnet werden, wenn sie mit weiteren Informationen verknüpft werden, die zum Beispiel von Analysefirmen aus Datenspuren der Nutzer im Internet ermittelt werden können. So nehmen in den USA etwa die Firmen Walmart und Time Warner bereits die Dienste von Gesundheitsanalysefirmen wie Castlight Health in Anspruch, um sich über mögliche Schwangerschaften von Angestellten zu informieren.[98]

---

[94] Vgl. https://business.linkedin.com/content/dam/business/talent-solutions/regional/de-de/c/pdfs/BigDataimPersonalmanagement_LinkedIn_Bitkom.pdf [17.10.2017].
[95] Vgl. Grimm/Maiß 2015.
[96] So beispielsweise der Hersteller des Fitness-Armbands „Fitbit" (vgl. https://www.fitbit.com/de/product/corporate-solutions [17.10.2017]) oder das Startup Soma Analytics, das Apps anbietet, die das Stresslevel der Nutzer messen und die Daten dann an das Personalmanagement weitergeben (vgl. http://www.soma-analytics.com [17.10.2017]).
[97] In Großbritannien werden solche Vorteile mit dem „Britain's Healthiest Workplace" prominent beworben (vgl. https://www.vitality.co.uk/business/healthiest-workplace [17.10.2017]).
[98] Vgl. http://fortune.com/2016/02/17/castlight-pregnancy-data [17.10.2017.

2.5.4 Kommerzielle Verwertung gesundheitsrelevanter Daten durch global agierende IT- und Internetfirmen

Ein großer Teil der Datenerhebung und -verwendung liegt in der Hand von global operierenden Unternehmen wie Facebook, Google (Alphabet), Apple, Amazon oder Microsoft. Diese Konzerne treten in erster Linie als Dienstleister auf, die auf der Grundlage ihres Zugangs zu riesigen Datenmengen und der geeigneten Dateninfrastruktur, Suchmaschinen, interaktive Informationsplattformen und Angebote wie Online-Shopping, aber auch eine breite Auswahl an multifunktionalen Geräten bereitstellen. Teilweise finanzieren sie sich durch personalisierte Bewerbung zielgruppenspezifischer Produkte bzw. durch kontextsensitive Werbung. Dabei werden unterschiedliche Nutzerdaten in großem Stil gesammelt, gespeichert und verwertet. Solchen Unternehmen, die zunehmend auch in Gesundheitsbereichen agieren[99], ist es daher in besonderer Weise möglich, primär gesundheitsrelevante Daten mit zahlreichen anderen Informationen in Verbindung zu setzen.

Sie bieten Software, Hardware, Technologieentwicklung und Online-Dienste für Big-Data-Anwendungen an und stellen datenorientierten Unternehmen und sonstigen Institutionen Systeme, Algorithmen, Geräte und Infrastruktur zur Datenerhebung, Auswertung, Verwaltung und Speicherung zur Verfügung, mit denen Prozesse beschleunigt und verbessert werden sollen, um eine hocheffiziente Nutzung jeweils relevanter Informationen zu gewährleisten. Auch soweit diese Unternehmen oder ihre Produkte bislang nicht auf gesundheitsrelevante Anwendungen spezialisiert sind, spielen sie eine wichtige Rolle im Bereich Big Data und Gesundheit. So hat beispielsweise Amazon Web Services als führender internationaler Anbieter im Cloud-Computing Verträge mit Krankenhäusern, Universitäten und Pharmaunternehmen[100], und Google arbeitet mit der amerikanischen Mayo-Klinik zusammen, um Suchergebnisse zu Krankheiten und deren Symptomen zu verbessern.[101] Da die Datengenerierung oft nutzergetrieben ist, ergibt sich eine positive Rückkopplung (mehr Nutzer – bessere Produkte – noch mehr Nutzer), die zugleich einen Wettbewerbsvorteil für große Unternehmen mit vielen Kunden bzw. Nutzern bedeutet. Hinzu kommen weitere Skaleneffekte, die mit dem technischen Aufwand und dem technologischen Know-how sowie den globalen Aktivitäten dieser Dienste zusammenhängen. Daher sind Daten sowohl im Hinblick auf ihre Qualität als auch auf ihre Relevanz ein wesentlicher Erfolgsfaktor für entsprechende Internetdienste.

Die datengetriebene Optimierung erstreckt sich auf alle kommerziellen Funktionen des Internets. Dazu gehören vor allem die Bereiche des E-Commerce und der Online-Werbung, die für den überwiegenden Teil der Geschäftsmodelle im Internet eine wesentliche Rolle spielen. Weil

---

[99] Beispiele hierfür sind etwa Googles Gesundheitsplattform Google Fit, Googles Schwesterfirmen Verily Life Sciences und Calico, Microsofts Gesundheitsdatenspeicher HealthVault oder Apples ResearchKit und CareKit.
[100] Vgl. https://aws.amazon.com/de/health [17.11.2017].
[101] Siehe https://www.mayoclinic.org/giving-to-mayo-clinic/your-impact/features-stories/google-works-with-mayo-clinic-to-share-health-knowledge [17.11.2017].

Nutzerdaten mittels der hier operierenden „Werbeoptimierungsmaschinerie“ systematisch in Geldwert übersetzt werden, findet man auch eine Kopplung der Werbemärkte an Datenmärkte. Über verschiedene Mechanismen der Nutzeridentifikation (siehe Abschnitt 2.4.1) können Daten verknüpft und angereichert werden, sodass die zielgerichtete Werbung noch präziser ausgewählt werden kann. Dabei erlangen einige wenige Anbieter eine marktbeherrschende Stellung. Eine Ursache hierfür liegt in der Ausbildung von Datenmonopolen, die den Marktzugang für Konkurrenten erschwert.

Es gibt im privatwirtschaftlichen Bereich eine Reihe von gesundheitsrelevanten Anwendungsfeldern, auf denen Tochterfirmen oder enge Partner großer IT-Konzerne agieren. Viele Firmen in den Bereichen Gesundheit, Fitness und Lifestyle arbeiten mit persönlichen Daten, um eine große Bandbreite individueller Angebote zur gesundheitsbezogenen Gestaltung des Alltags bereitzustellen. Dies erfolgt sowohl in Kooperation mit der institutionalisierten Medizin, etwa bei Herstellern von Medizinprodukten, als auch unabhängig davon, zu Zwecken eines unabhängigen persönlichen Gesundheitsmanagements.[102] Nutzer sollen im Tausch gegen ihre Daten ermächtigt werden, durch technische Hilfestellungen in Form von tragbaren Geräten, elektronischen Tagebüchern und Apps ihren jeweils präferierten Lebensstil zu realisieren, Trainings- oder Gesundheitsziele zu erreichen und sich mit anderen Personen zu vernetzen, auszutauschen und zu vergleichen. Darüber hinaus gibt es jedenfalls im Ausland bereits Bestrebungen von Unternehmen, zum Beispiel im Rahmen von Kooperationen auf gesundheitsrelevante Daten in öffentlichen Einrichtungen zuzugreifen.[103]

Die zunehmenden Aktivitäten digitaler Firmen im Gesundheitsbereich können als forschungsförderlich betrachtet werden, da große Internetkonzerne im Vergleich zum öffentlichen Sektor Zugriff auf wesentlich größere Datenmengen haben und mit leistungsfähigeren Analysemöglichkeiten sowie besseren technischen und finanziellen Ressourcen ausgestattet sind.[104] Da manche Unternehmen Datengebern den vollumfänglichen Zugang und die Weitergabe bzw. -verwendung der eigenen Daten verwehren und auch Nutzungsanfragen unternehmensexterner Wissenschaftler abweisen, besteht allerdings die Befürchtung, dass diese Entwicklung aus zwei Gründen den medizinischen Fortschritt hemmt und die Chancen- und Teilhabegerechtigkeit (siehe Abschnitt 4.5) einschränkt: zum einen, weil eine derart intransparente und abgekapselte

[102] Beispielsweise verfügt Google, unter anderem durch den Dienst Google Analytics, über eine Vielzahl von Gesundheitsdaten, die an Pharmaunternehmen zur Verbesserung von Medizinprodukten weitergegeben werden. So arbeitet Google etwa mit dem Biotechnologieunternehmen Biogen an der Verbesserung von Therapien in den Bereichen Multiple Sklerose sowie Alzheimer und entwickelt mit dem Schweizer Unternehmen Novartis eine Blutzucker messende Kontaktlinse (vgl. https://www.medizintechnologie.de/infopool/politik-wirtschaft/2015/der-angriff-der-giganten [17.10.2017]). Für das persönliche Gesundheitsmanagement bieten IT-Konzerne zahlreiche Apps und Wearables an. So kann beispielsweise die Apple Watch mithilfe von Sensoren die Herzfrequenz, den Puls und den Blutdruck des Trägers analysieren. Fitness-Tracker zeichnen Schritte, verbrannte Kalorien und zurückgelegte Strecken der Nutzer auf. Ernährungs-Apps, wie beispielsweise eatSimply zählen, nach Angabe der Körpermaße, Kalorien und stellen individuelle Ernährungspläne zusammen.
[103] Vgl. McGoogan 2017.
[104] Vgl. Wilbanks/Topol 2016.

Datenanalyse einer kritischen Reflexion weitgehend entzogen würde, und zum anderen, weil die enorme Marktmacht dieser Unternehmen dazu führte, eine methodische Pluralität der Untersuchung und Interpretation von Gesundheitsdaten zu vereiteln.

2.5.5 Erhebung gesundheitsrelevanter Daten durch Betroffene selbst

Der Erfolg datenbasierter Unternehmen und Anwendungen einschließlich der Forschung hängt entscheidend von der Erzeugung und teilweise auch gezielten Erhebung von Daten durch Individuen ab. Bürger bzw. Patienten zeichnen immer mehr unterschiedliche personen- und gesundheitsbezogene Daten über Sensoren in mobilen Endgeräten auf und speichern, verwalten und teilen sie online. Solche mobilen Gesundheitsanwendungen (M-Health) umspannen dabei ein heterogenes Feld von Gesundheits- und Fitness-Apps, das von der Gesundheitsförderung und Prävention bis zur Diagnostik und Therapiekontrolle reicht. So standen bereits im Jahr 2015 weltweit über 380.000 Apps mit Bezug zu Sport, Lifestyle, Ernährung, Medizin, Gesundheit und Fitness zur Verfügung.[105] Neben Smartphones kommen auch mehr und mehr mit Sensoren ausgestattete tragbare Geräte, sogenannte Wearables, zum Einsatz. Schätzungen für das Jahr 2017 gingen von einem weltweiten Verkauf von etwa 310 Millionen Geräten aus.[106] Der Markt zeichnet sich durch eine hohe Dynamik und Unübersichtlichkeit in Bezug auf Anbieter, Produkte, Datenformate und Geschäftsbedingungen aus und ist bis heute weitestgehend unreguliert.

Gesundheits- und Fitness-Apps kommen dabei einerseits im Rahmen einer neuen Bewegung zur Selbstvermessung (Quantified Self, siehe unten) zum gezielten Einsatz, finden andererseits aber auch außerhalb spezifisch interessierter Kreise aufgrund eines gesteigerten Interesses an Themen wie Ernährung und Fitness immer breitere Anwendung.[107] Es gibt zudem Apps und Messgeräte, die im Kontext der Telemedizin auf spezielle Zielgruppen zugeschnitten sind, etwa auf chronisch Kranke oder Heilberufsgruppen. Einen solchen Ansatz verfolgen beispielsweise bereits mehrere Apps und Portale zur Überwachung und Regulierung der Blutzuckerwerte von Diabetes-Patienten.[108] Einige dieser Apps speisen Daten in Systeme der Gesundheitsversorgung ein, zum Beispiel Vitaldaten aus Geräten zur Selbstüberwachung oder Daten aus digitalen Patiententagebüchern, die mit Therapeuten oder Versicherern geteilt werden.

Der Einzelne sieht sich angesichts der neuen Möglichkeiten von Big Data als Bürger, Patient und Versicherter, eventuell als Proband und sicherlich als Internet- und Gerätenutzer sowie als Kunde digital agierender Firmen mit vielfältigen Rollen und Schnittstellen zu anderen Akteuren konfrontiert. Für ihn können die Möglichkeiten und Konsequenzen der Erhebung gesundheitsrelevanter Daten vielfach schwer zu überblicken sein, zumal mitunter – zum Beispiel bei

---

[105] Lucht/Boeker/Kramer 2015, 6.
[106] Vgl. https://www.gartner.com/newsroom/id/3790965 [27.11.2017].
[107] Siehe hierzu beispielsweise den Bericht vom Munich Digital Institute: Zhelyazkova et al. 2017.
[108] Vgl. http://diadigital.de/zertifizierte-apps [17.10.2017] sowie Kitsiou et al. 2017.

Internetsuchen oder in sozialen Netzwerken – kaum klar sein dürfte, inwieweit Daten ausgewertet werden oder gesundheitsrelevant sind oder werden könnten.

Die bewusste Generierung individueller Gesundheitsdaten im Rahmen der sportlichen bzw. allgemein gesundheitlichen Selbstkontrolle, beispielsweise durch Wearables, wird von vielen Bürgern als Element der Förderung von Gesundheit und Wohlbefinden betrachtet. Dabei dürfte vielen Nutzern sicher klar sein, dass diese Daten nicht zwingend im Raum individueller Kontrolle verbleiben, vor allem dann nicht, wenn sie an Ärzte, Trainer, sportliche Mitbewerber und andere weitergegeben werden. Tatsächlich ist es heute beinahe unmöglich, *keine* Daten zu produzieren, die verschiedenen Interessenten zur möglicherweise ökonomischen Verwendung dienen könnten. Datenschutz und informationelle Selbstbestimmung werden, glaubt man statistischen Umfragen, in der Bevölkerung durchaus wertgeschätzt[109], allerdings ist die Digitalisierung der Lebenswelt so weit fortgeschritten, dass alltägliche Verhaltensweisen und Kommunikationsformen auch jenseits sozialer Netzwerke, Lifestyle-Apps etc. eine automatische Datenproduktion nach sich ziehen. Gleichzeitig aber schätzen viele die oben angesprochenen Dienstleistungen und Angebote nicht nur aus Gründen der Zeit- und Geldersparnis, sondern auch wegen der Möglichkeit, trotz räumlicher Distanzen soziale Kontakte zu pflegen.

Ein bereits genannter aktueller Trend, der zur gesundheitsrelevanten Datenflut beiträgt, ist die Bereitstellung und Nutzung von Sensoren und Apps, die immer mehr individuelle Gesundheitsdaten sowie tägliche Aktivitäts- und Umweltdaten erfassen, aufbereiten und mit vorhandenen Datenbeständen verknüpfen können. Solche Sensoren und Apps können zum Beispiel in Mobiltelefonen oder Smartwatches integriert sein, aber auch in eigens konzipierten tragbaren Geräten oder sogar Alltagsgegenständen wie Kleidung, die je nach Ausstattung körperliche Daten wie zum Beispiel Pulsfrequenz, Temperatur, Schlaf- oder Bewegungsmuster sammeln. Die Daten werden üblicherweise in einem größeren Datenspeicher (Cloud) abgelegt und verarbeitet, um später wieder darauf zurückgreifen zu können oder Auswertungen vorzunehmen.

Solche Geräte und Programme können als zeit- und ortsunabhängiger Zugang des Betroffenen zu seinen Gesundheitsinformationen und als Grundlage für eine faktengestützte Gesundheitsversorgung dienen. Sie werden ebenso zur Realisierung eines modernen, gesundheitsbewussten Lebensstils oder zur Förderung des persönlichen Wohlergehens angeboten. Bisweilen sind sie auch, versehen mit Schlagwörtern wie Lifelogging, Self-Tracking oder Selbstvermessung, Ausdruck einer speziellen, als besonders authentisch empfundenen Form von Selbstbestimmung. Dies schließt die Weitergabe der Daten an bestimmte Dritte ein. Dem bestehenden Gesundheitssystem wird mit dem Vorbehalt begegnet, zu wenig auf die aktive gesundheitliche Selbstgestaltung und Selbstverantwortung des Einzelnen zu setzen und die medizinische Behandlung

[109] Vgl. zum Beispiel Deutsches Institut für Vertrauen und Sicherheit im Internet 2017; Europäische Kommission 2017 sowie Mooy 2017.

zu wenig zu individualisieren. Beiden Punkten soll durch Selbstvermessung Rechnung getragen werden.

In der Tat kann die Eigenüberwachung gesundheitsrelevanter Parameter auf der einen Seite zur individuellen Gesundheit und zum individuellen Wohlbefinden beitragen und zu einer positiven Änderung des Lebensstils und des individuellen Selbstbildes führen. Auf der anderen Seite dürfen damit einhergehende Risiken nicht übersehen werden. Überzogene Selbstkontrolle kann ein übertriebenes, der Gesundheit abträgliches Optimierungsstreben sowie die Medikalisierung „natürlicher" Lebensvorgänge befördern. Zudem ist zweifelhaft, ob der Einsatz solcher Technologien zur Selbstvermessung tatsächlich immer Ausdruck persönlicher Souveränität (siehe Abschnitt 4.3) ist. Instrumente einer Selbstvermessung können eine Sogwirkung entwickeln, die auf den Nutzer wie ein innerer Zwang zurückwirkt. So entsteht ein Phänomen, das man als „selbstinduzierte Fremdbestimmung" bezeichnen kann.[110] Befürchtet wird ferner die Diskriminierung von Personen, die sich an solchen Messungen nicht beteiligen können oder wollen.

Werkzeuge zur Selbstvermessung können aber auch zu wissenschaftlichen Zwecken, als wichtige quantitative und qualitative Erweiterung der Datengrundlage mit neuer Detaildichte in der medizinischen und Gesundheitsforschung verwendet werden. So liefert etwa das Apples ResearchKit bereits einen Beitrag zur Epilepsie-Forschung.[111] Zugleich haben die Anbieter und Betreiber solcher Geräte und Apps und andere Unternehmen ein großes wirtschaftliches Interesse an solchen Datensammlungen, um die Daten durch Aufbereitung, Verknüpfung und Generierung neuer Korrelationen für andere zu jedwedem Zwecke nutzen zu können. Die bisherige Orientierung der Angebote an diesen wirtschaftlichen Interessen sowie Mängel bei Nutzerfreundlichkeit, Transparenz und Datenschutz werden kritisiert und eine stärkere Ausrichtung an den Bedürfnissen der Nutzer gefordert.[112] Demnach sollen auch durch technische Maßnahmen Transparenz und Kontrollierbarkeit gestärkt, der unkontrollierbare Zugang zu den Daten eingedämmt und der Bereich der Verwendung überschaubar und der informierten Einwilligung zugänglich gemacht werden.

[110] Lob-Hüdepohl 2007, 84.
[111] Vgl. https://www.apple.com/de/newsroom/2015/10/15Apple-Announces-New-ResearchKit-Studies-for-Autism-Epilepsy-Melanoma [17.10.2017].
[112] Vgl. zum Beispiel Albrecht 2016.

2.6 Zwischenfazit

Die bisherigen Ausführungen haben gezeigt, dass der Einsatz von Big Data in der klinischen Praxis, der medizinbezogenen Forschung und im weiteren gesundheitsrelevanten Bereich bestimmte Stärken und Schwächen sowie Chancen und Risiken mit sich bringt. Diese lassen sich in Anlehnung an die Technik der SWOT-Analyse[113] wie folgt skizzieren:

Stärken	**Schwächen**
1. Vergrößerung und Diversifizierung der Datenbasis sowie Beschleunigung der Informationsgewinnung 2. wechselseitig verstärkte Entwicklung innovativer Instrumente der Datenverarbeitung und erweiterter Datengrundlagen 3. hoher Grad der Vernetzung und ubiquitäre Zugangsmöglichkeiten	1. heterogene Datenqualität 2. Intransparenz von Datenflüssen und Kontrollverluste 3. höherer Aufwand hinsichtlich Koordination, Regulierung und Qualifikationen
Chancen	**Risiken**
1. verfeinerte Stratifizierung bei Diagnostik, Therapie und Prävention auf der Grundlage einer verbreiterten Wissensbasis 2. Effektivitäts- und Effizienzsteigerungen 3. Unterstützung gesundheitsförderlichen Verhaltens	1. Entsolidarisierung und Verantwortungsdiffusion 2. Monopolisierung und Datenmissbrauch 3. informationelle Selbstgefährdung

Die bisherigen Ausführungen haben jedoch auch gezeigt, dass die konkrete Beurteilung von Big-Data-Anwendungen mit Gesundheitsbezug maßgeblich von den jeweils beteiligten Akteuren, mit ihren unterschiedlichen Interessen und eigenen Chancen- und Risikoeinschätzungen sowie dem gesellschaftlichen Kontext abhängt. Zu berücksichtigen ist ferner, wer die jeweiligen Nutznießer bzw. Schadensträger einer Datenerhebung und -verwendung sind, worin der jeweilige Nutzen bzw. Schaden besteht, ob dieser Nutzen als erheblich bzw. der Schaden als erträglich eingestuft werden kann und wie die jeweiligen Wahrscheinlichkeit, dass bestimmte Ereignisse eintreten, abzuschätzen ist.

Die folgenden rechtlichen (siehe Kapitel 3) und ethischen (siehe Kapitel 4) Analysen berücksichtigen sowohl die skizzierten allgemeinen Chancen und Risiken als auch die jeweiligen spe-

[113] Die SWOT-Analyse ist ein situationsanalytisches Instrument, dass vor allem im Marketing, im Management, in der Persönlichkeitsentwicklung, aber auch in der politischen Strategieentwicklung und in der strategischen Analyse im Wissenschaftsmanagement verwendet wird. Im Zentrum der Analyse steht die Gegenüberstellung und Abwägung von Stärken (*strenghts*), Schwächen (*weaknesses*), Chancen (*opportunities*) und Risiken (*threats*) einer gegebenen oder geplanten Situation (vgl. Pelz 2017 und Schröder 2011, 43-50).

zifischen Kontextualisierungen. Für die bilanzierende Abwägung wird dann die verantwortliche informationelle Freiheitsgestaltung des Einzelnen als entscheidender normativer Maßstab entwickelt (siehe Kapitel 5).

3 Rechtliche Vorgaben für Big Data

Die unzähligen, in unterschiedlichen und oft grenzüberschreitenden Bereichen und Sachzusammenhängen gesammelten, zumindest potenziell gesundheitsrelevanten Daten und die zahlreichen Optionen ihres Einsatzes stellen eine erhebliche Herausforderung für das Rechtssystem dar. Gerade im Gesundheitssektor sind die mit Big Data verbundenen Möglichkeiten und Hoffnungen, wie in Kapitel 2 beschrieben, klar erkennbar. Zugleich besteht aber auch ein berechtigtes Misstrauen hinsichtlich der Qualität, Zuverlässigkeit und Aussagekraft der Big-Data-basierten Gesundheitsanwendungen. Die dem bisherigen (Datenschutz-)Recht zugrunde liegende Zielsetzung, Informationsungleichgewichten entgegenzuwirken, wird deshalb unter den Bedingungen von Big Data keineswegs obsolet. Im Gegenteil gilt: „Die immer größer werdenden Datenmengen, die als Folge der Tendenz, Big Data zu nutzen, verfügbar sind und verarbeitet werden, werden diese Asymmetrie im Wissensstand nur verstärken und die Kluft zwischen für die Verarbeitung Verantwortlichen und Nutzern vergrößern."[114] Das verlangt nach einem Rechtsrahmen, der die hinreichende Sicherheit und Zuverlässigkeit der neuen und sich mit großer Geschwindigkeit weiterentwickelnden Anwendungen gewährleistet, gleichzeitig aber auch diesen und den sich aus ihnen ergebenden Chancen hinreichend Raum zur Entfaltung bietet.

Bestandsaufnahme und Analyse der rechtlichen Steuerungsvorgaben haben dabei eine wichtige Unterscheidung zu beachten: die Differenz zwischen verfassungsnormativen und einfachrechtlichen Direktiven. Erstere entfalten Bindungswirkung auch gegenüber der Legislative und können so deren Gestaltungsspielräume einengen. Allerdings steckt das – deutsche wie europäische – Verfassungsrecht lediglich einen relativ offenen Ordnungsrahmen ab. In aller Regel lassen sich ihm keine bis ins Detail gehende Problemlösungsmodelle auf interpretatorische Weise entnehmen. Deshalb wäre es auch ein Trugschluss, das derzeit geltende einfachrechtliche Datenschutzrecht als authentische und „alternativlose" Konkretisierung verfassungsrechtlicher Normen zu deuten (siehe Abschnitt 3.2). Gerade weil erhebliche Gestaltungsspielräume existieren, ist es geboten, das bestehende Regelungsregime des einfachen Gesetzesrechts und seine (Dys-)Funktionalität im Hinblick auf Big-Data-Nutzungen näher zu untersuchen (siehe Abschnitt 3.2). Sachangemessen ist dabei eine Konzentration auf das allgemeine Datenschutzrecht, die speziellen Datenschutzbestimmungen des Gesundheitssektors sowie das Medizinprodukterecht. Darüber hinaus sind die zugrunde liegenden Anreizmechanismen einzubeziehen; namentlich ist darauf zu achten, ob und inwieweit sich Big-Data-basierte Dienste in das System der (privaten und gesetzlichen) Krankenversicherung einfügen.

In einem nächsten Schritt ist diese funktions- bzw. defizitorientierte Analyse zu erweitern um eine Betrachtung der zukünftig sinnvollerweise zu verwendenden Regelungskonzepte (dazu

[114] European Data Protection Supervisor 2015, 11.

Abschnitt 3.3). Schon infolge des gerade für datenbezogene, digitale Dienste auf der Hand liegenden Problems der territorialen Reichweite, aber auch mit Blick auf die massive Veränderungsdynamik des Sektors sind dabei nicht nur klassisch-hoheitliche Handlungsmodi zu berücksichtigen. Vermehrte Aufmerksamkeit verdienen vielmehr auch private (selbstregulative) sowie hybride Steuerungsmechanismen. Allerdings geht es an dieser Stelle primär um abstraktere Beobachtungen und grundlegende rechtstechnische Aussagen zu den unterschiedlichen in Betracht kommenden Instrumenten. Einzelne rechtspolitische Vorschläge sind, die Überlegungen des Rechts- und des Ethikkapitels aufnehmend, im Empfehlungsteil platziert.

3.1 Grundrechtliche Steuerungsdirektiven

Die wesentlichen Bauelemente des Datenschutzrechts sind grundrechtskonstituiert. Die zentrale verfassungsrechtliche Maßstabsnorm auf nationaler Ebene ist dabei das sogenannte Recht auf informationelle Selbstbestimmung, das vom Bundesverfassungsgericht im Volkszählungsurteil aus dem Jahre 1983 als spezifische Ausprägung des allgemeinen Persönlichkeitsrechts gemäß Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG entwickelt worden ist.[115] In weitgehender Parallelität dazu entfalten auf der Ebene der Europäischen Union die Bestimmungen der Art. 7 und 8 der Grundrechtecharta bzw. des Art. 16 Abs. 1 AEUV ihre verfassungsnormative Direktionskraft. Entsprechendes gilt für die auf der Grundlage von Art. 8 EMRK entwickelten Datenschutzgrundsätze.[116]

Bereits in der genannten grundlegenden Entscheidung hat das Bundesverfassungsgericht mit feinem Gespür für die Konsequenzen der technischen Weiterentwicklungen hervorgehoben, aufgrund der „der Informationstechnologie eigenen Verarbeitungsmöglichkeiten und Verknüpfungsmöglichkeiten“ gebe es „kein ‚belangloses‘ Datum mehr“.[117] Andererseits hat es aber – mit Blick auf die für Big Data charakteristischen Dekontextualisierungs- und Rekontextualisierungsmöglichkeiten zweifelsohne überschießend – formuliert, mit dem Recht auf informationelle Selbstbestimmung sei eine Gesellschaftsordnung und eine diese ermöglichende Rechtsordnung unvereinbar, „in der Bürger nicht mehr wissen können, wer was wann und bei welcher Gelegenheit über sie weiß“.[118] Trotz dieser missverständlichen Formulierung, die das Recht auf informationelle Selbstbestimmung eigentumsähnlich erscheinen lässt, hat dieses Recht einen instrumentellen Charakter. Es steht im Dienst freier Entfaltung der Persönlichkeit im Sinne der Doppelgewährleistung des Art. 2 Abs. 1 GG. Das Recht auf informationelle Selbstbestimmung zielt nicht nur auf den Schutz des allgemeinen Persönlichkeitsrechts[119], sondern auch auf die

[115] BVerfGE 65, 1.
[116] Siehe hierzu im Kontext des Gesundheitsdatenschutzrechts Kühling/Seidel 2015, 155 ff., 162 ff.
[117] BVerfGE 65, 1 (45).
[118] BVerfGE 65, 1 (43).
[119] So die traditionelle Rekonstruktion.

Gewährleistung der allgemeinen Handlungsfreiheit und der anderen freiheitlichen Teilgarantieelemente des Art. 2 Abs. 1 GG.[120] Auch das Bundesverfassungsgericht hat in jüngeren Entscheidungen mehrfach betont, das Recht auf informationelle Selbstbestimmung flankiere und erweitere den grundrechtlichen Schutz von Privatheit und Verhaltensfreiheit.[121] Dabei ist gegenüber verbreiteten Fehldeutungen zu betonen, dass das Recht auf informationelle Selbstbestimmung auch die Befugnis umfasst, selbst zu bestimmen, mit welchen Inhalten und in welchen Beziehungen jemand in den Prozess interaktiver Persönlichkeitsentfaltung mit seiner Umwelt eintritt. Dies gilt sowohl für das deutsche als auch das europäische Verfassungsrecht.[122] Diese Entfaltungsfreiheiten können ihrerseits kollidieren mit wichtigen Gemeinwohlbelangen wie der Förderung des wissenschaftlichen Fortschritts oder der Gewährleistung einer effektiven Gesundheitsversorgung. Aber auch die Grundrechtspositionen anderer Privatrechtssubjekte, die ihnen zugängliche Informationen aufgreifen und verarbeiten wollen, sind als potenziell gegenläufige Abwägungsgesichtspunkte zu berücksichtigen. Angesichts dieser möglichen Konflikte konstituieren die verfassungsrechtlichen Vorgaben lediglich einen Ordnungsrahmen, dessen nähere Ausgestaltung dem Gesetzgeber zugewiesen ist.

Soweit es um hoheitliche Beschränkungen des Rechts auf informationelle Selbstbestimmung (einschließlich der Privatheit) geht, hat der Staat dabei die abwehrrechtliche Funktion des Grundrechts unter Wahrung des Verhältnismäßigkeitsgrundsatzes zu beachten. Das verfassungsrechtlich geleitete Datenschutzrecht kennt insoweit nicht nur einen klassischen Eingriffsvorbehalt, sondern durchaus auch „modernere" Elemente wie institutionelle Sicherungen (etwa Datenschutzbeauftragte) und Verfahrens- und Transparenzanforderungen.

Big-Data-Problematiken betreffen indes vielfach horizontale Rechtsbeziehungen zwischen privaten Grundrechtssubjekten.[123] Datenschutz ist kein Selbstzweck, sondern dient der informationellen Selbstbestimmung der beteiligten Personen. Dementsprechend besteht auch eine grundsätzliche staatliche Schutzpflicht, Vorkehrungen gegenüber privaten Datenschutzgefährdungen zu ergreifen.[124] Hierauf reagiert das Datenschutzrecht etwa mit der Zurückdrängung der Einwilligungswirkung bei fehlender Transparenz und Kontrollierbarkeit des Datenverarbeitungsprozesses.[125] Das Bundesverfassungsgericht hat Kooperationspflichten gefordert, die sicherstellen, dass Nutzer und Datenverarbeiter im Dialog ermitteln, welche Daten erforderlich

[120] Dazu etwa (die Richterin des Bundesverfassungsgerichts) Britz 2010, 573 ff.
[121] Beispielsweise BVerfGE 120, 378 (397).
[122] So hinsichtlich Art. 8 Abs. 1 GRC ausdrücklich Klement 2017, 169; vgl. auch BVerfGE 120, 180 (197): Das allgemeine Persönlichkeitsrecht habe auch „die Aufrechterhaltung der Grundbedingungen sozialer Beziehungen zwischen dem Grundrechtsträger und seiner Umwelt zum Ziel".
[123] Vgl. auch Buchner 2006, 46 ff.
[124] Siehe etwa BVerfGE 84, 192 (194 f.); ferner beispielsweise Rudolf 2011, Rn. 26 ff.; zum leistungsrechtlichen Gehalt von Art. 8 Abs. 1 GRC, Art. 16 Abs. 1 AEUV siehe auch Augsberg 2015, Rn. 8.
[125] Siehe etwa Simitis 2014a, Rn. 7.

sind.[126] Allerdings sind auch die vielfältigen und sich verändernden Formen informationeller Selbstgefährdung[127] und subtile Formen der Machtausübung (siehe Abschnitt 4.3) verstärkt in den Blick zu nehmen. Sie verlangen vom freiheitsaustarierenden Gesetzgeber Regulierungsansätze, die der Komplexität und Dynamik der Regelungsmaterie gerecht werden. Erneut zeigt sich hier, dass das Verfassungsrecht keine abschließend-unmittelbaren Antworten geben kann.

3.2 Einfachrechtliche Vorgaben

Zum gegenwärtigen Zeitpunkt existieren – der Aktualität des Phänomens geschuldet – nur allgemeine rechtliche Vorgaben für Big-Data-basierte Gesundheitsdienste. Spezielle, den grundlegenden Wandel zu Big Data explizit aufnehmende und vorstrukturierende Normen fehlen weitestgehend. Die rechtliche Analyse hat sich deshalb zunächst mit der Frage zu beschäftigen, ob und wie die neuen technischen Möglichkeiten zu bestehenden, ursprünglich für andere Verwendungskontexte geschaffenen Regelungen passen. Zu klären ist, ob und inwieweit das geltende Recht auf die neuen Big-Data-Konstellationen angemessen vorbereitet ist bzw. wo Spannungen bestehen. Dazu werden das allgemeine Datenschutzrecht und das Datenschutzrecht im Gesundheitsbereich näher untersucht und anschließend das Medizinprodukterecht und das Krankenversicherungsrecht in den Blick genommen.

3.2.1 Big Data als Herausforderung für das geltende Datenschutzrecht

Das Datenschutzrecht ist, wie die nachfolgenden Ausführungen zeigen werden, auch nach seinen jüngsten, durch die europäische Datenschutz-Grundverordnung (DSGVO)[128] veranlassten Veränderungen auf das Phänomen Big Data unzureichend eingestellt. Unzweifelhaft sind zwar aufgrund der aktuellen regulatorischen Anstrengungen signifikante Verbesserungen im Bereich des Datenschutzes erreicht worden. Die DSGVO und die derzeit noch im europäischen Gesetzgebungsverfahren befindliche sogenannte E-Privacy-Verordnung[129] bedeuten insbesondere mit Blick auf die Etablierung grenzüberschreitender Standards sowie die stärkere Einbeziehung des Konzepts von *privacy by design* (Art. 25 DSGVO)[130] einen klaren Fortschritt. Ungeachtet der diesbezüglich bereits geäußerten Kritik gilt das im Grundsatz auch für das auf die

[126] Siehe Urteil des Bundesverfassungsgerichts in NJW 2013, 3086 (3088 f.).
[127] Siehe dazu Hermstrüwer 2016.
[128] Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (ABl. EU 2016 Nr. L 119/1).
[129] Vgl. den Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über die Achtung des Privatlebens und den Schutz personenbezogener Daten in der elektronischen Kommunikation und zur Aufhebung der Richtlinie 2002/58/EG vom 10. Januar 2017 (COM(2017) 10 final). Nachdem zuletzt das Europäische Parlament mehrheitlich für den (leicht abgeänderten) Entwurf gestimmt hat, finden nun im Rahmen der sogenannten Trilog-Prozesses Verhandlungen zwischen dem Parlament, dem Ministerrat und der Kommission statt. Zum Ganzen etwa Engeler/Felber 2017; kritisch siehe Hanfeld 2017.
[130] Näher hierzu etwa Baumgartner/Gausling 2017; ferner Gossen/Schramm 2017 und Schmitz/Dall'Armi 2017.

DSGVO bezogene deutsche Umsetzungsgesetz[131] und das darin enthaltene neue Bundesdatenschutzgesetz (BDSG n. F.). Allerdings genügt es nicht, eine adäquate, rechtliche Interoperabilität sicherstellende Umsetzung der unionalen Vorgaben in den Mitgliedstaaten einzufordern.[132] Erst recht kann nicht davon ausgegangen werden, bei „vollständiger Umsetzung und Durchsetzung der hierzulande [geltenden] datenschutzrechtlichen Vorgaben" sei auch gegenüber modernen Big-Data-basierten Anwendungen (hier: Gesundheits-Apps) ein „entsprechender Datenschutz gewährleistet".[133] Vielmehr sind grundlegende Annahmen, zentrale Prinzipien und Zielvorgaben des überkommenen Datenschutzrechts mit den Besonderheiten von Big-Data-Anwendungen kaum in Einklang zu bringen. Die traditionellen datenschutzrechtlichen Grundsätze des Personenbezugs, der Zweckbindung und Erforderlichkeit der Datenerhebung, der Datensparsamkeit, der Einwilligung und Transparenz stehen in ihrer gegenwärtigen Ausgestaltung der spezifischen Eigenlogik von Big Data entgegen. Big Data kann deshalb nicht reibungslos an die bestehenden rechtlichen Vorgaben und die sie konkretisierenden Vereinbarungen angepasst werden. Das hat erhebliche Konsequenzen. Will man weder den Einsatz von Big Data grundsätzlich untersagen noch relevante Einbußen am Schutzniveau hinnehmen, müssen alternative Gestaltungsoptionen und Regelungsmechanismen entwickelt werden. Dies wird bei genauerer Analyse der zentralen Elemente des allgemeinen Datenschutzrechts deutlich.

Personenbezug

Wie erwähnt, hat das Bundesverfassungsgericht schon im Volkszählungsurteil betont, es gebe „unter den Bedingungen der automatischen Datenverarbeitung kein ‚belangloses' Datum mehr."[134] Allerdings knüpft das geltende Datenschutzrecht an den Personenbezug von Daten an und legt besonderen Wert auf die damit einhergehenden spezifischen Zweckbindungen. Das gilt sowohl für die DSGVO[135] wie das (alte und neue) BDSG als auch für den zivilrechtlichen Schutz personenbezogener Daten durch das allgemeine Persönlichkeitsrecht.[136]

Für Big Data ist demgegenüber entscheidend, dass bei der Erfassung der Daten die künftigen Anwendungen nicht vorhersehbar sind und auch der Personenbezug bzw. der Bezug zu ihrer Gesundheit unter Umständen erst nachträglich hergestellt wird. Die Datenmengen entstammen gerade nicht einem bestimmten Verwendungszusammenhang und sind nicht auf diesen

---

[131] Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 vom 30. Juni 2017 (BGBl. I, 2097).
[132] So die Europäische Kommission (vgl. Habl et al. 2016, 54).
[133] Albrecht 2016, 29: „Die Schwächen dürften in der Umsetzung durch die Anbieter und der mangelnden Transparenz bei der Einholung der Einwilligung und der Aufklärung sowie der Sensibilität der Anwender im Zusammenhang mit datenschutzrechtlichen Fragen gegeben sein." Als problematisch bezeichnet werden vor allem die unzureichende Regulierung internationaler App-Angebote und die fehlende Umsetzung existierender Vorgaben.
[134] BVerfGE 65, 1 (45).
[135] Vgl. etwa Hofmann/Johannes 2017 und Krügel 2017.
[136] Vgl. zum Personenbezug etwa Werkmeister/Brandt 2016, 234 f.; siehe ferner das Urteil des Europäischen Gerichtshofes in NJW 2016, 3579 (3580 ff.).

beschränkt. Im Gegenteil „wird nicht ein bestimmtes Untersuchungsdesign mit einer fest vorgegebenen Fragestellung entworfen, sondern es werden Daten ohne einen bestimmten Zweck, jedenfalls ohne feste Verknüpfung mit einem Zweck, gesammelt, die über unterschiedliche, auch jeweils selbst auf Wandel, also Lernen angelegte ‚Algorithmen' nach produktiven Verknüpfungsmustern abgesucht werden, die selbst allerdings Aussagen insbesondere über Krankheitsverläufe oder Therapien ermöglichen sollen".[137] Diese besondere „Offenheit der Verknüpfungsmöglichkeiten" stellt ersichtlich schon den datenschutzrechtlichen Ausgangspunkt des Schutzes personenbezogener Daten infrage.[138]

Zweckbindung

Big-Data-Anwendungen stehen darüber hinaus in einem Spannungsverhältnis zu dem grundlegenden, in der DSGVO noch einmal betonten[139] datenschutzrechtlichen Gebot der Zweckbindung personenbezogener Daten. Diese Daten dürfen demnach, um Transparenz herzustellen und eine Kontrolle der Rechtmäßigkeit der Datenverarbeitung zu ermöglichen[140], nur für diejenigen Zwecke gespeichert, verändert oder genutzt werden, für die sie erhoben wurden. Eng verknüpft hiermit ist das Prinzip, dass die Datenverwendung zur Erfüllung der (in einer Erlaubnisvorschrift oder im Rahmen der Einwilligung festgelegten) Zwecke erforderlich sein muss (vgl. § 28 Abs. 1 S. 1 Nr. 2 BDSG a. F.). Auch dies widerspricht indes der durch Ergebnisoffenheit und permanente Reteleologisierungsprozesse gekennzeichneten Binnenlogik von Big Data. Im Rahmen von Big Data werden oft Daten, die zu anderen Zwecken gespeichert wurden, für neue Zwecke ausgewertet. Entsprechend ist es für Big-Data-Anwendungen auch geboten, Daten für noch unbestimmte Zwecke zu erheben und zu speichern.[141] Im Rahmen der Big-Data-typischen statistikbasierten und wahrscheinlichkeitsorientierten Erkenntnismethoden ist deshalb eine strikte Zweckbindung kaum möglich. Damit läuft der datenschutzrechtliche Grundsatz der Erforderlichkeit leer: „Wenn Daten für beliebige Zwecke ausgewertet werden sollen, sind sie immer erforderlich."[142]

Datensparsamkeit

In augenfälligem Widerspruch zu Big Data steht ferner der in Erwägungsgrund 156, Art. 5 Abs. 1 lit. c, Art. 25 Abs. 1 DSGVO sowie in § 71 Abs. 1 BDSG n. F. normierte Grundsatz der Datensparsamkeit/Datenvermeidung bzw. – in der Terminologie der DSGVO – Datenminimierung. Nach diesem Grundsatz sollen so wenig personenbezogene Daten wie möglich erhoben, verarbeitet oder genutzt werden. Damit korrespondiert die prinzipielle Pflicht, Daten zu löschen,

[137] Ladeur 2016, 360.
[138] Ebd., 363 mit Verweis auf Tene/Polonetsky 2012, 63; Schwartz/Solove 2011, 1814. Vgl. zu dem Konflikt auch Schneider 2017.
[139] Vgl. Schantz 2016, 1843 f.
[140] Vgl. Simitis 2014b, Rn. 111.
[141] Vgl. Roßnagel/Nebel 2015, 458.
[142] Roßnagel/Nebel 2015, 458.

sobald sie nicht mehr benötigt werden (vgl. etwa § 35 BDSG n. F.). Wie in Kapitel 2 ausführlich dargelegt, entfaltet sich das Potenzial von Big Data jedoch dann am wirkungsvollsten, wenn in möglichst unbegrenztem Ausmaß Daten gesammelt und miteinander verknüpft werden können. Löschungen führten dazu, dass bestimmte Korrelationsmuster nicht gefunden werden. Wenn an der pauschalen Verpflichtung auf Datensparsamkeit festgehalten würde, führte dies deshalb zu einem weitgehenden Ausschluss der Möglichkeiten von Big Data. Weil aber mit der Menge an gespeicherten Daten zugleich das Gefährdungspotenzial für das Recht auf informationelle Selbstbestimmung wächst, bedarf es wirksamer alternativer Schutzmechanismen.

Einwilligungserfordernis

Verdeutlichen lassen sich die Inkompatibilitäten zwischen Big Data und dem traditionellen Datenschutzrecht ferner an dem das gegenwärtige Datenschutzrecht dominierenden, etwa in Art. 6 Abs. 1, Art. 7 DSGVO normierten Erfordernis der Einwilligung.[143] Vorbehaltlich einer gesonderten (durch einen bestimmten Zweck eingegrenzten) gesetzlichen Erlaubnisvorschrift ist demnach eine Datenverwendung nur erlaubt, wenn der Betroffene eingewilligt hat. Die Einwilligung ist damit für den Betroffenen die wichtigste Möglichkeit, sein Recht auf informationelle Selbstbestimmung wahrzunehmen; sie soll nach dem Willen des Gesetzgebers die Verwendung personenbezogener Daten einschränken und kontrollieren.[144] Hieran hält prinzipiell auch die DSGVO fest. Sie sieht zwar neben einer Lockerung mit Blick auf das im deutschen Recht relativ strikt gehandhabte Schriftformerfordernis[145] auch die Möglichkeit vor, eine Einwilligung für „einen oder mehrere bestimmte Zwecke“ zu erteilen (Art. 6 Abs. 1 lit. a DSGVO). In der Sache ist damit indes keine Abkehr von dem bisherigen Konzept verbunden. Denn gemäß der Begriffsbestimmung des Art. 4 Nr. 11 DSGVO (entsprechend auch § 46 Nr. 17 BDSG n. F.) ist eine Einwilligung „jede freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung, mit der die betroffene Person zu verstehen gibt, dass sie mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist“.[146]

---

[143] Vgl. allgemein skeptisch zur Leistungsfähigkeit individueller Einwilligungen angesichts von „on the one hand, well-documented cognitive biases, and on the other hand the increasing complexity of the information ecosystem“, Tene/Polonetsky 2012, 67.
[144] Siehe zur Einwilligung gemäß § 4a BDSG a. F. auch Simitis 2014a, Rn. 4.
[145] Vgl. Krohm 2016 sowie Thüsing/Schmidt/Forst 2017.
[146] Siehe auch Erwägungsgrund 32 der DSGVO: „Die Einwilligung sollte durch eine eindeutige bestätigende Handlung erfolgen, mit der freiwillig, für den konkreten Fall, in informierter Weise und unmissverständlich bekundet wird, dass die betroffene Person mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist, etwa in Form einer schriftlichen Erklärung, die auch elektronisch erfolgen kann, oder einer mündlichen Erklärung. […] Die Einwilligung sollte sich auf alle zu demselben Zweck oder denselben Zwecken vorgenommenen Verarbeitungsvorgänge beziehen. Wenn die Verarbeitung mehreren Zwecken dient, sollte für alle diese Verarbeitungszwecke eine Einwilligung gegeben werden. Wird die betroffene Person auf elektronischem Weg zur Einwilligung aufgefordert, so muss die Aufforderung in klarer und knapper Form und ohne unnötige Unterbrechung des Dienstes, für den die Einwilligung gegeben wird, erfolgen.“ Siehe außerdem Buchner/Kühling 2017, Rn. 61. ff.

Vor diesem Hintergrund ist für das unionale wie nationale Recht weiterhin davon auszugehen, dass die Einwilligung nur wirksam ist, wenn der Betroffene die Bedeutung und Tragweite der beabsichtigten Datenverwendung überblicken kann. Aus diesem Grund ist er über die Verarbeitungsziele, die Identität der Verarbeitungsverantwortlichen und den Umfang der Verarbeitung zu informieren (vgl. Erwägungsgrund 42 der DSGVO). Im Falle einer Übermittlung muss sich aus der Einwilligung ergeben, wem konkret die Daten zugänglich gemacht werden dürfen.[147] Als Grundlage einer angemessenen Entscheidungsfindung werden darüber hinaus Informationen über die betroffenen Daten, die beabsichtigte Speicherdauer und die Folgen einer verweigerten Einwilligung gefordert.[148] Gerade mit Blick auf besonders problembeladene Anwendungskonstellationen enthält Art. 8 DSGVO nunmehr eine bereichsspezifische Weiterentwicklung des bisherigen Einwilligungsmodells. Bislang setzt eine wirksame Einwilligung nach unionalem ebenso wie nach nationalem Datenschutzrecht nicht die Volljährigkeit (Geschäftsfähigkeit), sondern lediglich eine anlassbezogen zu ermittelnde Einsichtsfähigkeit voraus.[149] Die Neuregelung des Art. 8 DSGVO reagiert nun auf die Tatsache, dass Minderjährige sich der Risiken bei der Verarbeitung personenbezogener Daten weniger bewusst sind als Erwachsene, mit einer pauschalen Schutzerhöhung. Für den besonders praxisrelevanten Bereich der „Dienste der Informationsgesellschaft“[150] etabliert die Norm ein Mindestalter von 16 Jahren. Wird diese Grenze nicht erreicht, ist stets der Träger der elterlichen Verantwortung einzubeziehen.[151] Art. 8 Abs. 2 DSGVO verpflichtet zudem die Datenverwender, technisch-organisatorische Vorkehrungen zu treffen, um sicherzustellen, „dass die Einwilligung durch den Träger der elterlichen Verantwortung für das Kind oder mit dessen Zustimmung erteilt wurde“. Ein bestimmtes (Alters-)Verifikationssystem wird indes nicht vorgegeben; zudem beschränkt sich die Verpflichtung im Sinne des Verhältnismäßigkeitsgrundsatzes auf „angemessene Anstrengungen“.[152]

Die Wirksamkeit und die Funktionalität dieser Anforderungen sind bislang schon umstritten. Nicht nur bei Minderjährigen, sondern auch bei Erwachsenen ist häufig mehr als zweifelhaft,

---

[147] Siehe zu § 4a BDSG a. F. auch Simitis 2014a, Rn. 80 und für Art. 7 DSGVO Buchner/Kühling 2017, Rn. 20 ff.; vgl. ferner Schaar 2017.
[148] Vgl. etwa Sivridis/Seidel/Kühling 2015, Rn. 319.
[149] Vgl. Ernst 2017, 111 m. w. N. sowie Kampert 2017, Rn. 2 m. w. N.
[150] Erfasst sind damit nach Art. 4 Nr. 25 DSGVO Dienstleistungen im Sinne von Art. 1 Nr. 1 lit. b der Richtlinie 2015/1535/EU, also (nur) „jede in der Regel gegen Entgelt elektronisch im Fernabsatz und auf individuellen Abruf eines Empfängers erbrachte Dienstleistung“. Wichtig ist dabei der Hinweis, dass entgeltlich in diesem Sinne auch werbefinanzierte Dienste sind, die für den Nutzer selbst daher kostenlos sind (vgl. Kampert 2017, Rn. 268 m. w. N.). Die im Interesse des Kindeswohls einleuchtende explizite Ausnahmeregelung für Präventions- oder Beratungsdienste, die unmittelbar einem Kind angeboten werden (vgl. Erwägungsgrund 38 der DSGVO), dürfte hingegen primär deklaratorischen Charakter besitzen, weil diese Dienste regelhaft kostenfrei sind.
[151] Art. 8 Abs. 1 S. 3 DSGVO gestattet es den Mitgliedstaaten, von dieser Vorgabe bis zu einer absoluten Untergrenze von 13 Jahren abzuweichen. Der deutsche Umsetzungsgesetzgeber hat indes von dieser Öffnungsklausel keinen Gebrauch gemacht.
[152] Vgl. Kampert 2017, Rn. 13 f.

dass sie insbesondere die Verwendungszwecke und die damit verbundenen Implikationen tatsächlich verstehen. Aus einer realitätsnahen Perspektive wird die Einwilligung deshalb nicht selten als eine Formalie bezeichnet. Gerade im Gesundheitswesen wird eine autonome Entscheidung des Patienten angesichts der Informationsasymmetrie zwischen Patient und Gesundheitsdienstleistern vielfach infrage gestellt. Erst recht gilt dies dort, wo die Preisgabe von Daten zur Voraussetzung für die Inanspruchnahme bestimmter medizinischer Leistungen wird und auf diese Weise das ohnehin schon bestehende Abhängigkeitsverhältnis zwischen Heileinrichtung und Betroffenem vertieft wird.[153]

Bei Big-Data-Anwendungen verstärkt sich diese allgemeine Problematik noch einmal erheblich[154]: Zunächst sind künftige Verwendungsarten zum Zeitpunkt der Datenerhebung oftmals nicht einmal als vage Möglichkeit bekannt, sodass der Betroffene über die Verwendung nicht unterrichtet werden kann.[155] Big Data ermöglicht nämlich nicht nur das Echtzeit-Monitoring, sondern auch die Simulation und Vorhersage zukünftiger Szenarien.[156] Somit kann der Betroffene auch nicht absehen, welche Ergebnisse aus seinen Daten generiert werden könnten. Auch können Big-Data-Analysen zeitlich kaum eingegrenzt werden. Eine Blankoeinwilligung für beliebige Zwecke reicht nach geltendem Datenschutzrecht jedoch nicht aus, da sie einem abstrakten Bekenntnis zur Selbstbestimmung gleichkäme.[157]

Selbst wenn man eine teilweise Abschätzbarkeit der Datenverwendung seitens der verarbeitenden Stelle unterstellt, ist zweifelhaft, ob dem Betroffenen derart komplexe und umfangreiche Informationen überhaupt in verständlicher Weise vermittelt werden können. Mittlerweile sind viele Datenschutzeinwilligungen so detailliert und schwer verständlich, dass der Verbraucher mit Informationen regelrecht überflutet wird.[158] Derartige Einwilligungen bieten in der Sache ebenfalls keine hinreichende Legitimationsgrundlage. Im Rahmen von Big-Data-Anwendungen verkommt die Einwilligung damit zu einer bloßen Formalie und ist als Legitimationsmittel kaum mehr geeignet.

Insgesamt wird deutlich, dass das geltende Datenschutzrecht auf die Einwilligung fokussiert ist und damit punktuell ansetzt. Abgesehen vom Recht auf Widerruf (Art. 7 Abs. 3 DSGVO) und den (sogleich zu erörternden) Rechten auf Auskunft, Berichtigung, Sperrung und Löschung bietet es wenig Möglichkeiten, prozesshaft auf das weitere Schicksal der Daten Einfluss zu nehmen. Im Grundsatz bedarf eine Verwendung zu anderen, von der bisherigen Einwilligung nicht gedeckten Zwecken stets einer neuen Einwilligung, und diese muss zeitlich der Verarbeitung

[153] Vgl. am Beispiel des Transplantationsregisters Augsberg 2016.
[154] Vgl. kritisch deshalb etwa Becker 2017, 173; Ulbricht/Weber 2017 und Mostert et al. 2016.
[155] Vgl. Mayer-Schönberger/Cukier 2013, 171 ff.
[156] Vgl. Langkafel 2015, 27.
[157] Vgl. Ernst 2017, 113, und Simitis 2014a, Rn. 3, 77.
[158] Vgl. Katko/Babaei-Beigi 2014, 362.

vorausgehen.[159] Nachträgliche Erklärungen können die Unzulässigkeit der Verarbeitung nicht beseitigen, da eine Genehmigung (§ 184 BGB) keine vorhergehende Einwilligung (§ 183 BGB) ist und das BDSG auch keine Heilung kennt.[160] Sind die Daten einmal mit Einwilligung erhoben, können sie von dem Betroffenen nicht mehr weiterverfolgt werden. Zwar bieten die Rechte auf Auskunft, Berichtigung, Sperrung und Löschung als Kontrollrechte des Betroffenen (etwa Art. 13 ff. DSGVO, § 32 ff. BDSG n. F)[161] zur effektiven Wahrnehmung seines Rechts auf informationelle Selbstbestimmung diesem die Möglichkeit zu überprüfen, ob seine Daten auch entsprechend der Zweckbestimmung verwendet werden. Eine missbräuchliche Verwendung kann dann mittels der Ansprüche auf Unterlassung einer unzulässigen Verwendung der Daten und auf Ersatz des hierdurch entstehenden Schadens abgewehrt bzw. sanktioniert werden. Diese Instrumente sind jedoch darauf ausgerichtet, dass die Fragestellung oder das Begehren des Betroffenen eingrenzbar bleibt (zum Beispiel ein bestimmter Datensatz, ein bestimmter Zeitraum), und damit eher als punktuelle Interventionsmöglichkeiten konzipiert. Die Dynamik von Big Data passt nicht in dieses Regelungskonzept. Gerade wenn man die Zustimmung der Betroffenen für ein unabdingbar gebotenes basales Erfordernis des Datenschutzes erachtet, ist deshalb nach Wegen zu suchen, wie dies auch unter Big-Data-Bedingungen funktional sinnvoll möglich ist.

Anonymisierung und Pseudonymisierung

Vor dem Hintergrund der (ihrerseits computer- und datengestützten, unter Big-Data-Bedingungen intensivierten) Möglichkeiten der Reidentifizierung bestehen zudem grundlegende Zweifel an der Effektivität des Anonymisierungs- bzw. Pseudonymisierungsgebots.[162] Zwar widerspricht das damit eingeforderte Auflösen des Personenbezugs, wann immer dies möglich ist, nicht prinzipiell dem Konzept von Big Data. Im Gegenteil ist ein entsprechender, nach herkömmlichem Verständnis für die Bestimmung, ob das Datenschutzrecht auf einen bestimmten Sachverhalt Anwendung findet oder nicht, entscheidender Personenbezug für Big-Data-Nutzungen regelhaft unwesentlich.

Gleichwohl ist insbesondere bei medizinischen Daten grundlegend zu fragen, ob sich ein Personenbezug im Zeitalter von Big Data überhaupt vollständig vermeiden bzw. beseitigen lässt. Schon in der Vergangenheit wurde argumentiert, dass eine Anonymisierung genomischer Daten aufgrund des einzigartigen „genetischen Fingerabdrucks“ eines jeden Menschen faktisch

[159] Vgl. Simitis 2014a, Rn. 27.
[160] Vgl. ebd., Rn. 29.
[161] Vgl. schon BVerfGE 65, 1 (46): „Als weitere verfahrensrechtliche Schutzvorkehrungen sind Aufklärungspflichten, Auskunftspflichten und Löschungspflichten wesentlich.“
[162] Vgl. grundlegend Ohm 2010; bereichsbezogen etwa Mostert et al. 2016.

nicht mehr zu leisten ist.[163] Bereits jetzt werden zur Ursachenforschung komplexer Erkrankungen biologische, klinische und soziodemografische Daten sowie Daten zum Lebensstil (zum Beispiel Grad der sportlichen Betätigung, Ernährungsverhalten etc.) erhoben.[164] Im Wege einer systemorientierten Medizin sollen in Zukunft auch im Behandlungskontext umfassend molekulare Daten (sogenannte Omik-Daten, siehe Abschnitt 2.4.2), klinische Daten, soziodemografische Daten, Daten zum Lebensstil, psychosoziale Daten und Daten über Umwelteinflüsse eines jeden Patienten zwecks Analyse und Vorhersage seines Gesundheitszustands zusammengeführt werden. Damit werden die Datensätze der Patienten zunehmend individueller. Um eine effektive Anonymisierung zu gewährleisten, müssten die einzelnen Parameter vor der Überführung in einen anderen Kontext unabhängig voneinander gespeichert werden. Dies jedoch liefe dem Konzept von Big Data zuwider, dass neue Aussagen hinsichtlich der Entdeckung unbekannter Zusammenhänge gerade aufgrund der umfassenden Verknüpfung unterschiedlichster individueller Daten getroffen werden sollen. Selbst im Falle teilweise voneinander getrennter, jeweils anonymisierter Datensätze bestünde bei einer Überführung in andere Kontexte die Gefahr, dass diese anonymisierten, individuellen Patienten- bzw. Probandendaten mithilfe von Big-Data-Analysen durch Auswertung und Zusammenführung mit weiteren im Internet zugänglichen oder in sämtlichen sonstigen Datenbeständen enthaltenen Daten wieder einzelnen Personen zugeordnet werden könnten.[165] Je individueller der Datensatz ist, umso leichter kann er durch in anderen Beständen enthaltenes Zusatzwissen konkretisiert und schließlich rückführbar gemacht werden. Dabei kommt es nicht darauf an, welche Datenquellen „vernünftigerweise“ vorstellbar sind und ob der Zugriff auf diese Daten rechtlich zulässig oder unzulässig ist.[166] Besonders riskant sind in diesem Zusammenhang die unzähligen Daten, die Verbraucher teils bewusst, teils unbewusst mit ihrem Verhalten im Internet erzeugen.[167] Insgesamt wird so der Reidentifizierungsaufwand durch die Verbesserung der Analysetools und der wachsenden verfügbaren Datenbestände von Jahr zu Jahr kleiner.[168] Die Frage, inwieweit und ab welchem Grad die Gefahr einer Reidentifizierung für sich genommen anonymisierter Daten für die Annahme eines Personenbezugs der Daten ausreichend ist[169], verschärft die Problematik um den ohnehin schon umstrittenen Begriff des Personenbezugs im Datenschutzrecht.[170]

---

[163] Vgl. etwa Berdin 2017, 194; Vossenkuhl 2013, 97.
[164] Kollek 2012, 26.
[165] Vgl. Weichert 2014a, 170.
[166] Vgl. ebd.
[167] Vgl. Chatziastros/Drepper/Semler 2014.
[168] Vgl. Weichert 2014b, 836.
[169] Das Datenschutzrecht stellt mit Blick auf die Anonymisierung darauf ab, ob die Daten „nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft“ einer bestimmten oder bestimmbaren Person zugeordnet werden können (vgl. § 3 Abs. 6 BDSG a. F.).
[170] So wird es für die Annahme eines Personenbezugs von den Datenschutzbehörden als ausreichend erachtet, dass irgendein Dritter die Zuordnung von (pseudonymisierten) Daten wiederherstellen kann. Damit wären Daten für jeden Empfänger personenbezogen. Die Gegenansicht stellt hingegen darauf ab, ob die speichernde

Auskunft, Berichtigung, Löschung, Sperrung

Die bereits angesprochenen Rechte auf Auskunft, Berichtigung, Löschung und Sperrung (Art. 13 ff. DSGVO, §§ 32 ff. BDSG n. F.) bestehen, sofern und solange die Daten einen Personenbezug aufweisen.[171] Damit dienen sie dem Grundsatz der Transparenz. Sie bieten aber häufig keinen effektiven Schutz, weil sie praktisch nicht oder nur sehr schwer geltend gemacht und durchgesetzt werden können und damit faktisch leerlaufen.[172] Schon die Reichweite und Detailtiefe der Auskunftsansprüche sind angesichts der vielen Einzeldaten, die der Betroffene in ihrer Gesamtheit nicht versteht und aus denen auch stetig immer neue Erkenntnisse generiert werden können, problematisch. Ähnlich wie bei der Einwilligung geht auch das Recht auf Auskunft mit einem (mitunter problematischen) erheblichen zusätzlichen Informationsbedarf und gegebenenfalls sogar ärztlichem Beratungsbedarf[173] einher. Ferner wird der Betroffene kaum alle potenziellen Anspruchsgegner kennen, was aber erforderlich wäre, um sich umfassende Kenntnis davon zu verschaffen, welche erzielbare Gesamtinformation als zerstreutes Mosaik über ihn in Umlauf ist. Aus genau diesem Grunde bestehen auch berechtigte Zweifel, dass das in Art. 20 DSGVO enthaltene Recht auf Datenportabilität bzw. Datenübertragbarkeit die damit verbundenen Hoffnungen und Erwartungen erfüllen wird. Die Vorschrift gibt der betroffenen Person unter anderem das Recht, den auf sie bezogenen Datensatz von dem Verantwortlichen in einem strukturierten, gängigen und maschinenlesbaren Format zu erhalten. Das soll eine bessere Kontrolle über die eigenen Daten ermöglichen und zugleich sicherstellen, dass Lock-in-Effekte vermieden und Anbieter unkompliziert gewechselt werden können.[174] Gerade die Feststellung, wer Verantwortlicher in diesem Sinne ist, dürfte aber unter Big-Data-Bedingungen größte Schwierigkeiten bereiten.

Von den Auskunftsrechten umfasst ist grundsätzlich auch die Nachvollziehbarkeit des Datenverarbeitungsprozesses. Insoweit sind die Vorgaben des Art. 22 DSGVO von besonderer Relevanz. Weil der Einzelne nicht zum bloßen Objekt rein maschineller Entscheidungen werden soll, sind demnach voll automatisierte Verfahren grundsätzlich verboten. Die vorgesehenen Ausnahmen (für Vertragsverhältnisse, aufgrund spezieller Regelungen und bei Einwilligung) setzen zudem angemessene Schutzmaßnahmen – insbesondere ein Diskriminierungsverbot – voraus und sind mit bestimmten Informationspflichten gekoppelt.[175] Der Betroffene muss er-

Stelle mit eigenen Mitteln einen Personenbezug herstellen kann. Somit sind nach dieser Position die Daten nur für diejenigen, die den Personenbezug herstellen können, personenbezogen Daten.

[171] Vgl. Weichert 2013, 252; siehe auch Paal/Hennemann 2017, 1700 f.

[172] Vgl. auch Keppeler/Berning 2017.

[173] Vgl. Weichert 2014b, 837.

[174] Vgl. Keßler 2017, 591 sowie Paal/Hennemann 2017, 1701.

[175] Vgl. dazu etwa Buchner 2017, Rn. 11 ff. und Martini/Nink 2017.

kennen können, welche Daten und welche Analysemethoden in den Verarbeitungsprozess eingehen und wie das Analyseergebnis zustande kommt.[176] Dies beträfe damit insbesondere die Analyse-Algorithmen. Die Algorithmen selbst unterfallen jedoch aufgrund der in ihnen liegenden geistigen Leistungen und zum Schutz der Betriebs- und Geschäftsgeheimnisse nicht dem Auskunftsanspruch.[177] Im Übrigen wäre, selbst wenn im Gesundheitswesen Algorithmen zwecks Vergleichbarkeit offengelegt würden, damit für den Auskunftsanspruch des Einzelnen mangels Verständlichkeit der komplexen Rechenformeln, insbesondere unter den Bedingungen sich selbst fortschreibender Algorithmen im Bereich Deep Learning (siehe Abschnitt 2.3.2), kaum etwas gewonnen. Da Algorithmen auch zu Fehlschlüssen führen können, ist diese Intransparenz datenschutzrechtlich hochproblematisch, zumal es auch keine staatliche Algorithmuskontrolle gibt.[178] Damit läuft auch das bislang bestehende Recht auf Berichtigung und Löschung leer, da der Betroffene diese Rechte ohne eine umfassende Auskunft nicht wahrnehmen kann.

3.2.2 Gesundheitsdatenschutzrecht

Die vorstehende, auf das allgemeine Datenschutzrecht bezogene Defizitanalyse kann mit gewissen Einschränkungen auf das besondere Gesundheitsdatenschutzrecht übertragen werden. Dieses zeichnet sich durch ein doppeltes Schutzregime aus: Zum einen werden personenbezogene Daten im Gesundheitsbereich durch das zum Teil bereichsspezifisch ausgestaltete Datenschutzrecht geschützt, zum anderen unterliegen sie als Patientendaten den zivil-, straf- und berufsrechtlichen Vorgaben der ärztlichen Schweigepflicht.

Das Gesundheitsdatenschutzrecht ist indes denselben Grundgedanken verpflichtet wie das allgemeine Datenschutzrecht: Es ist ebenfalls auf die Einwilligung fokussiert und greift auch im Übrigen auf vergleichbare Regulierungsansätze zurück (vgl. zum Beispiel die einschlägigen Regelungen des GenDG). Zwar gibt es beispielsweise im Krebsregisterrecht Modifikationen des Einwilligungsmodells[179] und im Recht der Biobanken greift die Praxis zum Teil in rechtlich zweifelhafter Weise auf Globaleinwilligungen zurück[180]. Im Kern bleiben die normativen Lösungsansätze des Gesundheitsdatenschutzrechts jedoch weitgehend einer Problemperspektive aus der „Vor-Big-Data-Zeit“ verhaftet.[181]

---

[176] Vgl. Martini 2014, 1484.
[177] Vgl. zum Auskunftsanspruch nach Art. 22 Abs. 1, Art. 15 Abs. 1 lit. h DSGVO etwa Helfrich 2017, Rn. 77 ff.; siehe auch Martini 2014, 1485; ferner das Urteil des Bundesgerichtshofes in NJW 2014, 1235 ff.
[178] Vgl. Martini 2014, 1485.
[179] Zum Beispiel die Weitergabe von Daten für ein konkretes Forschungsprojekt ohne Einwilligung des Betroffenen auf der Grundlage einer behördlichen Zulassung; siehe etwa § 9 Hamburgisches Krebsregistergesetz.
[180] Siehe Thorbom 2015, 367.
[181] Zum Biobankenrecht ähnlich Berdin 2017, 54 ff., 159, 379: Der rechtliche Regelungsrahmen sei weiterhin zentral am Erfordernis einer individuellen Einwilligung orientiert, womit keine angemessene Regulierung der Biobankforschung zu leisten sei.

Besonders deutlich wird dies an der traditionellen Unterscheidung zwischen Daten unterschiedlicher Sensibilität.[182] Herkömmlich werden dabei zu den besonders sensiblen und damit besonders schutzbedürftigen Daten auch personenbezogene Daten mit Gesundheitsrelevanz gezählt. Sowohl die DSGVO (Art. 4 Nr. 13, 15, Art. 9, Erwägungsgrund 35, 52 ff.) als auch das Datenschutzrecht von Bund und Ländern (siehe etwa § 3 Abs. 9 BDSG a. F., § 46 Nr. 13 BDSG n. F., für Sozialdaten ferner § 35 SGB I, §§ 67 ff. SGB X und § 284 SGB V) enthalten entsprechende gesonderte Vorgaben für Gesundheitsdaten.

Die zugrunde liegende Annahme, die Sensibilität dieser Daten stehe aufgrund ihres Gesundheitsbezugs fest, trifft auf Big-Data-Konstellationen nicht oder allenfalls nur eingeschränkt zu. Dafür sind nämlich die eingangs beschriebenen Möglichkeiten kontinuierlicher De- und Rekontextualisierung „durch neue Zwecke (Bottom up) oder durch spontan generierte Korrelationen und Muster [charakteristisch], die eine Dynamik des Wissens in Anschlag bringen, die über variable Konstellationen prozessiert wird, nicht aber ‚die' Wirklichkeit abbilden".[183] Angesichts der möglichen Kombination molekularer und klinischer Patientendaten mit psychosozialen, soziodemografischen, Lebensstil- und anderen Daten zu einer „Gesamtgesundheitsinformation" eines jeden Patienten ist es äußerst zweifelhaft, ob sich weiterhin derartige kategorische Unterscheidungen treffen lassen. Unter Big-Data-Bedingungen gewinnen auch prima facie gesundheitsferne Daten, zum Beispiel das Einkaufsverhalten in einem Supermarkt, eine unter Umständen erhebliche Gesundheitsrelevanz. Big-Data-Analysen operieren zudem nahezu ausschließlich auf Basis von statistischen Korrelationen und Wahrscheinlichkeitswerten. Dadurch können beispielsweise Informationen über die genetische Veranlagung eines Menschen in Verbindung mit zusätzlichen, unter Umständen prima facie gesundheitsirrelevanten (Umwelt-)Informationen einen hohen Aussagewert gewinnen und es ermöglichen, spezielle Risikoprofile zu erarbeiten. Weder das geltende Datenschutzrecht noch das Gendiagnostikgesetz sind hinreichend auf die Besonderheiten der Analysemethoden für molekulare Daten aus der Humanmedizin und die mit Big Data einhergehenden Verknüpfungsmöglichkeiten zugeschnitten. Allgemein gesprochen, besteht eine erhebliche Inkongruenz zwischen den Charakteristika von Big Data und den bestehenden normativen Anforderungen an die Nutzung von Daten. Letztere sind – nicht nur, aber auch im Gesundheitswesen – auf die Nutzung begrenzter Daten durch legitimierte Nutzer zum Zweck der Versorgung eines bestimmten Patienten ausgerichtet.

Erschwerend tritt hinzu, dass die Vorgaben für den Schutz personenbezogener Daten im Gesundheitssektor besonders komplex und damit schwer nachvollziehbar sind. Zum einen unterliegen die personenbezogenen Daten im Gesundheitsbereich als Patientendaten den zivil-, straf- und berufsrechtlichen Vorgaben der ärztlichen Schweigepflicht. Zum anderen ergeben

[182] Vgl. Ohm 2015.
[183] Ladeur 2016, 361.

sich die spezifischen datenschutzrechtlichen Vorgaben aus einem komplizierten Zusammenspiel von bundes- und landesrechtlichen Regelungen. So unterfallen etwa die privaten Krankenversicherungen (nur) dem allgemeinen Datenschutzrecht. Im Bereich der gesetzlichen Krankenversicherung sind demgegenüber spezifische Regelungen der Sozialgesetzbücher (SGB V und X) sowie besondere landesrechtliche Vorschriften (etwa des Gesundheitsdatenschutzgesetzes NRW) zu beachten, wobei umstritten ist, ob und inwieweit dies auch die Leistungserbringer (etwa niedergelassene Ärzte und medizinische Versorgungszentren) betrifft. Für Krankenhäuser normiert das einschlägige Landeskrankenhausrecht teilweise ebenfalls besondere datenschutzrechtliche Vorgaben. Fehlen diese, unterliegen Häuser in privater Trägerschaft grundsätzlich dem Bundesdatenschutzgesetz, öffentliche Kliniken den jeweiligen Landesdatenschutzgesetzen, die zum Teil ihrerseits auf das BDSG verweisen. Für Häuser in kirchlicher Trägerschaft existieren kircheneigene Datenschutzregime.[184] Dieses komplexe Geflecht wird überwölbt von den vorrangig zu beachtenden Vorgaben der europäischen DSGVO.

3.2.3 Zwischenfazit

Zwischen den Anforderungen des traditionellen Datenschutzrechts und den Wirkungsbedingungen von Big Data besteht also eine erhebliche Diskrepanz. Selbst mit einem extrem hohen regulativen und organisatorischen Aufwand dürfte es deshalb kaum möglich sein, diese überkommenen Regelungsmechanismen friktionslos und effektiv auf die neue Situation anzuwenden. Das gilt zumal im selbstverwalteten Gesundheitssystem, das durch eine Vielzahl von unterschiedlichen Leistungserbringern mit teilweise gegenläufigen Interessen und entsprechenden Einflussnahmeversuchen gekennzeichnet ist (siehe Kapitel 1). Will man nun aber nicht aus den beschriebenen Regelungsdefiziten auf eine grundlegende Unzulässigkeit von Big-Data-Nutzungen schließen, die den datenschutzrechtlichen Vorgaben nicht (zur Gänze) genügen, müssen alternative, gleichermaßen bereichsadäquate und ein angemessenes Schutzniveau garantierende Mechanismen identifiziert werden. Weil aber andererseits auch eine Absenkung der verfassungsnormativ vorgegebenen Mindeststandards nicht in Betracht kommt, sind nicht nur mögliche Weiterentwicklungen des Datenschutzrechts, sondern auch kompensatorische Effekte weiterer, bereichsbezogener Regelungen in den Blick zu nehmen.[185]

3.2.4 Medizinprodukterecht

Entsprechende kompensatorische Wirkung könnten dabei zunächst die Bestimmungen des Medizinprodukterechts entfalten, das grundsätzlich das Ziel verfolgt, den freien Verkehr mit Medizinprodukten zu regeln und dabei gleichzeitig die Sicherheit, Eignung und Leistung der Medizinprodukte zum Schutz von Patienten, Anwendern und Dritten zu gewährleisten. Da es

[184] Vgl. umfassend Karaalp 2016, 59 ff.
[185] Vgl. demgegenüber zum allgemeinen Wettbewerbsrecht etwa Paal/Hennemann 2017, 1698 f. und Drexl 2017.

um die Produktsicherheit geht, kann man das Medizinprodukterecht auch als Teil des Verbraucherschutzrechts betrachten.[186]Die Rechtsgrundlagen des Medizinprodukterechts finden sich im Europarecht und nationalen Recht. Relevante Vorgaben enthalten vor allem das deutsche Medizinproduktegesetz (MPG) bzw. entsprechende Verweisungen, durch die auch drei europäische Richtlinien – für aktive implantierbare medizinische Geräte[187], für In-vitro-Diagnostika[188] und für Medizinprodukte[189] – umgesetzt werden.[190] Seit Kurzem existieren auf europäischer Ebene zwei neue Verordnungen: für Medizinprodukte[191] und für In-vitro-Diagnostika[192]. Beide traten am 25. Mai 2017 in Kraft und sind nach einer drei- bzw. fünfjährigen Übergangszeit verpflichtend anzuwenden.

Anders als Arzneimittel bedürfen Medizinprodukte keiner staatlichen Zulassung. Gemäß § 6 MPG dürfen Medizinprodukte aber grundsätzlich nur in den Verkehr gebracht oder in Betrieb genommen werden, wenn sie zertifiziert und mit einer CE-Kennzeichnung[193] versehen sind. Das Medizinprodukt muss dafür den grundlegenden Anforderungen entsprechen, die nach § 7 MPG in den Anlagen I zur Medizinprodukterichtlinie aufgelistet sind. Hierfür bedarf es einer produktspezifischen Risikobewertung, eines Verfahrens des Risikomanagements im Sinne einer Risikominimierung und einer Risiko-Nutzen-Analyse. Zudem und vor allem muss ein dem Risiko des Produkts angemessenes Konformitätsbewertungsverfahren erfolgreich durchlaufen worden sein. Das bedeutet: Je nachdem, in welche Risikoklasse das Produkt gemäß § 13 MPG eingestuft wurde[194], kann der Hersteller dies in eigener Verantwortung durchführen und damit selbst die Konformität bestimmen (Risikoklasse I), oder er muss eine „Benannte Stelle" beteiligen.[195]

---

[186] Vgl. hierzu und zum Folgenden etwa Zirfas 2017, 106 ff.; Pannenbecker 2013, Rn. 250.
[187] Richtlinie des Rates vom 20. Juni 1990 zur Angleichung der Rechtsvorschriften der Mitgliedstaaten über aktive implantierbare medizinische Geräte (90/385/EWG) (ABl. EG 1990 Nr. L 189/17).
[188] Richtlinie 98/79/EG des Europäischen Parlaments und des Rates vom 27. Oktober 1998 über In-vitro-Diagnostika (Abl. EG 1998 Nr. L 331/1).
[189] Richtlinie 93/42/EWG des Rates vom 14. Juni 1993 über Medizinprodukte (ABl. EG 1993 Nr. L 169/1).
[190] Die Richtlinien 90/385/EWG und 93/42/EWG wurden durch die Richtlinie 2007/47/EG vom 5. September 2007 (ABl. EU 2007 Nr. L 247/21) novelliert.
[191] Verordnung (EU) 2017/745 des Europäischen Parlaments und des Rates vom 5. April 2017 über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates (ABl. EU 2017 Nr. L 119/1).
[192] Verordnung (EU) 2017/746 des Europäischen Parlaments und des Rates vom 5. April 2017 über In-vitro-Diagnostika und zur Aufhebung der Richtlinie 98/79/EG und des Beschlusses 2010/227/EU der Kommission (ABl. EU 2017 Nr. L 119/176).
[193] Die CE-Kennzeichnung ist eine verbindliche Konformitätskennzeichnung, die angibt, dass ein Produkt mit den Harmonisierungsvorschriften der Europäischen Union übereinstimmt. Mit der Anbringung an ein Produkt erklärt der Hersteller gegenüber den Behörden, dass das Produkt den Vorschriften entspricht und den vorgeschriebenen Konformitätsbewertungsverfahren unterzogen wurde.
[194] Ausgenommen von dieser Anforderung sind In-vitro-Diagnostika.
[195] Zur Risikoklassifizierung von Stand-alone-Software im Einzelnen vgl. Oen 2009, 57.

Medizinprodukte unterliegen folglich keiner präventiven staatlichen Kontrolle, sondern lediglich einer nachgeschalteten Kontrolle in Form einer Marktüberwachung durch Landesbehörden und einem Medizinprodukte-Beobachtungs- und -Meldesystem durch eine Bundesoberbehörde (vgl. §§ 26 ff. MPG), nämlich das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) und das Paul-Ehrlich-Institut. Allerdings bedroht das MPG einen Verstoß gegen die genannten Normen (§§ 4, 6 MPG) mit Strafe (§§ 40, 41 MPG), einen Verstoß gegen § 4 MPG auch bei lediglich fahrlässigem Verhalten.

Mit Blick auf Big-Data-basierte Anwendungen ist von erheblicher Bedeutung, dass auch Software als Medizinprodukt zu klassifizieren sein kann. Voraussetzung hierfür ist eine in § 3 Nr. 1 MPG aufgelistete medizinische Zweckbestimmung.[196] Nach § 3 Nr. 10 MPG ist Zweckbestimmung die Verwendung, für die das Medizinprodukt in der Kennzeichnung, der Gebrauchsanweisung oder den Werbematerialien nach den Angaben des Herstellers (§ 3 Nr. 15 MPG) bestimmt ist. Aus diesen Regelungen folgt, dass es für die Zweckbestimmung eines Produkts und folglich auch für seine Qualifizierung als Medizinprodukt maßgeblich auf die Angaben des Herstellers ankommt.[197] Nur soweit sich die vom Hersteller definierte Zweckbestimmung als wissenschaftlich unhaltbar oder widersprüchlich erweist, stellt die Rechtsprechung auf die objektive Eignung des Produkts ab.[198]

Ersichtlich können hieraus erhebliche Abgrenzungsschwierigkeiten zwischen medizinischen Anwendungen und bloßen Lifestyle- oder Fitness-Apps resultieren, weil etwa Ernährungstipps oder Trainingsübungen je nach Kontext auch medizinischen Zwecken dienen können.[199] Zudem können Hersteller auch im medizinischen Kontext therapeutische oder diagnostische Funktionen ausschließen. Als Auslegungshilfen lassen sich die (rechtlich unverbindlichen) Leitlinien der Europäischen Kommission zu Stand-alone-Software[200] und das durch die Kommission veröffentlichte Handbuch zur Abgrenzung von Medizinprodukten[201] heranziehen. Apps, die keine Auswertung vornehmen, sondern sich auf die bloße Erfassung von Daten und deren grafische Darstellung beschränken, dienen keinen diagnostischen oder therapeutischen

[196] „a) der Erkennung, Verhütung, Überwachung, Behandlung oder Linderung von Krankheiten, b) der Erkennung, Überwachung, Behandlung, Linderung oder Kompensierung von Verletzungen oder Behinderungen, c) der Untersuchung, der Ersetzung oder der Veränderung des anatomischen Aufbaus oder eines physiologischen Vorgangs oder d) der Empfängnisregelung". Vgl. näher etwa Heimhalt/Rehmann 2014, 200 ff.
[197] Vgl. Heimhalt/Rehmann 2014, 201.
[198] Vgl. das Urteil des Bundesgerichtshofes in NJW-RR 2014, 46 (47).
[199] Vgl. Heimhalt/Rehmann 2014, 202.
[200] Vgl. Europäische Kommission 2012b.
[201] Vgl. Europäische Kommission 2015.

Zwecken.[202] So fallen auch Informationssysteme, die lediglich der Speicherung, Wiedergabe oder dem Transfer von Daten dienen, nicht unter das MPG.[203] Sie werden jedoch dann zum Medizinprodukt, sobald sie umfassende medizinische Daten von Patienten speichern, um eine Behandlung zu optimieren.[204] Ebenso fällt Software, die durch Zusammenführung und Auswertung von Patientendaten einen Diagnose- oder Therapievorschlag abgibt, eine Medikationsdosierung errechnet oder Laborwerte mit Referenzwerten abgleicht, unter das Medizinprodukterecht.[205]

Dennoch sind die für eine klinische Bewertung erforderlichen Kriterien bezogen auf M-Health-Software nicht immer deutlich erkennbar. Hinsichtlich der Anforderungen an die Bewertung im Einzelnen verweist § 19 Abs. 1 MPG auf die Richtlinien und gegebenenfalls einschlägige harmonisierende Normen.[206] Allerdings finden sich weder dort noch in der aktuellen Medizinprodukteverordnung entsprechende, die beschriebenen Abgrenzungsschwierigkeiten umfassend beseitigende Vorgaben.[207] Die Neuregelungen enthalten vor allem Anpassungen der bisherigen Rechtslage, aber keine grundlegenden Änderungen des Regelungskonzepts. So werden einige Medizinprodukte in höhere Risikoklassen eingestuft sowie für Benannte Stellen und Überwachungsbehörden höhere Anforderungen und schärfere Kontrollpflichten statuiert. Auch die Anforderungen an die klinische Bewertung sind detaillierter und umfangreicher ausgestaltet. Interessant ist insbesondere der Erwägungsgrund 19 der EU-Medizinprodukteverordnung.[208] Demnach „muss eindeutig festgelegt werden, dass Software als solche, wenn sie vom Hersteller speziell für einen oder mehrere der in der Definition von Medizinprodukten genannten medizinischen Zwecke bestimmt ist, als Medizinprodukt gilt, während Software für allgemeine Zwecke, auch wenn sie in Einrichtungen des Gesundheitswesens eingesetzt wird, sowie Software, die für Zwecke in den Bereichen Lebensstil und Wohlbefinden eingesetzt wird, kein Medizinprodukt ist. Die Einstufung der Software entweder als Produkt oder als Zubehör ist unabhängig vom Ort der Software und von der Art der Verbindung zwischen der Software und einem Produkt.“ Gerade der zweite Satz verdeutlicht, dass mittlerweile ein klares Bewusstsein für die Notwendigkeit besteht, zwischen als Medizinprodukte zu qualifizierenden Gesundheits-

---

[202] Vgl. Heimhalt/Rehmann 2014, 202 m. w. N.
[203] Vgl. Rübsamen 2015, 487; Europäische Kommission 2012b, 20.
[204] Vgl. Dierks, nach Gärtner 2010, 17.
[205] Vgl. Rübsamen 2015, 487; Europäische Kommission 2012b, 11, 20.
[206] Stimmen die Medizinprodukte mit harmonisierten Normen oder ihnen gleichgestellten Monografien des Europäischen Arzneibuches oder Gemeinsamen Technischen Spezifikationen, die das jeweilige Medizinprodukt betreffen, überein, greift gemäß § 8 MPG die Vermutungsregel, dass die Bestimmungen des MPG eingehalten sind.
[207] Vgl. zum Entwurf Heimhalt/Rehmann 2014, 203; Gassner 2015, 77.
[208] Verordnung (EU) 2017/745 des Europäischen Parlaments und des Rates vom 5. April 2017 über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates (ABl. EU 2017 Nr. L 119/1).

Apps einer- und Lifestyle-Apps andererseits genau zu unterscheiden. Eine solche bewusst vorgenommene Negativabgrenzung dürfte zukünftig bei der Qualifikation gesundheitsbezogener Apps als Medizinprodukt gebührend zu beachten sein.[209]

3.2.5 Big-Data-Dienste im Kontext der (gesetzlichen) Krankenversicherung

Nicht nur im Sinne eines Ausgleichs unzureichender Schutzstandards, sondern zumal im Sinne einer Sicherstellung einer dauerhaften Nutzung der mit Big Data verbundenen Chancen erweisen sich ferner die Vorgaben des Krankenversicherungsrechts als relevant. Dabei geht es vorliegend weniger um Apps, die von privaten oder gesetzlichen Krankenkassen angeboten werden.[210] Vielmehr sind aus rechtlicher Sicht drei besonders neuralgische Punkte hervorzuheben, die im Folgenden exemplarisch für das System der gesetzlichen Krankenversicherung (GKV) beleuchtet werden sollen: erstens die erforderliche Einpassung von Big-Data-basierten-Anwendungen in das bestehende Vergütungsmodell, zweitens die Kompatibilität der so generierten Erkenntnisse mit den epistemischen Standards des Leistungsrechts und drittens die spezifischen Möglichkeiten und Grenzen mit Blick auf die Beitragsgestaltung.

Der erste Komplex verdeutlicht, inwieweit auch über an sich datenschutzferne Finanzierungskonstellationen kompensatorischer (Daten-)Schutz gewährt werden kann. Denn bekanntlich zählt es zu den problematischen Aspekten der Internetökonomie, dass viele Dienstleistungen zwar vordergründig kostenfrei angeboten, letztlich aber die Nutzerdaten als finanziell relevante Gegenleistung eingesetzt werden. Vor diesem Hintergrund wird deutlich, dass eine separate Vergütung, namentlich die Einordnung von M-Health-Anwendungen in die Vergütung der gesetzlichen wie privaten Krankenversicherung, zumindest problemabspannend wirken kann, weil die damit sichergestellte Finanzierung den Anreiz reduziert, die Daten auf andere Weise zu verwerten und in datenschutzrechtlich problematischer Weise unkontrolliert weiterzuverbreiten. Dass dies auch schon im Rahmen des bestehenden Rechts prinzipiell möglich ist, zeigen einzelne Pilotprojekte der sogenannten Telemedizin, etwa die ambulante augenärztliche Betreuung von Kindern mit gestörter Entwicklung des beidäugigen Sehens, denen eine webbasierte Stimulationstherapie unter anderem über eine App für Tablet-PC und Smartphone angeboten wird. Die Vergütung der Leistungserbringer erfolgte in diesem Fall durch die Krankenkasse, allerdings nicht im Rahmen der Regelversorgung, sondern zunächst auf Basis einer einzelvertraglichen Regelung (im Sinne des allerdings zwischenzeitlich aufgehobenen § 73c SGB V) zwischen der Kasse und einem Qualitätsverbund von Augenärzten.[211]

[209] So Gassner 2016, 111.
[210] Hierzu im Überblick Aumann/Frank/Pramann 2016.
[211] Vgl. GKV-Spitzenverband 2014, 11.

Prinzipielle Bedenken, Big-Data-basierte Dienstleistungen auch im Rahmen der Regelversorgung der gesetzlichen Krankenversicherung anzubieten, bestehen nicht – solange deren Grundprinzipien beachtet werden. Das betrifft ganz zentral den der gesetzlichen Krankenversicherung zugrunde liegenden und eine spezifische „Blindheit" gegenüber bestimmten Erkenntnissen verlangenden Solidaritätsgrundsatz (etwa gegenüber Vorerkrankungen oder entsprechenden genetischen Dispositionen). Die privaten Krankenversicherungen folgen naturgemäß nicht einem entsprechend strikten Solidaritätsgedanken. Dennoch bzw. gerade deshalb ist aber bei ihnen genau darauf zu achten, dass zulässige Differenzierungen nicht in unzulässige Diskriminierungen und Stigmatisierungen übergehen.

Besondere Anforderungen ergeben sich daneben aus der übergreifenden Verpflichtung, mit den begrenzten Ressourcen des Gesundheitssystems schonend umzugehen. Das hieraus entwickelte komplexe Programm der Rechtskonkretisierung bezieht namentlich auch die Erkenntnisse der medizinischen Wissenschaft mit ein. Grundlegend fordert dementsprechend schon § 2 Abs. 1 S. 3 SGB V, dass „Qualität und Wirksamkeit der Leistungen [...] dem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen und den medizinischen Fortschritt zu berücksichtigen" haben. Mit dieser Verweisung erkennt das Rechtssystem eigene Erkenntnisgrenzen an und versteht Wissenschaft dabei zutreffend als dynamischen, niemals abgeschlossenen Prozess. Allerdings sind dem Gesetz keine präzisen Anforderungen dazu zu entnehmen, wie diesem Wissenschaftsgebot genügt werden kann.[212] In der Rechtspraxis sind unterschiedliche Strategien zu beobachten: Die sozialgerichtliche Rechtsprechung beschränkt sich teilweise (noch) auf eher klassische Bezugnahmen auf Expertenwissen, insbesondere in Form der Suche nach einem Konsens oder doch einer Mehrheitsmeinung der einschlägigen Fachkreise. Zunehmende Bedeutung besitzen allerdings die teilweise explizit (etwa in § 35 Abs. 1b S. 4, § 35a Abs. 1 S. 7 Nr. 2, § 35b Abs. 1 S. 5, § 139a Abs. 4 S. 1 SGB V[213]), teilweise nur implizit in Bezug genommenen Vorgaben der evidenzbasierten Medizin.[214] Indem das Recht diese in Bezug nimmt, weist es ihr eine über den fachinternen Entstehungskontext hinausreichende Bedeutung zu. Die auf den Umgang mit sich verändernden Erkenntnissen angepasste, spezifisch rationalisierte und formalisierte, ursprünglich auf die individuelle klinische Entscheidungsfindung des Arztes bezogene (Verfahrens-)Maxime gewinnt damit übergreifende normative Bedeutung.[215] An diesen Maßstäben hat sich auch die Beurteilung von Big-Data-gestützten Verfahren zu orientieren.

---

[212] Vgl. allgemein zur Konkretisierung der komplexen Vorgaben des SGB V Francke/Hart 2008; siehe auch Axer 2011, 203 ff.
[213] Siehe ferner § 73b Abs. 2 Nr. 2, § 137f Abs. 1 S. 2 Nr. 3, Abs. 2 S. 2 Nr. 1, § 139a Abs. 3 Nr. 3 SGB V, wo ebenfalls auf die Evidenzbasierung abgestellt wird. Vgl. hierzu auch Wigge 2000.
[214] Vgl. zu § 35b SGB V etwa Deutscher Bundestag 2005, 8.
[215] Vgl. dazu etwa Stallberg 2010, 6 ff.; ferner Hart 2000.

Dem Bundesversicherungsamt als der für die (bundesunmittelbaren) gesetzlichen Krankenkassen zuständigen Aufsichtsbehörde genügen die durch Smartphones, Fitness-Tracker oder ähnliche elektronische Geräte und Anwendungen vom Versicherten selbst an die Krankenkasse übermittelten personenbezogenen Daten nicht als valider Nachweis einer Teilnahme des Versicherten an einer qualitätsgesicherten Maßnahme im Sinne des § 65a Abs. 1 Nr. 3 SGB V. Schon deshalb fehlt es derzeit an der datenschutzrechtlichen Erforderlichkeit für die Erhebung, Verarbeitung und Nutzung dieser Daten.[216] Ungeklärt ist noch, ob unabhängig von diesen datenschutzrechtlichen Erwägungen grundsätzliche Überlegungen gegen eine auch auf M-Health-Anwendungen und den auf diesem Wege gesammelten Daten basierende Beitragsgestaltung sprechen. Auf der einen Seite bestehen ersichtlich erhebliche Gefahren einer Entsolidarisierung; denn es gehört gerade zu den Grundprinzipien der gesetzlichen Krankenversicherung, gegenüber individuellen Morbiditätsrisiken „blind" zu sein. Auf der anderen Seite kann aber nicht verkannt werden, dass der Aspekt der Eigenverantwortlichkeit und der Prävention im SGB V zunehmend an Bedeutung gewinnt.[217]

3.3 Regelungsoptionen

Die mit dem Schlagwort Big Data umschriebene Öffnung der Nutzung der in der Patientenversorgung gewonnenen Daten für weitere Anwendungen und für die Gesundheitsforschung wirft neue Fragen und Probleme auf, vergrößert die Komplexität und kann zu Unübersichtlichkeit führen. Das herkömmliche Datenschutzrecht ist, wie gezeigt, auf diese Herausforderungen nur unzureichend eingestellt. Eine strikte Anwendung der vorhandenen Regelungen bedeutete deshalb letztlich, Big-Data-Anwendungen sowohl in Deutschland wie auf der Ebene der Europäischen Union für weitgehend unzulässig zu erklären. Dem steht indes entgegen, dass die Nutzung der in der Patientenversorgung gewonnenen Daten für die Gesundheitsforschung grundsätzlich zu begrüßen ist, da damit ihre Erkenntnisgrundlagen erheblich erweitert und ihre Effektivität erheblich gesteigert werden können. Dem kann und sollte im Rahmen der durch das Verfassungsrecht gewährten Handlungsspielräume Rechnung getragen werden. Der notwendigen Stabilisierungsfunktion des Rechts korrespondiert gerade in modernen und stark in der Entwicklung befindlichen Lebensbereichen die Erforderlichkeit hinreichend flexibler, innovationsoffener Regelungen. Der Regulierungsrahmen muss demnach den grundlegenden ethischen und verfassungsnormativen Vorgaben entsprechen; er muss aber gleichzeitig so ausgestaltet sein, dass Neuerungen nicht verhindert werden.[218] Das spricht etwa gegen die Etablierung eines umfassenden Vorsorgeprinzips. Es lässt sich zudem als Argument für die Verwen-

[216] Vgl. Deutscher Bundestag 2016, 9.
[217] Vgl. die Beiträge in Weilert 2015.
[218] Vgl. in diese Richtung schon Roßnagel/Nebel 2015, 459.

dung komplexerer, privatrechtliche wie privat-staatlich kooperative Steuerungsbeiträge mit berücksichtigender Regulierungsstrategien verstehen. Gerade dort, wo private Regeln keine reine Selbstregulierung implizieren, bedarf es institutioneller Absicherungen, gegebenenfalls auch mittels hoheitlicher Vorgaben. Diese sollten sich dabei nach Möglichkeit weniger auf den konkreten Inhalt als auf die organisatorischen wie prozeduralen Arrangements beziehen. Eine entsprechend begleitete private Regelsetzung kann einen wichtigen Baustein im Rahmen umfassender konzipierter, koordiniert-kooperativer und zeitlich gestaffelter Rechtsetzungsprozesse bilden.

An dieser Stelle kann es nicht darum gehen, dem Gesetzgeber konkrete Vorschläge hinsichtlich einer präzisen Neuordnung des bestehenden Normengefüges zu unterbreiten. Wegen der Vielzahl ihrer Einsatzmöglichkeiten ist es bereits kaum möglich, eine einheitliche und abschließende Beurteilung der rechtlichen Zulässigkeit von Big-Data-Anwendungen im Gesundheitswesen zu treffen. Erst recht können keine umfassenden und detaillierten, sämtliche beschriebenen Probleme aufnehmenden Regelungskonzepte entfaltet werden. Angesichts der jüngst erfolgten umfassenden Neuordnung des Datenschutzrechts durch die DSGVO und das BDSG n. F. ist zwar noch nicht abschließend einzuschätzen, ob und wie sich die neuen Normen und Mechanismen bewähren werden. Indes dürfte in inhaltlicher Hinsicht nach dem Vorgesagten feststehen, dass einige Grundprinzipien des geltenden Datenschutzrechts mit dem Konzept von Big Data kaum in Einklang zu bringen sind. Mit Blick auf die bleibende, verfassungsfundierte Bedeutung des Datenschutzes kann dies nicht schlicht zu einer Anpassung in dem Sinne führen, dass dessen Anwendungsbereich reduziert wird. Ziel darf nicht eine bloße Deregulierung, sondern kann nur eine Umstellung auf eine bereichsadäquate *smart regulation* sein. Eine entsprechende Fortentwicklung der bestehenden Datenschutzkonzeption sollte die Verbreitung von Daten weder nur als problematisch verstehen noch mit ihr verbundene reale Gefahren verharmlosen. Anzustreben ist ein differenziertes, Chancen wie Risiken berücksichtigendes Gestaltungs- und Regulierungskonzept: „In der Informationsgesellschaft erhält der Datenschutz eine neue Qualität, aber nicht nur deshalb, weil so viele Daten anfallen, genutzt und gespeichert werden, sondern vor allem deshalb, weil ein neues Datenschutzkonzept gefordert ist. In den Informationsströmen darf nicht mehr nur [...] eine Gefahr für Persönlichkeitsrechte gesehen werden, sondern Information und Kommunikation müssen als Chance, sogar als Grundlage der Entfaltung der Menschen, gesehen und damit im Hinblick auf Möglichkeiten der Teilhabe an Kommunikationsprozessen gewertet werden.“[219] Dementsprechend dienen die nachfolgenden Ausführungen vor allem dazu, in einem stärker formal orientierten Sinne auf die unterschiedlichen zur Verfügung stehenden Handlungsoptionen hinzuweisen. Innerhalb der über-

[219] Hoffmann-Riem 2000, 55.

greifenden, nachfolgend näher zu erläuternden Vorstellung eines Instrumenten- bzw. Regelungsmixes werden dabei allerdings immer wieder exemplarisch einzelne Steuerungsdesiderata herausgegriffen.

3.3.1 Weiterentwicklung bestehender Gesetze

Zunächst und vor allem bleiben selbstredend Gesetze die zentralen Steuerungsinstrumente des demokratischen Verfassungsstaates. Namentlich grundrechtswesentliche Entscheidungen sind weiterhin (nur) hier vorzunehmen. Allerdings ist in diesem Zusammenhang noch einmal auf die beschriebene mit Big Data verbundene Schwierigkeit hinzuweisen, weiterhin in bestimmten Bereichen Daten mit einem besonderen Sensibilitätsgrad zu identifizieren. Zumindest langfristig könnte es sich deshalb als sinnvoll erweisen, auf bereichsspezifische Regelungen – etwa des Sozialdatenschutzes – zu verzichten und stattdessen möglichst einheitliche Vorgaben festzulegen.[220] Schon an dieser Stelle ist im föderalen System allerdings die Pluralität der Gesetzgeber zu berücksichtigen. Darüber hinaus sind in eine künftige Regelungsstrategie schon auf Gesetzgebungsebene – ungeachtet der bestehenden, nicht allein terminologischen Unterschiede – notwendig auch die unionalen Rechtsetzungsinstanzen mit einzubeziehen. Inhaltlich ist dabei im Sinne der eingangs erwähnten Rekalibrierung des bestehenden Steuerungsansatzes insbesondere zu überlegen, ob der Mangel an Konkretheit von Big-Data-Anwendungen (im Gesundheitswesen) durch zusätzliche technisch-organisatorische sowie materiell- und verfahrensrechtliche Sicherungen kompensiert werden kann.[221] Vorbildhaft wirken kann hier etwa die Regelung zur Datenportabilität (Art. 20 DSGVO). Sie verpflichtet nicht dazu, bestimmte technisch kompatible Datenverarbeitungssysteme einzuführen, soll aber doch interoperable Formate fördern und trägt damit zum Be- bzw. Entstehen eines Wettbewerbs um datenschutzfreundliche bzw. datenschutzfreundlichere Technologien bei.[222] In diesem Sinne ist etwa zu fragen, inwieweit sich die Herausforderungen, die sich aus den kontinuierlichen De- und Rekontextualisierungen ergeben, durch eine komplexere Einwilligungskonzeption, durch klarere Verantwortungszuweisungen und durch (technische, aber normativ eingeforderte) ex post ansetzende Einwirkungsoptionen bewältigen lassen. Beispielhaft ist an folgende Ansätze zu denken:

Gerade mit Blick auf das Einwilligungserfordernis hat sich gezeigt, dass es einer konsequenten Weiterentwicklung der festgefahrenen und dysfunktionalen Vorgaben des Datenschutzrechts bedarf. An die Stelle einer pauschalen Einheitslösung müsste eine stärker ausdifferenzierte, den Besonderheiten eines Regelungsbereichs und den Präferenzen der Betroffenen Raum gebende

[220] Ähnlich Fetzer 2015, 778.
[221] Vgl. Weichert 2014b, 835.
[222] Vgl. Paal/Hennemann 2017, 1701.

Konzeption treten.[223] Die in bioethischen Kontexten schon länger diskutierten (siehe Kapitel 4)[224] und im Gesundheitsrecht zumindest ansatzweise bereits verwendeten (siehe Abschnitt 3.2.2)[225] Möglichkeiten modifizierter Einwilligungserfordernisse können hierzu ebenso beitragen wie die mit dem Stichwort „Einwilligungsassistent" umschriebenen technischen Konzepte.[226]

Daneben ist aber auch an die datenschutzrechtlich grundsätzlich zulässige Möglichkeit zu erinnern, nicht nur Opt-in-, sondern auch Opt-out-Modelle zu wählen, also die Erhebung und Nutzung von Daten auf Basis gesetzlicher Erlaubnisnormen durchzuführen. Denn mit Blick auf die Datenqualität und die Bedeutung möglichst vollzähliger Datensätze (Repräsentativität) kann ein verabsolutiertes Einwilligungskonzept Probleme verursachen. Weil Big-Data-Analysen vor allem auch den Zugriff auf Datenbestände benötigen, die ursprünglich für andere Zwecke angelegt wurden, würde erstens ein striktes und umfassendes Einwilligungserfordernis die Zahl der zur Verfügung stehenden Daten massiv beschränken. Weil hierdurch nur Teile der betroffenen Daten erfasst werden, erzeugt zweitens so gerade ein umfassendes datenschutzrechtliches Einwilligungserfordernis bestimmte Verzerrungen. Das kann die von Big-Data-Anwendungen erhofften positiven Effekte jedenfalls reduzieren. Ersichtlich besteht mithin ein Widerspruch zwischen der (Ideal-)Vorstellung einer – jedenfalls dem Grundansatz nach – freiwilligen Datenübermittlung und dem gleichzeitigen Streben nach einer möglichst vollständigen und repräsentativen Datengrundlage.[227]

Entsprechende gesetzliche Erlaubnisregelungen enthalten namentlich die §§ 22, 27 BDSG n. F., die damit in europarechtskonformer Weise von dem in Art. 9 Abs. 1 DSGVO statuierten Verbot der Verarbeitung besonderer Kategorien personenbezogener Daten abweichen. Nach Art. 9 Abs. 2 DSGVO kann das nationale Recht Ausnahmen von diesem Verbot normieren. Dementsprechend legt § 22 Abs. 1 BDSG n. F. fest, unter welchen Voraussetzungen die Verarbeitung besonderer Kategorien personenbezogener Daten ausnahmsweise zulässig ist, ohne damit unmittelbar sich aus Art. 9 Abs. 2 DSGVO ergebende oder auf Grundlage der Verordnung erlassene bereichsspezifische Regelungen auszuschließen.[228] Die Vorschrift setzt dabei eine kom-

[223] Eine gewisse Bedeutung besitzt die pauschale Einwilligung (nur) noch im strafrechtlichen Kontext: Sie wirkt mit Blick auf § 203 StGB tatbestandsausschließend, und zwar auch dann, wenn sie nicht den datenschutzrechtlichen Anforderungen des § 4a Abs. 1, Abs. 3 BDSG genügt. Vgl. Heimhalt/Rehmann 2014, 204 f.
[224] Siehe ferner etwa Budin-Ljøsne et al. 2017; Kaye et al. 2014; Steinsbekk/Kåre Myskja/Solberg 2013; Wee 2013; Williams et al. 2015.
[225] Siehe ferner Richter/Buyx 2016 m. w. N. Auch die Stellungnahmen des Nationalen Ethikrates (2004) und des Deutschen Ethikrates (2010) zu Biobanken weisen darauf hin, dass ein eng gefasstes Einwilligungskonzept problematisch ist. Siehe ähnlich auch Europäische Kommission 2012c, 57 f.
[226] Siehe hierzu zuletzt ausführlich Riechert 2016, insbesondere 24 ff.
[227] Vgl. exemplarisch mit Blick auf das Transplantationsregistergesetz Augsberg 2016.
[228] Vgl. Deutscher Bundestag 2017, 93.

plexe Interessenabwägung voraus (siehe Kapitel 1) und ergänzt diese um prozedurale und organisatorische Anforderungen (siehe Abschnitt 2.1). Ähnlich erlaubt § 27 BDSG auf Basis einer umfassenden Interessenabwägung insbesondere auch die (Weiter-) Verarbeitung zu wissenschaftlichen Zwecken. Auf diese Weise wird die Regelung des Art. 5 Abs. 1 lit. b DSGVO, der zufolge eine Weiterverarbeitung für wissenschaftliche oder historische Forschungszwecke und für statistische Zwecke nicht als unvereinbar mit den ursprünglichen Zwecken gilt und deshalb auf die bereits für die Erstverarbeitung geltende Rechtsgrundlage gestützt werden kann, auf den Bereich der besonderen Kategorien personenbezogener Daten ausgedehnt.[229] Diese Erlaubnisregelungen müssen dann sowohl beim allgemeinen zivil- und strafrechtlichen Schutz personenbezogener Daten, als auch im Rahmen des besonderen Schutzes von Patientendaten gleichermaßen berücksichtigt werden.

Der damit bereits angesprochene Schutz der Datenqualität könnte durch weitere legislative Maßnahmen flankiert werden. Das betrifft bereits die Datengewinnung. Hier muss spezifischen Verzerrungen (*biases*) entgegengewirkt werden, die wie gesehen etwa auch als ungewollte Nebenfolge des Einwilligungserfordernisses ergeben können. Es zählt zu den allgemeinen Herausforderungen von Big Data, diskriminierungsfreie[230] und aussagekräftige[231] Datensätze sicherzustellen. Zum anderen ist indes nicht nur fraglich, wie kritisch die Zuverlässigkeit der Datenauswertung und der Datenaussagen von den Herstellern und Dienstleistern selbst überprüft wird. Zumindest teilweise dürfte diesen auch eine belastbare Überprüfung der Reliabilität und Validität der Daten kaum möglich sein. Das gilt etwa, wenn Verbraucher oder Patienten diese selbst bei Benutzung einer App eintragen oder mit ihrem Verhalten den Mess- und Analyseprozess beeinflussen können. Dies ist selbstredend besonders kritisch zu sehen, wenn an die ausgewerteten Daten eine Behandlungsstrategie anknüpfen soll, wirft aber auch spezifische Probleme auf, wenn entsprechend generierte Daten für Forschungszwecke genutzt werden. Prospektiv ist deshalb nach Möglichkeiten zu forschen, um wissenschafts- und wirtschaftsinterne Qualitätssicherungsmechanismen noch stärker hoheitlich unterstützen, aber gegebenenfalls auch überprüfen und absichern zu können. Sollten sich diese Maßnahmen insgesamt als wenig zielführend erweisen, könnten mittel- bis langfristig auch rigidere rechtliche Vorgaben etabliert werden.

3.3.2 Regulierungsfunktion des Privatrechts

Neben der Neujustierung der Datenschutzgesetze kommt dem Privatrecht eine große Bedeutung für die Weiterentwicklung des Datenschutzes zu. Verdeutlichen lässt sich das hier bestehende regulatorische Potenzial[232] anhand von drei Referenzbereichen: dem Verbraucherrecht,

[229] Vgl. ebd., 98.
[230] Vgl. etwa Barocas/Selbst 2016, 673.
[231] Siehe hierzu etwa Kim/Huang/Emery 2016.
[232] Vgl. Binder 2012, 50 ff.; Hellgardt 2016; Poelzig 2012. Speziell für den Datenschutz siehe auch Buchner 2006.

dem Haftungsrecht sowie den Regelungen für die Zuordnung von Daten und die Befugnis, über ihre Verwendung zu bestimmen („Eigentum“ an Daten).

Das Verbraucherschutzrecht ist dabei von doppeltem Interesse: Erstens basiert es, ähnlich wie das Datenschutzrecht, auf der Erkenntnis, dass bestehende Informationsasymmetrien zu missbräuchlichen Verwendungen führen können. Denn durch derartige tatsächliche Einschränkungen kann die grundrechtlich gewährleistete Privatautonomie eingeschränkt sein und eine ungleiche Verhandlungsposition bestehen. Solche Nachteile sollen durch das auf Verträge bezogene Verbraucherschutzrecht abgemildert werden. Der Verbraucher soll vor Täuschungen und Übervorteilungen im Wirtschaftsleben geschützt werden. Darüber hinaus dient das Verbraucherschutzrecht auch der Sicherstellung, dass der Verbraucher als Rechtssubjekt generell befähigt ist, optimale Marktentscheidungen zu treffen.[233] Interessant hieran sind zweitens die im Vergleich zum Datenschutzrecht deutlich komplexeren Kombinationen von Informationspflichten, Widerrufsoptionen und Vertragsabschlusserfordernissen. Drittens bestehen Überschneidungen, soweit im Vertragsrecht diskutiert wird, ob und inwieweit das Zurverfügungstellen von personenbezogenen Daten als geldwerte Gegenleistung anzuerkennen ist.[234]

Das Haftungsrecht (hier verstanden als Haftung für den missbräuchlichen Umgang mit Daten) ist in Deutschland charakterisiert durch den negatorischen, unabhängig von einem etwaigen Verschulden bestehenden Anspruch auf Unterlassung des missbräuchlichen Umgangs einerseits und den Anspruch auf Schadensersatz andererseits, der nach allgemeinem Zivilrecht Verschulden voraussetzt, soweit nicht besondere gesetzliche Haftungsregelungen (zum Beispiel des Produkthaftungsgesetzes und/oder des Medizinproduktegesetzes) eingreifen. So ist beispielsweise bei Medizinprodukten[235] die Haftung in der Erprobungsphase von der Haftung in der Nutzungsphase zu unterscheiden, und es sind die diversen potenziell haftbar zu machenden Betroffenen zu identifizieren (Krankenhäuser bzw. Ärzte, Sponsoren, Hersteller, Zulieferer, Vertriebshändler und Benannte Stellen).[236] Hier gilt es nach neuen Wegen zu suchen, um die bestehenden Haftungsansprüche wirksamer durchzusetzen oder neue und wirksamere Sanktionen bei der unbefugten Erhebung und Verwendung von Daten zu etablieren. Die DSGVO stellt hier bereits einen bedeutenden Fortschritt dar. Sie sieht neben strafrechtlichen Sanktionen (Art. 83, 84 DSGVO sowie § 42 BDSG n. F.) und Bußgeldern (Art. 83 DSGVO, § 43 BDSG n. F.) auch einen Schadensersatzanspruch des Betroffenen bei unbefugter Datenverwendung gegen den Verantwortlichen und den Verarbeiter vor, der kein Verschulden voraussetzt und den Ersatz von materiellen wie immateriellen Schäden umfasst (Art. 82 DSGVO). Von dieser Haftung kann sich nur befreien, wer nachweist, dass er „in keinerlei Hinsicht für den Umstand, durch

[233] Vgl. Tamm 2011, 19.
[234] Vgl. Fezer 2017a, 100; Specht 2017, 763 ff.
[235] Vgl. dazu etwa Reich 2014.
[236] Näher Ortner/Daubenbüchel 2016, 2921 ff.

den der Schaden eingetreten ist, verantwortlich ist“ (Art. 82 Abs. 3 DSGVO). Angesichts des mittlerweile unbestreitbaren Umstandes, dass Daten ein erhebliches ökonomisches Wertschöpfungspotenzial haben, steht die regulatorische Wirkung solcher haftungsrechtlichen Ansätze außer Zweifel. Allerdings gewährleisten weder die DSGVO noch das deutsche Haftungsrecht einen Ausgleich dafür, dass Big-Data-Anwendungen wegen der mit ihnen verbundenen Chancen im Allgemeininteresse zugelassen werden, die wirtschaftlichen Vorteile jedoch – jedenfalls zunächst – bei den Datenverwendern liegen, während die Datengeber die Risiken tragen. Denn wenn zum Beispiel eine Person über ihre von einem Unternehmen befugterweise genutzten Daten reidentifiziert wird und etwa in der Folge einen gewünschten Versicherungsschutz nicht erhält oder verliert, begründet dies nach Art. 82 DSGVO nur dann einen Schadensersatzanspruch, wenn die Datenverwendung unbefugt erfolgt.[237] Materielle oder immaterielle Schäden, die diese Person infolge einer befugten Datenverwendung erleidet, werden hingegen weder nach Art. 82 DSGVO noch nach deutschem Haftungsrecht ersetzt.

Es ist deshalb von besonderem Interesse, wem diese ökonomischen Potenziale auf welche Weise zugerechnet werden können. Hier bestehen insbesondere in der klinischen Diagnostik und der medizinbezogenen Forschung große Unsicherheiten. Der umgangssprachlich genutzte Begriff „Eigentum an Daten“ bildet den Ausgangspunkt für zahlreiche Konflikte. Es gibt eine Vielzahl von Personen, die aufgrund der Tatsache, dass sie einen wichtigen Beitrag zur Erfassung, Analyse und/oder Verknüpfung von Daten leisten, sich als deren Eigentümer bzw. Miteigentümer (miss-)verstehen und deshalb ausschließliche Nutzungsrechte für sich reklamieren. Demgegenüber ist festzuhalten, dass ein Eigentum an Daten im Rechtssinne nicht existiert. Eigentum kann nur an (beweglichen oder unbeweglichen) Sachen bestehen. Daten sind aber gerade keine körperlichen Gegenstände und daher keine Sache. Eine Anwendung der Regelungen für das Sacheigentum auf Daten scheidet daher nach geltendem Recht aus.[238] Daten unterfallen für sich gesehen mangels einer ausreichenden Bearbeitung auch nicht den Regeln des Immaterialgüterrechts und können daher nicht einmal als „geistiges Eigentum“ bezeichnet werden.[239] Die oben beschriebene eigentumsanaloge Ausgestaltung des Rechts auf informationelle Selbstbestimmung ändert hieran ebenso wenig wie das neue Recht auf Datenportabilität aus Art. 20 DSGVO. Dieses besitzt zwar faktische Auswirkungen auf die Zuordnung von Rechten an Daten, verhält sich aber nicht zur Frage des Eigentums.[240]

Hält man die derzeitige Rechtslage für unbefriedigend und möchte sie weiterentwickeln, bedarf es kreativer rechtspolitischer Überlegungen. Gerade für die besonders drängende Problematik,

---

[237] Vgl. Kreße 2017, Rn. 18 ff.
[238] Vgl. etwa Ensthaler 2016, 3475 f.; Paal/Hennemann 2017, 1698; Wiebe/Schur 2017, 463 f.
[239] Vgl. Ensthaler 2016, 3473 f. Denkbar ist demnach nur ein Leistungsschutz für bereits existierende (nicht hingegen für erst zu generierende) Daten über das Datenbankrecht der §§ 87a ff. UrhG. Siehe dazu näher Wiebe 2017.
[240] Vgl. Keßler 2017, 591 m. w. N.

wem die ökonomisch und wissenschaftlich relevanten Daten von Menschen „gehören", die von Forschern und/oder Unternehmen bearbeitet werden, liegen bereits interessante Forschungsansätze vor. Analysiert wird etwa, ob und inwiefern das Datenschutzrecht den betroffenen Personen, also den Personen, deren Daten bearbeitet werden, bereits heute eine Rechtsposition verschafft, die einem Eigentum an Personendaten – im Sinne eines übertragbaren Ausschließlichkeitsrechts – zumindest nahekommt. Auf dieser Basis werden unterschiedliche Varianten für die Ausgestaltung eines Eigentums an Personendaten und deren Inkorporation in die bestehende Eigentumsordnung untersucht.[241] An anderer Stelle finden sich erste Überlegungen für ein „originäres Immaterialgüterrecht sui generis an verhaltensgenerierten Informationsdaten der Bürger"[242], das den „Nutzern als Datenproduzenten ein eigentumsrechtliches Abwehr- und Vermögensrecht"[243] verschaffen soll.

3.3.3 Möglichkeiten grenzüberschreitender Regulierung

Die erforderliche normative Begleitung hat insbesondere im Bereich von Big Data eine starke internationale Komponente. Die beschriebenen klassischen Steuerungsmodi sind sämtlich mit dem Problem konfrontiert, mit einer territorial begrenzten Rechtsetzung auf ein seiner Natur nach nicht an nationale Grenzen gebundenes, in digitalen Netzen verbundenes Datenallokationsphänomen zu reagieren.[244]

Die jeweiligen Datenschutzrechte sind jedoch international gesehen sehr unterschiedlich, was sowohl die Betroffenen als auch die Regulierungsinstanzen vor besondere Herausforderungen stellt: Datenströme sind zum großen Teil cloudbasiert. Server und Firmen, die Daten erheben, weiterleiten und auswerten, sind ebenfalls zu einem großen Teil nicht in Deutschland angesiedelt bzw. physisch verortet. Demgegenüber unterfallen etwa nach dem sogenannten Territorialprinzip (nur) solche verantwortlichen Stellen dem Anwendungsbereich der Datenschutzgesetze, die personenbezogene Daten in Deutschland erheben, verarbeiten oder nutzen. Entsprechende Vorgaben enthalten auch die speziellen datenschutzrechtlichen Bestimmungen für Telemedien- bzw. Telekommunikationsdienste (§§ 11 ff. Telemediengesetz; §§ 91 ff. Telekommunikationsgesetz). Hier bringt die DSGVO eine deutliche Verbesserung (insbesondere Kapitel V: Übermittlungen personenbezogener Daten an Drittländer oder an internationale Organisationen, Art. 44 ff.).

[241] Hierzu Thouvenin 2017.
[242] Fezer 2017a, 99 ff.
[243] Fezer 2017b, 3 ff.
[244] International Bioethics Committee 2017.

Die grenzüberschreitende Dimension stellt unter anderem auch besondere Anforderungen an die Zusammenarbeit regionaler Datenschutzbehörden bzw. erfordert eine komplexe internationale Koordination.[245] Ein entsprechendes Instrument im europäischen Raum ist der Europäische Datenschutzbeauftragte (EDSB). Zu den mit dem EDSB kooperierenden Institutionen gehören unter anderen Europol (Strafverfolgung), das Schengener Informationssystem (SIS, eine Datenbank, die für die polizeiliche Zusammenarbeit und den Grenzschutz eingesetzt wird), EURODAC (eine Datenbank, die Fingerabdrücke von Asylbewerbern und irregulären Einwanderern in die EU enthält) oder auch der Europarat, um nur einige zu nennen. Fragen zum Thema Gesundheitsdaten sind Gegenstand der Zusammenarbeit von EDSB und OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung). Letztere hat als Zusammenschluss aus derzeit 35 Mitgliedstaaten Richtlinien zum Umgang mit personenbezogenen Daten in der Wirtschaft verabschiedet, in denen auch Empfehlungen zu Gesundheitsdaten enthalten sind.[246] Der EDSB beteiligt sich außerdem an Netzwerken, um regionale Initiativen zu unterstützen, deren Ziel es ist, den Datenschutz weltweit zu stärken. Dies sind unter anderem das globale Netzwerk für die Durchsetzung des Rechts auf Schutz der Privatsphäre (Global Privacy Enforcement Network, GPEN), das Forum der asiatisch-pazifischen Datenschutzbehörden (Asia Pacific Privacy Authorities, APPA), der französischsprachige Verbund von Behörden, die für den Schutz personenbezogener Daten zuständig sind (Association Francophone des Autorités de Protection des Données Personnelles, AFAPDP), und das iberoamerikanische Datenschutznetzwerk (Red Iberoamericana de Protección de Datos, RIPD).

Als Nachfolgeregelung zu dem bis dahin bestehenden Safe-Harbor-Abkommen[247] zwischen der Europäischen Union und den Vereinigten Staaten ist am 1. August 2016 der EU-US-Datenschutzschild (Privacy Shield)[248] in Kraft getreten. Beide Abkommen ähneln einander mit Blick auf die verwendete Regelungstechnik: Mit Blick auf die Datenschutzverpflichtungen US-amerikanischer Unternehmen wird ein Selbstregulierungssystem etabliert, in dem diese sich freiwillig zur Einhaltung bestimmter, näher spezifizierte *privacy principles* verpflichten. Im Vergleich zum Safe-Harbor-Abkommen enthält der Privacy Shield eine Stärkung der Betroffenenrechte, soweit teilnehmende Unternehmen und US-Behörden Beschwerdestellen einrichten müssen und sich Betroffene nunmehr mit Beschwerden über ihre nationalen Datenschutzbe-

[245] Vgl. näher etwa Lewinski/Herrmann 2016.
[246] Siehe OECD 2013.
[247] Siehe Entscheidung der Kommission vom 26. Juli 2000 gemäß der Richtlinie 95/46/EG des Europäischen Parlaments und des Rates über die Angemessenheit des von den Grundsätzen des „sicheren Hafens" und der diesbezüglichen „Häufig gestellten Fragen" (FAQ) gewährleisteten Schutzes (ABl. EG 2000 Nr. L 215/7).
[248] Siehe Durchführungsbeschluss (EU) 2016/1250 der Kommission vom 12. Juli 2016 gemäß der Richtlinie 95/46/EG des Europäischen Parlaments und des Rates über die Angemessenheit des vom EU-US-Datenschutzschild gebotenen Schutzes (ABl. EU 2016 Nr. L 207/1). Vgl. zum Folgenden Towfigh/Ulrich 2017, Rn. 44 ff. m. w. N.; Determann/Weigl 2016, 811 ff. m. w. N.

hörden auch an das US-Handelsministerium wenden können. Vorgesehen ist zudem eine regelmäßige Überprüfung der Einhaltung und Effektivität des Privacy Shields durch die Europäische Kommission.

Unabhängig von den Bemühungen um eine Harmonisierung des Datenschutzrechts und eine bessere Kooperation und Koordination auf dem Gebiet des Datenschutzes geht es aus der Sicht der Betroffenen in erster Linie um die grenzüberschreitende Geltendmachung ihrer Rechte und deren Anerkennung durch eine ausländische Rechtsordnung, das heißt um Fragen des Internationalen Verfahrensrechts und des Internationalen Privatrechts. Die für die Durchsetzung der Rechte zum Schutz personenbezogener Daten maßgeblichen Regelungen sind im europäischen Rechtsraum weitgehend harmonisiert, weltweit jedoch weiterhin sehr unterschiedlich. Selbst dort, wo sie harmonisiert sind, gibt es jedoch zahlreiche praktische Hindernisse, die einer effektiven Rechtsverfolgung im Wege stehen.

3.3.4 Ergänzungsfunktion nicht hoheitlicher Steuerungsinstrumente

Ungeachtet der diesbezüglich geäußerten, teilweise scharfen Kritik[249] verdeutlicht das oben genannte Beispiel des Privacy Shield, warum es gerade mit Blick auf die beschriebenen territorialen Begrenzungen klassischer rechtlicher Steuerung, aber auch angesichts der spezifischen Dynamik und Volatilität des Regelungsbereichs angezeigt sein kann, auch nicht hoheitliche und kooperative Steuerungsmechanismen mit einzubeziehen.[250] Entsprechend schlägt ein Branchenverband der digitalen Wirtschaft vor, mittels „Zertifizierung mit einem anerkannten Datenschutz- bzw. Datensicherheits-Siegel (zum Beispiel ePrivacyseal, ULD-Siegel) […] für transparente Prozesse, Vertrauen und Sicherheit zu sorgen".[251] Neben eher allgemein gehaltenen Verhaltensempfehlungen[252] existieren dabei auch konkrete Vorschläge für komplexere Regelungsarrangements, beispielsweise vom Rat für Informationsinfrastrukturen (RfII).[253]

Eine derartige Ausweitung des Blickfeldes auch in den Bereich der Selbst- bzw. Koregulierung besitzt im Kontext des Datenschutzrechts eine gewisse Tradition.[254] Sie sollte in ihrer Bedeutung nicht überschätzt werden[255] und darf in grundrechtssensiblen Bereichen weder als „Feigenblatt" noch als Substitut für zwingende Vorgaben eingesetzt werden. In Kombination mit Letzteren kann sie indes dazu beitragen, die Qualität bzw. Vertrauenswürdigkeit von Anwendungen und Anbietern zu erhöhen, und Datengebern wichtige Entscheidungshilfen geben. Bei-

[249] Vgl. allgemein Lewinski 2016, 414 f., 418; speziell mit Blick auf den Datenaustausch in der medizinischen Forschung vgl. Molnár-Gábor/Kaffenberger 2017.
[250] Vgl. als ein entsprechender Verhaltenskodex für die Wissenschaft Zook et al. 2017.
[251] Bundesverband Digitale Wirtschaft 2017, 15.
[252] Vgl. etwa Zook et al. 2017.
[253] Vgl. Rat für Informationsinfrastrukturen 2017.
[254] Vgl. insbesondere die Beiträge in Wright/De Hert 2016; siehe auch Martini 2016, 354.
[255] Vgl. etwa Gellman/Dixon 2016.

spielhaft zu nennen sind etwa überobligatorische Erklärungen dazu, wie eine App oder ein Forschungsprogramm funktioniert und was sie/es leisten kann und erreichen will, wie mit den Daten gearbeitet wird, wie ihre Auswertung funktioniert, ob und wie Daten weitergegeben werden, inwieweit der Nutzer diese selbst wieder aus dem System holen und mit anderen teilen kann, modular über die Weiterverwendung entscheiden kann usw.

Hinzuweisen ist an dieser Stelle auch darauf, dass die hoheitliche und die private Regulierung nicht notwendig strikt voneinander getrennt zu denken sind, sondern miteinander verschränkt werden können. So erlaubt etwa die DSGVO in Art. 42 die Zertifizierung von konkreten Verarbeitungsvorgängen. Darüber hinaus hat die Europäische Kommission schon früher auf der Grundlage von Art. 27 der Datenschutzrichtlinie[256] darauf gedrängt, Selbstregulierung als Instrument des Datenschutzes in Gestalt von Verhaltenskodizes zu fördern. Es verwundert deshalb auch wenig, dass entscheidende Impulse für eine entsprechende Neuordnung von der Unionsebene ausgehen und gerade für M-Health-Anwendungen heute ein relativ komplexes selbstregulatives Gefüge existiert, in dem verbindliche Vorgaben der Datenschutz-Grundverordnung mit einem Verhaltenskodex verknüpft werden und Letzterer mit speziellen Institutionen und Sanktionsmechanismen abgesichert wird.

Die Initiative zu einer entsprechenden Koregulierungsmaßnahme lässt sich zurückführen auf das Konsultationsverfahren zum Grünbuch über Mobile-Health-Dienste[257] und die dabei zutage getretenen Zweifel hinsichtlich der Gewährleistung eines hinreichenden Datenschutzstandards.[258] Der Entwurf eines „Code of Conduct on privacy for mobile health applications“[259] wurde auf Grundlage eines umfangreichen, von der Europäischen Kommission moderierten und eine Vielzahl von unterschiedlichen Stakeholdern einbindenden Normsetzungsprozesses erarbeitet und am 7. Juni 2016 von der Kommission entsprechend der in Art. 27 Abs. 3 der Datenschutzrichtlinie vorgesehenen Option der Artikel-29-Datenschutzgruppe zur Stellungnahme zugeleitet. Von seinem Anwendungsbereich her umfasst der Kodex M-Health-Anwendungen, in denen individuelle Gesundheitsdaten verarbeitet und Dritten zu medizinischen Zwecken mitgeteilt werden. Während demnach einerseits Apps nicht geregelt sind, soweit sie bloße Lifestyle-Daten ohne „klare und enge“ Verbindung zum gesundheitlichen Zustand des Betroffenen verwenden, bleiben andererseits auch solche Anwendungen unberücksichtigt, die bereits als Medizinprodukte zertifiziert sind. Inhaltlich zielt der Kodex, seiner ursprünglichen Entstehungsgeschichte entsprechend, vor allem darauf ab, grundlegende datenschutzrechtliche

[256] Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr (ABl. EG 1995 Nr. L 281/31).
[257] Siehe Europäische Kommission 2014.
[258] Vgl. hierzu und zum Folgenden auch Gassner 2016, 114.
[259] Siehe Europäische Kommission 2016.

Vorgaben und Prinzipien, namentlich die Grundsätze der informierten Einwilligung, der Zweckbindung und der Datenminimierung, für die neuen Anwendungsfelder, insbesondere auch aus Sicht der App-Entwickler, operabel auszugestalten. Besondere Aufmerksamkeit erhalten dabei technische, der eigentlichen Datennutzung vorgelagerte, nämlich schon in der Entwicklungsphase zu berücksichtigende Aspekte von *privacy by design* und *privacy by default*.[260] Erwähnenswert ist ferner, dass der Kodex mit einem umfangreichen institutionellen Arrangement ausgestattet ist, insbesondere umfangreiche Regelungen zur Durchsetzung der enthaltenen Vorgaben existieren.

3.4 Fazit: Statik und Dynamik des Rechtsrahmens

Das Austarieren von Risiken und Chancen der neuen Technologien ist mit Recht als die vermutlich größte gesellschaftspolitische Herausforderung unserer Zeit bezeichnet worden.[261] Aus der Schwierigkeit der Aufgabe darf aber nicht auf ihre Unlösbarkeit geschlossen werden. Vielmehr gilt es, auf Basis der Einsicht in die rechtlichen Steuerungsdefizite nunmehr grundlegende ethische Vorgaben herauszuarbeiten, die auch unter Big-Data-Bedingungen unverzichtbar und zwingend einzuhalten sind. Ziel muss es sein, einerseits die großen Möglichkeiten, die mit Big Data nicht nur, aber gerade im Gesundheitssektor verbunden sind, nutzen zu können, andererseits aber dabei nicht nur nicht hinter das vorhandene Datenschutzniveau zurückzufallen. Im Gegenteil geht es darum, ein gegenüber den besonderen Herausforderungen der Digitalisierung angemessenes Leitprinzip und ein hieran ausgerichtetes Gestaltungs- und Regelungskonzept zu entwickeln.[262] Bevor indes diese hier zunächst nur angedeuteten Überlegungen zur Datensouveränität näher aus- und fortgeführt werden können (siehe Kapitel 5), bedarf es einer diese zusätzlich fundierenden detaillierten Untersuchung der ethischen Grundlagen (siehe Kapitel 4).

[260] Vgl. hierzu auch Becker 2017, 175 ff. mit einem Plädoyer für ein „Recht auf datenerhebungsfreie Produkte".
[261] Vgl. Polonetsky/Tene 2013, 26; ähnlich Roßnagel/Nebel 2015, 459.
[262] Vgl. ähnlich schon Krüger 2016, 190: „Notwendig sind gesetzliche Rahmenbedingungen, die einen angemessenen Ausgleich zwischen der Datensouveränität des Einzelnen und den legitimen wirtschaftlichen Interessen an der Nutzung personenbezogener Daten gewährleisten."

4 Zur Ethik von Big Data und Gesundheit

Angesichts der in Kapitel 2 beschriebenen Entwicklungen von Big-Data-Anwendungen im gesundheitsrelevanten Bereich sind gewichtige Änderungen im individuellen, wissenschaftlichen und gesellschaftlichen Umgang mit Gesundheit und Krankheit teils bereits erfahrbar, teils unschwer prognostizierbar. Zugleich wurde in Kapitel 3 deutlich, dass die bisherigen rechtlichen Regelungsregimes offensichtliche Unzulänglichkeiten aufweisen. Im Folgenden soll dargestellt werden, wie auch grundlegende ethische Normen und Konzepte von den hier beschriebenen technischen Anwendungen in der biomedizinischen Forschung und Praxis und in weiteren gesundheitsrelevanten Bereichen herausgefordert und verändert werden.

Von diesen erwartbaren Dynamiken sind zum einen ethische Orientierungsmuster betroffen, die normativ und evaluativ die Rolle, Funktion und Stellung des Individuums thematisieren, das Big-Data-Anwendungen nutzt. Zu den in dieser Hinsicht relevanten Begriffen gehören Freiheit und Selbstbestimmung, aber auch Privatheit und Intimität, Souveränität und Macht sowie Schadensvermeidung und Wohltätigkeit, die im Kontext intensiver Datensammlung und -verwertung eine Rolle bei der Gestaltung normativer Schutzkonzepte spielen. Sie alle bringen den Wunsch und den Anspruch des Individuums zum Ausdruck, nicht einfach zum Objekt von Datenströmen und den auf diese angewandten Algorithmen zu werden, sondern ein hinreichendes Maß an Kontrolle, Souveränität und Macht über die eigenen Daten zu behalten oder sich zumindest auf die Wahrung ihrer Interessen durch Dritte verlassen zu können.

Zum anderen sind von Big-Data-Anwendungen im Gesundheitsbereich Maßgaben sozialer Orientierung wie Gerechtigkeit und Solidarität betroffen. Sie machen deutlich, dass Menschen Ansprüche auf etwas haben und in Gemeinschaften wechselseitig Sorge füreinander tragen. Angesichts der Auswirkungen, die die tatsächlich oder vermeintlich zunehmende digitale Beobachtungsschärfe aller Lebensbereiche[263] auf das Selbstverständnis von Menschen hat, könnten sich solche Bindungen aber lockern oder gar auflösen – kurzum: deutlich verändern. So könnte etwa der Solidaritätsgrundsatz der gesetzlichen Krankenversicherung in Deutschland infrage gestellt werden, von möglichen Konsequenzen für den Bereich der privaten Krankenversicherungen ganz zu schweigen. Umgekehrt sind aber auch neue normative Solidaritätsmuster denkbar – beispielsweise mit Blick auf die Bereitschaft, eigene Proben oder Daten zur Verfügung zu stellen, um den medizinischen Fortschritt jenseits eigenen Nutzens zu fördern.

Vor diesem Hintergrund wird es schließlich auch schwieriger, Personen oder Institutionen unter Big-Data-Bedingungen moralische Verantwortung zuzuschreiben: Vor allem ist zu fragen, wer nicht nur rechtlich, sondern auch moralisch zur Rechenschaft gezogen werden kann, wenn

[263] Vgl. Kucklick 2016.

Algorithmen und die aus ihrer Mustererkennung gezogenen Konsequenzen nicht nachvollziehbar sind, wenn Maschinen „lernen“, auf Grundlage großer Datenmengen über Menschen zu urteilen, und Personen auf der Basis der so entwickelten Bewertungsmuster Entscheidungen treffen.

Im Folgenden sollen nicht nur ein grundlegendes Verständnis dieser Konzepte und ihrer möglichen Transformationen durch Big-Data-Anwendungen in Medizin und Forschung sowie im weiteren gesundheitsrelevanten Bereich skizziert werden. Es sollen in diesem Zusammenhang auch die mit solchen Veränderungen verbundene Vulnerabilität betroffener Individuen und Gruppen in den Blick genommen werden. Vor dem Hintergrund dieser Analysen werden in den Kapiteln 5 und 6 Grundzüge eines angemessenen Gestaltungs- und Regelungskonzepts entwickelt, das das Grundanliegen der genannten normativen Orientierungsmuster aufgreift und für das digitale Zeitalter in theoretischer wie praktischer Hinsicht umsetzen soll.

4.1 Freiheit: Handlungsurheberschaft und Selbstbestimmung

Das Bundesverfassungsgericht hat in seinem Volkszählungsurteil die informationelle Selbstbestimmung als wesentlichen Teilaspekt des allgemeinen Persönlichkeitsrechts herausgestellt.[264] Nach den Ausführungen von Kapitel 3 erscheint das zu ihrem Schutz entwickelte Datenschutzrecht aufgrund der unter Big Data zusammengefassten informationstechnischen Entwicklungen, die in Kapitel 2 dargestellt wurden, als in wichtigen Zügen reformbedürftig. Es kann informationelle Selbstbestimmung nicht mehr hinreichend gewährleisten. Vor diesem Hintergrund ergibt sich die Aufgabe, den normativen Kerngehalt dieses Schutzgutes neu bzw. genauer zu bestimmen.

4.1.1 Handlungsurheberschaft

Der Ausdruck Freiheit wird in vielen Bedeutungen, in mannigfachen Zusammenhängen und oft wenig präzise verwendet. Die im Folgenden entwickelten Unterscheidungen dienen vor allem einer begrifflichen Rekonstruktion menschlichen Handelns, die sich weitgehend auf die lebensweltliche Handlungserfahrung stützen. Eine vermutlich von allen geteilte grundsätzliche menschliche Handlungserfahrung besteht darin, dass Akteure sich selbst und andere bestimmten Ereignissen und Zuständen als Urheber zurechnen. Von dieser grundsätzlichen Annahme sind die konkreten Bedingungen zu unterscheiden, die mit der Ausübung der Handlungsurheberschaft verbunden sind und diese in wechselndem Maße beeinflussen. Diese konkret-personalen Handlungsumstände sind hier unter dem Begriff der Selbstbestimmung zusammengefasst. Die Unterscheidung von Handlungsurheberschaft (als grundsätzlicher Freiheitsbedin-

[264] Vgl. BVerfGE 65, 1.

gung) und Selbstbestimmung (als abhängig von mehr oder weniger deutlich erfahrbaren Umständen) entspricht einer verbreiteten Charakterisierung der Selbstbestimmung als des Praktisch-Werdens von Freiheit in einer einzelnen Person.[265] Mit dieser Erklärung des Begriffs der Freiheit wird ein breites Spektrum unterschiedlicher Formen normalen Handelns erfasst, vom reflektiert vollzogenen Beschluss über zahlreiche Verhaltensweisen des Alltags (wie zum Beispiel das unbewusste Lächeln beim Grüßen) bis hin zu absichtslos expressiven Akten (wie zum Beispiel dem Jubelschrei der Fans beim Fußball). Sie alle sind ohne Weiteres Urhebern zurechenbar und damit im hier gemeinten Sinne grundsätzlich frei.[266]

Die Fähigkeit zur Handlungsurheberschaft ist Grundlage dafür, dass handelnde Menschen ihre Handlungen nach Maximen ausrichten können, die sie sich selber setzen.[267] Verfügten sie über diese Fähigkeit nicht und wären sie gezwungen, Handlungen allein durch Befolgen von Autoritäten, Traditionen oder äußeren Zwang auszuführen oder wären Handlungen keiner anderen vernunftgemäßen Beschreibung zugänglich als in Begriffen kausaldeterministischer Vorgänge, seien sie naturhafter (physikalischer, genetischer, neuronaler) oder psychologisch-sozialer (familienspezifischer, schichtenspezifischer, wirtschaftlicher) Art, dann verlöre die Rede von der Freiheit den Kern ihres genuinen Sinns. Denn diese naturhaften oder sozialen Bestimmungen schließen nicht aus, sondern setzen in gewissem Sinn sogar voraus, dass der Mensch sein eigenes Dasein in ein Verhältnis zu solchen Bestimmungen (durch unter anderem Überwindung und Unterwerfung) setzen kann. Gewiss tut er dies stets auch in seiner Zugehörigkeit zur äußeren, beispielsweise von Biologie und Physik beschriebenen und erklärten Welt, also auch auf

---

[265] Eine eingehende Untersuchung dazu findet sich bei Gerhardt 1999, 107-147.

[266] Der Begriff der Handlungsurheberschaft ist allerdings in der neueren philosophischen Diskussion Gegenstand zahlreicher Kontroversen. Umstritten ist etwa, ob er allein ein jeweils objektiv feststellbares Geschehen (das In-Szene-Setzen eines Tuns durch einen Akteur) bezeichnet oder, ob er auch – und gegebenenfalls in welchem Modus und Umfang – ein subjektives Handlungserleben dieses Akteurs voraussetzt. Beides muss nicht konvergent nebeneinander gegeben sein. (In gewissen Grenzfällen, zum Beispiel solchen der Schizophrenie, empfinden Akteure ihr eigenes Tun als von Dritten gesteuert; in anderen Fällen empfinden sie ihr Tun, das allein von Dritten, etwa durch bestimmte Hirnstimulationen, ausgelöst wird, als genuin eigenes, selbstgesteuertes Handeln.) Umstritten sind darüber hinaus zahlreiche Einzelfragen, und zwar sowohl (1.) im Hinblick auf die objektiven Grundlagen einer Handlungszuschreibung als auch (2.) mit Blick auf die Phänomenologie des subjektiven Handlungserlebens und schließlich (3.) auf die gegebenenfalls für notwendig gehaltene Beziehung zwischen diesen beiden Typen von Bedingungen der Handlungsurheberschaft. Siehe dazu Roessler/Eilan 2003 sowie Hyman/Steward 2004. Für unsere Zwecke bedürfen diese Fragen jedoch keiner Erörterung. Wir konzentrieren uns vielmehr auf die „Normalfälle" menschlichen Handelns, die bei aller Vielfalt ihrer Formen regelmäßig *beide* konstitutiven Elemente einer Urheberschaft aufweisen, die objektiven wie die korrespondierenden subjektiven – in welchem Umfang, Maß und Verhältnis auch immer.

[267] Maximen im hier verwendeten Sinne sind Maßgaben mittlerer Allgemeinheit und Reichweite als subjektive Prinzipien eigenen Handelns; sie umfassen mehr als jeweils einzelne Handlungsmotive und weniger als die gesamte eigene Lebensführung. (Beispiele: Bedürftigen grundsätzlich zu helfen; Rechtsverletzungen regelmäßig abzuwehren; eigene Fähigkeiten nach Möglichkeit zu entwickeln etc.). Philosophiegeschichtlich ist der Begriff der Maxime (als Bezugsgegenstand des kategorischen Imperativs) vor allem von Immanuel Kant entwickelt worden. Siehe Grundlegung zur Metaphysik der Sitten (1785), AA IV, 385-464 (Kant 1903, 420 f.). Nach Kant setzt sich der Akteur die Maximen kraft „reiner praktischer Vernunft" selbst und handelt deshalb (und nur insofern) „autonom" (selbst gesetzgebend). Siehe ebd., 402, 431 und vgl. dazu auch Gerhardt 1999, 406-413.

der Grundlage von Vorgängen, die (einschließlich der in seinem Gehirn stattfindenden) naturhaften Regelmäßigkeiten unterliegen. Das schließt einen vernünftig verstandenen Begriff personaler Freiheit qua Handlungsurheberschaft aber ebenso wenig aus wie eine hinreichende Selbstbestimmung in konkreten Fällen einzelnen Handelns.[268]

Das Merkmal der Handlungsurheberschaft kommt einem Wesen entweder zu oder nicht (*kontradiktorischer Gegensatz*).[269] Wird mit Immanuel Kant unterstellt, dass der Akteur sich selbst als Letztzweck seiner Handlungen setzt und dafür Anerkennung fordert, die er umgekehrt auch anderen Akteuren zuzugestehen hat, und wird somit gefordert, dass ein Akteur niemals als bloßes Mittel für Zwecke anderer zu behandeln ist (Instrumentalisierungsverbot), dann kann man von der Anerkennung der Würde solcher Akteure sprechen.[270] Unbeschadet der Frage, ob auch nicht menschliche Wesen (beispielsweise Tiere oder „superintelligente" Maschinen) als Handlungsurheber in Betracht kommen, beschränken sich die weiteren ethischen Überlegungen auf Menschen.

4.1.2 Selbstbestimmung und Einwilligung

Freiheit im Sinne der konkreten Verwirklichung des Merkmals der Handlungsurheberschaft soll nach der hier oben genannten vorgeschlagenen Unterscheidung *Selbstbestimmung* heißen. Selbstbestimmung können Handlungsurheber mehr oder weniger oder auch gar nicht ausüben (*polar-konträrer Gegensatz*)[271]. In diesem Sinne wird Selbstbestimmung in unterschiedlichen Erscheinungsformen verwirklicht.[272]

Einmal bezeichnet Selbstbestimmung die *Fähigkeit* einer Person, ihr Leben jedenfalls im Großen und Ganzen nach ihren eigenen Vorstellungen zu gestalten. Eine solche tatsächliche Fähigkeit kann in unterschiedlichen Graden gegeben sein. Unterhalb einer bestimmten Minimalgrenze des dafür Erforderlichen wird man von Selbstbestimmung auch im Sinn einer bloßen Disposition nicht mehr sprechen können, etwa bei kleinen Kindern oder mental schwer Erkrankten. Wo diese Grenze zu ziehen ist, kann jedoch nicht abstrakt-generell, sondern muss in Abhängigkeit vom jeweiligen Kontext des Handelns und den damit jeweils verfolgten Zwecken bestimmt werden.

---

[268] Das philosophische Problem der Willensfreiheit kann hier nicht einmal andeutungsweise behandelt werden (Überblick in Kane 2011). Über allen Streit der zahlreichen Positionen hinweg gibt es einen weitreichenden Konsens darüber, dass ein vernünftiger Begriff menschlicher Freiheit für alle Seiten (wenngleich mit unterschiedlichen Anforderungen an diesen Begriff) explizierbar ist.

[269] Unter einem kontradiktorischen Gegensatz versteht man einen solchen, der im Interesse der Vermeidung eines Widerspruchs nur Ja-Nein-Behauptungen zulässt (Beispiel: farbig – farblos), während ein polar-konträrer Gegensatz diskrete oder kontinuierliche Übergänge zulässt (Beispiel: Schwarz – Weiß).

[270] Siehe etwa Grundlegung zur Metaphysik der Sitten (1785), AA IV, 385-464 (Kant 1903, insbesondere 2. Abschnitt, 427-429, 437) zum Instrumentalisierungsverbot insbesondere ebd., 438.

[271] Siehe Fn. 269.

[272] Die folgenden Ausführungen lehnen sich an die Untersuchungen von Feinberg 1989, 27-51.

Ferner bezeichnet Selbstbestimmung einen *tatsächlich gegebenen Zustand*: Auch wer die Fähigkeit zur Selbstbestimmung besitzt, muss diese in seinem persönlichen Leben nicht oder nicht durchgängig ausüben, sei es, weil äußere Gründe seiner Lebensumstände dies verhindern (zum Beispiel in einer Gefangenschaft oder in Umständen extremer Not), sei es, weil er die vorhandene eigene Fähigkeit verkümmern lässt und seine Lebensführung weitgehend den Direktiven anderer anheimgibt. Dann fehlt es am tatsächlichen Zustand ausgeübter Selbstbestimmung. Eine selbstbestimmte praktische Lebensführung im Sinne dieser Bedeutung muss eine Reihe von Elementen aufweisen, die als Kriterien einer solchen Lebensführung verstanden werden können, etwa die hinreichende Unabhängigkeit von den Entscheidungen anderer, eine hinreichende Authentizität der eigenen Entscheidungen sowie die hinreichende Selbstkontrolle der Handlungen, mit denen solche Entscheidungen ausgeführt werden.

Schließlich bezeichnet Selbstbestimmung eine als ideal vorgestellte Form der *Lebensführung*. Auch wer im Sinne der vorher erwähnten Bedeutungen fraglos als (hinreichend) selbstbestimmt zu gelten hat, mag dennoch mehr oder weniger weit entfernt sein von einem Idealbegriff der selbstbestimmten Person (gegebenenfalls auch seinem höchstpersönlichen) und von der entsprechenden Lebensführung. Akzeptiert er diesen Idealbegriff als für sich verbindlich, mag er ihm die Aufgabe entnehmen, die eigene Persönlichkeit in Richtung dieses Zieles weiterzuentwickeln. Als Ideal muss das Ziel nicht erreichbar sein. Vielmehr sollte es als regulative Idee zur Orientierung der eigenen Lebensführung aufgefasst werden.

Von diesen Formen personaler Selbstbestimmung ist der *rechtliche* Schutz ihrer Ausübung zu unterscheiden. Verfassungs- wie einfachgesetzliche Normen garantieren in vielerlei Hinsicht ein Handeln-Können nach höchstpersönlichen Maximen, Gründen, Wünschen und Interessen. Der Umfang ihres rechtlichen Schutzes ist nicht deckungsgleich mit der sachlichen Reichweite jener Formen höchstpersönlicher Selbstbestimmung (als Fähigkeit, Zustand, Ideal). Er reicht vielmehr regelmäßig darüber hinaus. Die meisten Menschen treffen zumindest manche ihrer Entscheidungen nach den Maßgaben anderer Personen, nach öffentlichen Moden oder kollektiven Stimmungen, und manche treffen die meisten ihrer Entscheidungen so. Den oben dargelegten Kriterien personaler Selbstbestimmung mag ein solcher Modus des Unselbstständigen nicht annähernd genügen. Gleichwohl fällt auch er – bis zur Grenze des Genötigt- oder Manifest-Getäuscht-Werdens – als rechtlich selbstbestimmt ohne Weiteres in den Schutzbereich der einschlägigen Normen. Handlungen, die in dieser Weise extern motiviert oder bestimmt werden, mangelt es deshalb keineswegs an rechtlicher Verbindlichkeit. Daher kann, wer derart unselbstständig, aber in rechtlicher Hinsicht gleichwohl selbstbestimmt handelt, sich jede (paternalistische) Einmischung Dritter verbitten.

Die skizzierte Differenz zwischen den Reichweiten einerseits höchstpersönlicher und andererseits rechtlicher Selbstbestimmung hat freilich eine Kehrseite, deren Bedeutung gerade im thematischen Zusammenhang mit Big-Data-Technologien auf der Hand liegt. Soweit externe Einflussnahmen auf die Entscheidungen anderer die Schutzgrenzen deren rechtlicher Selbstbestimmung nicht überschreiten (also etwa mit subtileren Methoden agieren als denen der Täuschung oder Nötigung), sind sie ihrerseits grundsätzlich rechtlich zulässig. Die Fähigkeit ihrer Adressaten zur höchstpersönlichen Selbstbestimmung ihres Handelns, zu einer darauf gegründeten Lebensführung und erst recht die Kriterien eines dafür richtungsweisenden Ideals mögen sie gleichwohl substanziell gefährden. Das gilt insbesondere für Strategien einer solchen Einflussnahme, deren Wirkungen unterhalb der Schwelle der Wahrnehmung ihrer Adressaten liegen. Gerade dieser Modus ist es aber, der Big-Data-basierte Formen des Einflussnehmens spezifisch kennzeichnet. Als massenhafte Nudges[273] in ökonomischen, politischen, religiösen oder weltanschaulichen Dingen mögen sie eine von den Nutzern einschlägiger Internetdienste nicht bemerkte und eben deshalb nachdrückliche Wirksamkeit entfalten, die zwar die Schwelle zur kollektiven Manipulation überschreitet, aber noch keine gesetzlichen Verbote verletzt. Damit stellt sich die Frage, ob es gegen solche neuen und subtilen Formen der Unterminierung höchstpersönlicher Selbstbestimmung auch neuer Formen rechtlichen Schutzes bedarf (siehe Abschnitt 3.1).

Während man auf die grundsätzliche Entscheidungs- und Handlungsfähigkeit nicht verzichten und sie nicht delegieren kann, verhält es sich mit der Selbstbestimmung anders. Formen und Grade ihrer Ausübung sind von erheblicher praktischer Bedeutung. Man kann etwa in bestimmten Zusammenhängen und für präzise bezeichnete Zwecke sein Recht auf Selbstbestimmung an jemand anderen vorsorglich delegieren. Diese Delegationsmöglichkeit wird besonders relevant bei nach Leistung und/oder Fähigkeit in ihrer Selbstbestimmung eingeschränkten Personen, etwa bei Personen, die Vorsorge für eine Demenz treffen, wenn ihnen im oben dargelegten Sinn auch eine Selbstbestimmung als Recht nicht zugebilligt werden kann. Ausfallende Fähigkeiten sind dann, soweit nötig und möglich, zu kompensieren, etwa durch einen von der Person individuell eingesetzten Vertreter (mandatierte Selbstbestimmung), durch einen von der Person mit anderen kollektiv gewählten Vertreter (repräsentative Selbstbestimmung) oder durch einen von anderen für die Person eingesetzten Vertreter (advokatorische oder assistierte Selbstbestimmung). Ein in dieser Form von Selbstbestimmung legitimierter „weicher" Paternalismus ist als Fürsorge nicht immer schon ein Gegensatz zur Selbstbestimmung – jedenfalls

[273] Der von Richard Thaler und Cass Sunstein eingeführte Begriff „Nudge" (engl. für Schubser) steht als Sammelbegriff für Maßnahmen der Umgestaltung der sogenannten Entscheidungsarchitektur (Thaler/Sunstein 2008). Letztere wird so verändert, dass ein bestimmtes erwünschtes Verhalten einfacher, attraktiver, offensichtlicher etc. wird, ohne dass gleichzeitig Handlungsoptionen verwehrt werden. (Siehe Abschnitt 4.3).

dann nicht, wenn er das Wohl des anderen aus dessen eigenem Verständnis heraus zum Maßstab nimmt. Paternalismus bei Kindern kann wiederum gerechtfertigt sein, wenn die stellvertretend entscheidende Person die Interessen des Kindes im Auge hat und durch ihre fürsorglichen Entscheidungen die Würde des Kindes nicht untergräbt.

Im Kontext von Big Data ist in den vergangenen Jahren eine Reihe von Instrumenten entwickelt worden, um die praktische Umsetzung der oben beschriebenen Selbstbestimmungsverständnisse zu ermöglichen – unter anderem durch Delegation. Dazu gehören beispielsweise verschiedene Formen der Einwilligung von Individuen in die Sammlung, Speicherung und Weiterverwertung eigenen Biomaterials und der damit zu verknüpfenden persönlichen Daten, wie dies beispielsweise in großen Biobanken[274] in der medizinischen Forschung erfolgt. Anders als bei der klinischen medizinischen Forschung ist zum Zeitpunkt der Speicherung von Material und Daten in Biobanken meist nicht absehbar, welche Forschungsfragen künftig bearbeitet werden sollen. Hinzu kommt die für viele Big-Data-Fragestellungen grundlegende, oft konkret nicht absehbare Weitergabe der Materialien und Daten an andere Forscher und Institutionen, sowie die Verknüpfung mit weiteren Datensätzen, die für den optimalen Forschungsertrag von Biobanken wichtig sind. All dies führt dazu, dass die klassische, eng zweckgebundene informierte Einwilligung aus der klinischen Forschung, in der die Forschungsteilnehmer detailliert über Forschungsziele, -fragen, -risiken und die Weiterverwendung von Ergebnissen aufgeklärt werden, im Kontext von Biobanken nicht zielführend ist.[275] Die Erlaubnis der Material- und Datennutzung in diesem Kontext muss unbestimmter und damit breiter sein, um Praktikabilität und Effektivität zu gewährleisten. Zugleich sollte sichergestellt werden, dass sie auf der Grundlage einer wirklich selbstbestimmten Entscheidung der Teilnehmer erfolgt.

Eine ursprünglich vor allem im angloamerikanischen Raum verbreitete Möglichkeit, auf diese Herausforderung zu reagieren, ist eine Blanko-Einwilligung der Teilnehmer. Dabei willigen die Teilnehmer einmal, nämlich zum Zeitpunkt der Materialentnahme und Datenspeicherung, in die (inhaltlich unbestimmte) zukünftige Nutzung oder Weitergabe ihrer Materialien und Daten ein; alle weiteren Entscheidungen hinsichtlich der Nutzung liegen dann bei den Betreibern der

---

[274] In der biomedizinischen Forschung waren bisher zwei Biobanktypen am weitesten verbreitet: erstens meist kleinere, krankheitsspezifische Biobanken, in denen Patientendaten und Biomaterialien zur Erforschung einer bestimmten Erkrankung oder Erkrankungsgruppe gesammelt werden, etwa um krankheitsbezogene Genomforschung zu ermöglichen, wobei eine genaue Zielbestimmung auch bei diesen Biobanken bereits schwerfällt. Der zweite Typ sind populationsbasierte Biobanken, in denen Daten und Materialien von großen Populationen bzw. Bevölkerungskohorten gesammelt werden, um daran gesundheitsbezogene Forschung zu betreiben, ohne enge, etwa krankheitsbezogene, Zweckbindung. (Oft erfolgt allerdings für jede Einzelstudie innerhalb der Biobank dann eine spezifische, zweckbezogene Einwilligung). Zunehmend verwischen die Grenzen zwischen den verschiedenen Typen. So werden etwa, wie in Kapitel 2 beschrieben, verschiedene vormals krankheitsspezifische Biobanken zu größeren Forschungsrepositorien zusammengezogen oder große Einrichtungen des Gesundheitswesens, etwa Unikliniken, sammeln alle Restmaterialien ihrer Patienten und machen diese verknüpfbar mit klinischen Daten (*healthcare-embedded biobanking*).

[275] Siehe hierzu Deutscher Ethikrat 2010.

Biobank. Ob eine solche unbestimmte, einmalige Einwilligung und de facto Entäußerung aller zukünftigen Entscheidungs- und Kontrollrechte ohne weitere Sicherungselemente den Anforderungen an eine selbstbestimmte Entscheidung entsprechen kann, wurde und wird sehr kontrovers diskutiert.[276] Inzwischen haben sich daher verschiedene andere Einwilligungsmodelle etabliert. Sie sollen eine Balance garantieren zwischen der unrealistisch engen Zweckbestimmung auf der einen und einer einmaligen, allzu breiten Freigabe auf der anderen Seite.[277]

Bekannt geworden ist das Modell der dynamischen Einwilligung.[278] Es soll den Voraussetzungen einer engen, informierten Zustimmung dadurch gerecht werden, dass in einem dynamischen Modus mehrfacher Wiederholungen in jeweils einzelne Elemente und Teilprojekte der Biobank eingewilligt werden kann. Die Teilnehmer stehen mit der Biobank in einer wechselseitigen Beziehung – zumeist über eine Online-Plattform oder telefonisch – und erhalten regelmäßig eine Art Optionen-Menü, anhand dessen sie entscheiden können, wie viele Informationen sie über weitere Projekte erhalten wollen und ob und an welchen Projekten sie mit der Material- und Datenfreigabe teilnehmen möchten.

Im ursprünglichen Modell der dynamischen Einwilligung war nur vorgesehen, die persönliche, zweckgebundene Einwilligung in zukünftige Projekte sicherzustellen, zu erleichtern und den Informationsprozess zu optimieren. Tatsächlich stellte sich heraus, dass Teilnehmer es mitunter bevorzugen, zumindest einige dieser zukünftigen Entscheidungen zu delegieren, etwa in Form der oben skizzierten repräsentativen Selbstbestimmung. Das Modell der dynamischen Einwilligung ist deswegen weiterentwickelt und erweitert worden. Gegenwärtig darf wohl das als Kaskaden- oder Meta-Einwilligung bezeichnete Verfahren als Gold-Standard gelten.[279] Mit ihm wird zum einen die Dynamisierung der Einwilligung beibehalten und um weitere Optionen, etwa um Delegationsmöglichkeiten, ergänzt. Zum anderen wird dem Prozess eine wichtige Entscheidung vorgelagert. Teilnehmer entscheiden zu Beginn, also anlässlich der Material- und Datenspende, welche Form der Einwilligung sie grundsätzlich bevorzugen. Diese Entscheidung beruht auf der sorgfältigen Information über die übergreifenden Ziele der Biobank und deren wichtigste Charakteristika, also die Art der Finanzierung, das jeweilige Datenschutz- und Governance-Konzept, etwaige etablierte Kooperationen und Weitergabevereinbarungen mit anderen Institutionen, sowie die verschiedenen Möglichkeiten der Einwilligung in die Teilnahme. Teilnehmer werden dann gebeten auszuwählen, ob sie in Zukunft in dynamischer Form über jedes weitere Projekt informiert und jeweils um eine erneute Teilnahme gebeten werden möchten. Sie können alternativ in verschiedene, breite Kategorien von Forschung einwilligen,

[276] Einschlägig ist hier etwa Caulfield 2007.
[277] Eine Übersicht der verschiedenen Modelle und der ethischen Argumente in dieser Debatte bietet Richter/Buyx 2016.
[278] Siehe Kaye et al. 2015.
[279] Siehe Ploug/Holm 2016.

ohne für einzelne Projekte jedes Mal die Erlaubnis zu geben; sie können die zukünftigen Entscheidungen stellvertretend an ein Expertengremium, etwa eine mit der Biobank assoziierte Ethikkommission oder ein unabhängiges Kontrollorgan delegieren; sie können in manchen Varianten des Konzeptes ihre Materialien und Daten für jegliche Nutzung gänzlich freigeben; und sie können natürlich auch eine Teilnahme ablehnen. Der Wechsel zwischen den Einwilligungsformen bleibt im Kaskadenmodell möglich und kann über die im dynamischen Modell etablierten Kommunikationsstrukturen erfolgen, ebenso wie der spätere Widerspruch für eine weitere Nutzung. Es sind zudem Kombinationen der verschiedenen Einwilligungsoptionen denkbar – etwa die Delegation verbunden mit einer grundsätzlich breiten oder umfassenden Freigabe.

Dieses Kaskadenmodell jenseits der bisherigen Pilotprojekte breiter für den Bereich der biomedizinischen Forschung – und darüber hinaus – zu implementieren, ist mit signifikanten organisatorischen Anstrengungen für die betroffenen Institutionen verbunden. Andererseits kann davon ausgegangen werden, dass dieses Modell dem Bemühen, die Nutzung der eigenen Daten durch andere entlang des eigenen Wohls und der persönlichen Wert- und Lebensvorstellungen auszurichten, am ehesten entsprechen könnte – vorausgesetzt, es ist sichergestellt, dass flankierende Schutzmechanismen, wie etwa entsprechende Datenschutzkonzepte, nachhaltig wirksam sind.

4.1.3 Äußere Rahmenbedingungen für die Realisierung von Freiheit

Wenn unter Selbstbestimmung das konkret-individuelle Praktisch-Werden der Freiheit zu verstehen ist, dann sind die Merkmale, mit denen etwa John Stuart Mill Freiheit charakterisiert, nämlich Individualität, Authentizität und Originalität – nicht anders als die oben skizzierten Kriterien – Formen der Selbstbestimmung.[280] Für deren jeweilige Beurteilung ist auch der soziale Kontext des Handelnden mit einzubeziehen. Freiheit als Fähigkeit kann man nur dann selbstbestimmt verwirklichen, wenn auch die äußeren Realisierungsbedingungen in hinreichendem Maße gegeben sind. Diese können die Selbstbestimmung positiv betreffen, insofern ein Einzelner in der Lage ist, eigene Lebenspläne zu entwickeln und umzusetzen. Eine plurale Gesellschaft setzt voraus, dass es für jeden Einzelnen hinreichend weite Handlungsspielräume gibt, den eigenen Lebensplänen und Intentionen (sofern sie andere nicht illegitim beeinträchtigen) störungsfrei nachgehen zu können, ohne sich dafür rechtfertigen zu müssen. Diese Selbstbestimmung als Ausdruck des Rechts, alleingelassen zu werden (oft negative Freiheit genannt)[281], ist verknüpft mit dem Anspruch auf geschützte Privatheit und Intimität (siehe Abschnitt 4.2), bedeutet aber nicht ein wechselseitiges Desinteresse der Menschen aneinander. Sie ist vielmehr personale Bedingung und Nährboden für die Entwicklung eigener Fähigkeiten zur

[280] Siehe Mill 1859.
[281] Siehe hierzu Berlin 1969.

selbstbestimmten Lebensführung unter den orientierenden Maßgaben eines regulativen Ideals (oft *positive Freiheit* genannt). Umgekehrt besteht die Gefahr, dass sich der Raum negativer Freiheit entleert oder ausdünnt, wenn er nicht darauf ausgerichtet ist, positive Freiheiten als Selbstbestimmung so auszubilden, dass Menschen sich auch für Gemeinschaften, in denen sie leben, und für die dort bedeutsamen Fragen des Gemeinwohls engagieren.

Frei sein und selbstbestimmt handeln zu können, bedeutet vor dem skizzierten Hintergrund zumindest die realistische Möglichkeit, die eigene Identität bewahren und gestalten sowie die eigenen Handlungen vor sich und anderen verantworten zu können. Dazu sind – gerade in einer pluralen und funktional ausdifferenzierten Gesellschaft – rechtsstaatliche Standards notwendig, die verlässlich und fair sind und die ohne Ansehen der Person gelten. Selbstbestimmung, negative und positive, innere und äußere Freiheit realisieren sich auf Dauer nur in einem solchen Rahmen gerechter und rechtssicherer Institutionen. Dies wiederum setzt voraus, dass Bürger ihn ihrerseits als wesentlich für die Möglichkeit ihrer eigenen Selbstbestimmung und damit für die Verwirklichung ihrer Vorstellungen von einem guten individuellen und sozialen Leben verstehen. Positive und negative Freiheit sowie äußere und innere Freiheit stehen insofern in einem Verhältnis wechselseitigen Bedingtseins.

Vor diesem Hintergrund sind Einschränkungen der äußeren Freiheit sehr genau daraufhin zu prüfen, ob sie mit einer bedrohlichen Wirkung auch auf die innere Freiheit einhergehen. In der Sozialethik und der politischen Philosophie wird beispielsweise unterstrichen, dass die Institutionengestaltung – zumindest die der öffentlichen Institutionen und Ämter – nur dann freiheitsfreundlich und -förderlich ist, wenn sie transparent erfolgt und einen grundsätzlichen Zugang für jedermann gewährleistet.[282] Ferner muss sie die von den Institutionsentscheidungen Betroffenen befähigen, diese Entscheidungen möglichst zu verstehen und den Umständen entsprechend souverän umzusetzen. Schließlich muss es Verfahren der Beteiligung und Kontrolle geben, um eine Revision der Institutionenstruktur sowie einzelner Entscheidungen grundsätzlich in die Wege leiten zu können. Bei allen faktischen Grenzen, solche Vorgaben in einer komplexen, zudem gerade im Informationsbereich global vernetzten Gesellschaftsstruktur umzusetzen, lässt sich nicht leugnen, dass solche Kriterien Prüfstandards markieren, wie Individuen ihre innere Freiheit in einer modernen Gesellschaft so bewahren und verteidigen können, dass sie nicht durch äußere Beschränkungen allzu leichtfertig gefährdet wird.

[282] Prominentes Beispiel: John Rawls, für den die Offenheit des Zugangs zu Ämtern und Positionen für jedermann unter Bedingungen der fairen Chancengleichheit als ein konstitutives Element zum zweiten seiner (zwei) grundlegenden Prinzipien der Gerechtigkeit gehört, dem sogenannten Differenzprinzip (mit dem allein sich nach Rawls soziale Ungleichheiten rechtfertigen lassen). Vgl. Rawls 1972, 75 f.

4.2 Privatheit und Intimität

Der Anspruch, über das eigene Leben selbst zu bestimmen, es entsprechend führen zu können und an eigenen Motiven, Gründen, Wünschen und Optionen auszurichten, zu denen der Handelnde nicht gezwungen wird, erstreckt sich grundsätzlich auf alle Bereiche des privaten wie des öffentlichen Lebens. Im öffentlichen Raum ist die selbstbestimmte Lebensführung einer Person in ein engmaschiges Netz der Koordination mit anderen und der Rechtfertigung ihnen gegenüber eingebunden. Die *Öffentlichkeit* ist potenziell immer das Medium ungewollter Beobachtung und Kontrolle; sie wirkt deshalb steuernd oder limitierend auf die selbstbestimmte Lebensführung jedes Einzelnen zurück. Dagegen bezeichnet *Privatheit* eine Lebenssphäre, in der solche ungewollten Kontrollmechanismen und Rechtfertigungsnotwendigkeiten weitgehend[283] zurückgedrängt sind und die betroffene Person selbst darüber entscheidet, wem sie zu diesem Bereich Zugang gewährt oder nicht.

Die klassische rechtliche Definition von Privatheit als Recht, in Ruhe gelassen zu werden („the right to be let alone") stammt aus dem Jahr 1890.[284] Die Unterscheidung zwischen einer privaten und einer öffentlichen Sphäre ist jedoch älter und kulturübergreifend geläufig[285] – auch wenn die zu schützende Sphären, die dafür verwendeten Begründungen und die Art, Verursacher von Bedrohungen zu definieren, im historischen Verlauf variiert haben und ihr Verhältnis zueinander ständig neu austariert worden ist. Schon die antiken Kulturen der Griechen und Römer unterscheiden zwischen der Öffentlichkeit (*polis* bzw. *res publica*) einerseits und der heute als Privatsphäre bezeichneten Hausgemeinschaft (*oikos* bzw. *domus*) andererseits. Ähnliches gilt auch für weite Teile der vom Judentum oder vom Islam geprägten Kulturräume.

Auch in der Moderne finden sich zahlreiche Traditionen, die an diese grundlegende Differenzierung anknüpfen und sie weiterentwickeln. Zwei Entwicklungslinien sind hier besonders relevant: Mit Blick auf die Abgrenzung des *oikos* zur *polis* wird die Privatsphäre zunehmend eigentumslogisch begründet. Zugleich soll sie einen Bereich markieren, der in erster Linie vor staatlichen Eingriffen schützt. So verbindet sich mit der neuzeitlichen Idee der Rechtsstaatlichkeit auch das Bestreben, den Bürgern einen klar umrissenen und verlässlichen Raum offenzuhalten, innerhalb dessen sie ohne Einflussnahme von außen ihre persönlichen Angelegenheiten regeln (Privatautonomie). Freilich wurde schon früh moniert, dass die Privatsphäre nicht nur vonseiten des Staates, sondern auch seitens der Gesellschaft von einem hohen Maß an informellem Druck bedroht ist. Deshalb bezieht sich die klassische Definition von Privatheit als Recht, in Ruhe gelassen zu werden, neben dem Staat auch auf andere gesellschaftliche Kräfte

[283] Auch innerhalb der Privatsphäre/im rechtlich geschützten Raum der eigenen Privatautonomie darf man natürlich nicht *alles* machen; was man will; auch hier gelten die Schutzrechte derer fort, mit denen man seine Privatsphäre teilt.
[284] Vgl. Warren/Brandeis 1890.
[285] Vgl. ebd.

wie etwa die Massenmedien oder Weltanschauungsgemeinschaften, die die Privatsphäre als Ort unbeeinflusst selbst gestalteten Lebens erheblich bedrohen können.

Dieses Verständnis von Privatsphäre als Ort unbeeinflusst und unbeobachtet selbstgestalteten Lebens spielt auch mit Blick auf eine bedeutsame Binnendifferenzierung des *oikos* eine erhebliche Rolle: Denn auch innerhalb des *oikos* oder – modern gewendet – innerhalb einer Familie entstehen abgegrenzte Räume des Privaten selbst zwischen ihren ansonsten eng verbundenen Mitgliedern, die es wechselseitig zu respektieren gilt. Die heimliche Lektüre der Tagebucheintragungen der Lebenspartnerin erscheint vor diesem Hintergrund vielen ebenso Tabu wie die der Facebook-Seite oder der Chatprotokolle im Mobiltelefon eines Jugendlichen, der seinen Eltern zunächst keinen Zugang gestatten will. Es handelt sich um Räume von Intimität, die ausschließlich den unmittelbar Betroffenen vorbehalten bleiben und deren Details nur von ihnen selbst – wenn überhaupt – einem ausgewählten Kreis selbstbestimmt zugänglich gemacht werden können. Und auch das gilt nur in begrenztem Umfang: Es gibt Lebensbereiche der Intimität, deren freiwillige, ja bewusst provozierende Offenlegung anderen oftmals als Exhibitionismus gilt. Zwar ist das, was als intim (oder auch exhibitionistisch) gilt, in erheblichem Umfang kulturvariant. Gleichwohl bezeichnet es immer einen Sachverhalt, der in besonderer Weise mit Scham oder (bloßstellender) Beschämung verbunden ist. Genau der Schamaspekt, der Privatheit und vor allem Intimität eigen ist, birgt aber immer auch ein Missbrauchspotenzial: Denn kommt es im intimen Bereich zu Übergriffen, im schlimmsten Fall zu sexualisierter Gewalt, ist die Offenlegung des Missbrauchs überaus mit Scham behaftet, weshalb sich mögliche Opfer vor einem solchen Schritt aus Selbstschutz nicht selten scheuen. Sich dieser zweiten Seite von Privatheit und Intimität bewusst zu bleiben, negiert nicht den grundsätzlichen Schutzwert dieser Güter und Räume, mahnt aber dazu, sie auch nicht unkritisch zu überhöhen. Privatheit und Intimität bilden keinen Legitimationskontext dafür, andere moralisch oder rechtlich zu verletzen.

Gegen die libertäre Begründung, dass nur eine möglichst extensive Privatsphäre die persönliche Freiheit eines Menschen garantieren könne, ist etwa aus kommunitaristischer Sicht eingewendet worden, hinter dem Bollwerk der Privatsphäre verschanzten sich vor allem egoistische Ansprüche gegenüber berechtigten Gemeinschaftsanliegen, die so die Gesellschaft in ihre individualisierten Einzelteile atomisierten. An diese Kritik knüpft im Kontext von Big Data neuerdings auch die Post-Privacy-Bewegung[286] an. Sie fordert im Interesse einer transparenten und insofern fairen Gesellschaft die Offenlegung sämtlicher Aktivitäten aller Bürger. Ihre Vertreter vergessen allerdings, dass die Wahrung der in der Privatsphäre geschützten Möglichkeit von Intimität und Vertraulichkeit vor allem auch dem Respekt vor der zwischenmenschlichen

[286] Siehe Heller 2011.

Kommunikation des sozialen Nahbereiches und nicht zuletzt der großen Verletzlichkeit geschuldet ist, der jede Person in den Fragmenten und Suchbewegungen ihrer eigenen Lebensgestaltung in wechselndem Maße ausgeliefert ist.

Damit deutet sich eine normative Begründung von Privatheit an, die in deren (sozial-) anthropologischer Bedeutsamkeit gründet. Nur in der Sphäre des Privaten können sich solche sozialen Nahbeziehungen (zu Lebenspartnern, Familienangehörigen, Freunden, Nachbarn) wie auch die Entwicklungsbedingungen personaler Identität (immer neu) ausbilden. Privatheit eröffnet Räume von Intimität und Vertraulichkeit, in der Personen Beziehungen pflegen, in denen sie unbefangen und unverstellt sie selbst sein können – nach außen abgeschirmt, nach innen aber offen. Solche Intimität und Vertraulichkeit ist vermutlich unerlässlich für das Wagnis, sich auf eine immer unsichere, suchende Lebensgestaltung einzulassen und damit das bleibend Fragmentarische jeder selbstbestimmten Lebensführung zu bejahen und zu praktizieren. Nur in dieser Privatheit können Personen ohne Furcht vor einer schonungslos voyeuristischen Öffentlichkeit ihre Unbefangenheit bewahren und wechselseitig erleben.

Das Verhältnis jedes Menschen zu sich selbst würde sich wohl dramatisch verändern, wenn er davon ausgehen müsste, dass er prinzipiell zu jeder Zeit und an jedem Ort beobachtet werden könnte. Davor schützt die gesicherte Privatheit. Umstritten ist jedoch, wie der Begriff des Privaten im Einzelnen aufgefasst werden sollte. Eine Möglichkeit besteht in einer von der Philosophin Beate Rössler entwickelten Konzeption, die sie auf wenige Funktionen beschränkt, aber mit einem starken normativen Anspruch ausstattet. Sie begreift Privatheit als Grundbedingung selbstbestimmter Lebensführung, allerdings nicht nur formal als Schutz gegenüber Ein- und Angriffen. Vielmehr verwirklicht sich dieser Konzeption zufolge Privatheit vor allem in der Kontrolle über den Zugang zu unmittelbarer Kontaktaufnahme zwecks Aufbau einer sozialen Nahbeziehung ebenso wie zu Informationen über intime Details der eigenen Lebensführung („Zugang zur eigenen Person“).[287]

Solche Kontrollmöglichkeiten ergeben sich demnach in dreierlei Hinsicht: Zum einen muss es möglich sein, Handlungs- und Entscheidungsspielräume überhaupt nutzen zu können, zum anderen muss man die Kontrolle darüber behalten, was andere über einen wissen dürfen, und schließlich muss es auch Rückzugsorte geben, in denen und aus denen heraus Selbstbestimmung entwickelt werden kann. Entsprechend können nach diesem Ansatz drei Formen von Privatheit unterschieden werden: entscheidungsrelevante, informationelle und räumliche Privatheit. Unter Hinweis auf die entscheidungsrelevante Privatheit wird in dieser Konzeption

[287] Vgl. Rössler 2001, insbesondere 23-26.

deutlich, dass es bestimmter Bedingungen und Fähigkeiten bedarf, um solchen Ein- und Angriffen widerstehen zu können. Eine solche Privatheitsdimension wird heute vielfach mit dem Gedanken der psychischen Widerstandskraft (Resilienz) verknüpft.[288] Resilienz stellt sich freilich nicht naturwüchsig ein; sie muss in Prozessen der Erziehung und Bildung erworben, gefördert und nicht selten gegen den uniformierenden Sog gesellschaftlicher Anpassungs- und Kontrollerwartungen verteidigt werden.

Andere Konzepte deuten den Begriff der Privatheit vor allem im Hinblick auf seine höchst unterschiedlichen Verwendungszusammenhänge zurückhaltender. Soziologische und juristische Untersuchungen zeigen, dass sich ein gemeinsamer Nenner für ein wenigstens rudimentär geteiltes Verständnis von Privatheit kaum finden lässt.[289] Gleichwohl wird auch in zurückhaltenden Konzeptionen von Privatheit eine differenzierte Taxonomie möglicher Privatheitsgefährdungen im Bereich digitaler Kommunikation entworfen, die mit der Erfassung, Analyse und neuen Verknüpfung von Daten und Informationen sowie mit vorsätzlichen Eingriffen in das Lebensumfeld des Datensubjekts einhergehen.[290] Gerade mit den in Abschnitt 2.4.1 genannten Möglichkeiten zur Deanonymisierung von ursprünglich anonymisierten Daten verbindet sich erkennbar eine mögliche Verletzung der Privatsphäre. Auch die zunehmenden Möglichkeiten, ursprünglich getrennte Daten – womöglich ohne Wissen und Zustimmung des Betroffenen – mit der Maßgabe zu verknüpfen, daraus Korrelationen zu ermitteln, haben das Potenzial, Privatheit im Sinne von Kontrollmöglichkeit über Datenströme einzuschränken.

Dennoch lehnen Menschen in der digitalen Gesellschaft die Weiterverwendung ihrer Daten, selbst dann, wenn sie dieser etwa aufgrund alternativloser Geschäftsbedingungen nur unfreiwillig zugestimmt haben, oftmals nicht rundum als Verletzung ihrer Privatsphäre ab. Sozialwissenschaftliche Studien zu individuellen und gesellschaftlichen Einstellungsmustern bestätigen, dass Menschen ihren Privatheitsansprüchen in der analogen wie auch in der digitalen Welt je nach Kontext und mitunter widersprüchlich unterschiedliches Gewicht beimessen.[291] Zudem erweisen sich Einstellungsmuster, Erwartungen und Erfahren als prägend: In sozialen Medien erweitert man den Kreis der „Freunde", mit dem man Privates teilt, und gibt dabei mehr preis, als man dies vor der Nutzung dieser Medien getan hätte. Man glaubt, durch die Begrenzung des „Freundeskreises" der Weiternutzung kontrolliert begegnen zu können.

Selbst das psychologisch nachvollziehbare Muster, die Regel, Vertrautes normalerweise nur mit Freunden zu teilen, in besonderen Situationen zu durchbrechen – sich beispielsweise auf einer Bahnfahrt einem Wildfremden mit sehr persönlichen Äußerungen zu offenbaren, dürfte im digitalen Zeitalter häufiger als früher sein. Auch wenige unter dem vermeintlichen Mantel der

[288] Siehe zum Beispiel Berndt 2013.
[289] Vgl. Solove 2007; 2008; 2011 und Solove/Hartzog 2014.
[290] Vgl. Solove 2007.
[291] Vgl. Acquisti/Brandimarte/Loewenstein 2015.

Anonymität unvorsichtig getätigte Äußerungen lassen sich dank der neuen digitalen Datenverknüpfungsmöglichkeiten und Entwicklungen, wie der sich langsam etablierenden Gesichtserkennung, immer besser einer Person zuordnen – und dies vielleicht Jahre später. Dass in solchen prekären Konstellationen gesundheitsrelevante Daten und ihre spätere, mit Big-Data-Methoden aufbereitete Zuordnung zu besonderer Vulnerabilität führen können, wenn beispielsweise die Analyse von bestimmten Äußerungen in sozialen Medien Hinweise auf psychische Auffälligkeiten ergibt, dürfte unmittelbar einleuchten.[292]

Auch wenn viele Menschen trotz der vielleicht nicht hinreichend eingeschätzten Risiken ahnen mögen, dass in der digitalen Gesellschaft eine vollständige Kontrolle der eigenen Datenspuren unmöglich geworden ist, bedeutet das noch nicht, dass ihnen gleichgültig ist, wie ihre Daten genutzt und weiterverwendet werden. Je nach Lebenskontext und Präferenzen entscheiden sich Menschen, wo sie einerseits verstärkt auf Datenschutz und Kontrolle setzen und höhere Vertrauenserwartungen an Datenverarbeiter richten und wo sie andererseits eher gelassen oder gar fahrlässig agieren. Zugleich ist vielen wichtig, dass sie die Option haben, ihre diesbezüglichen Präferenzen und Entscheidungen ad hoc zu ändern.[293]

In dem Maße, in dem die Forderung, als individuelles Datensubjekt die Verwendung der eigenen Daten selbst zu kontrollieren, zunehmend als Illusion wahrgenommenen wird, gewinnt – mit Blick auf die Zufriedenstellung der verbleibenden Ansprüche an den Umgang mit Daten – die Erwartung an Bedeutung, Datenströme würden von den darauf zugreifenden Organisationen oder Unternehmen vertraulich behandelt. So, wie viele Autofahrer das technische Innenleben ihres Fahrzeugs nicht begreifen, sondern auf die Reputation einer bestimmten Marke als Qualitätssiegel setzen, unterstellen Menschen bei der Erfassung und Weitergabe ihrer Daten eine hinreichend robuste Vertraulichkeit ihrer wichtigsten Internetdienste zum Schutz ihrer Privatheit. Ein solcher Vertrauensvorschuss ist offensichtlich unerlässlich. Denn darauf zu setzen, dass der Einzelne eigenverantwortlich die technische Kompetenz und auch Zeit aufbringen kann, sich durch komplexe allgemeine Geschäftsbedingungen zu arbeiten und die dort angesprochenen Konsequenzen der von ihm erwarteten Einwilligungen zu durchschauen, dürfte für normale Internetnutzer unrealistisch sein. Die Möglichkeiten, Daten zu nutzen, zu verknüpfen und zu verbreiten sind unter Big-Data-Bedingungen so komplex und undurchschaubar, dass die meisten Menschen mit der Aufgabe, diese für den jeweiligen Kontext zu verstehen und auf dieser Grundlage situationsgerecht einzuwilligen, weit überfordert wären. Für sie verbliebe entweder nur die wirklichkeitsfremde Option, sich gänzlich aus dem Online-Leben zu verabschieden[294] oder aber die durchaus realistische Option, sich der Daten-Politik einer entsprechenden

[292] Siehe hierzu in Abschnitt 4.7.2 das Beispiel des von Facebook entwickelten Algorithmus zur Erkennung von Suizidabsichten.
[293] Vgl. Lauss et al 2011.
[294] Vgl. unter anderem Conley et al. 2012; Acquisti/Brandimarte/Loewenstein 2015; Floridi 2014.

Organisation anzuvertrauen, die gleichsam treuhänderisch die Daten- und Privatheitsinteressen des Einzelnen vertritt.

Keinesfalls soll der Einzelne aus der Verantwortung für Dinge entlassen werden, die er selbst entscheiden kann, über die er sich informieren kann oder für die er durch Delegation an kompetente Stellen Vorsorge treffen kann – gerade im Bereich von gesundheitsrelevanten Lifestyle-Produkten (siehe Abschnitt 4.7). Dennoch gilt vorrangig: Um Privatheit als grundsätzliche Kontrollmöglichkeit über die eigenen Daten zu begreifen, rückt eine Konzeption zum Schutz der Privatheit ins Zentrum ethischer Reflexion, die weniger auf das individuelle Verhalten als auf organisationelle Vertraulichkeitsstandards setzt.[295] Auch diese Konzeption sieht einige elementare moralische und politische Normen und Regeln von Privatheit vor – besonders das Kriterium, nicht durch Privatheitsgefährdungen Schaden leiden zu müssen (siehe Abschnitt 4.4). Allerdings wird man den aktuellen Herausforderungen in der digitalen Gesellschaft mit einem fest gefügten normativen Kernbestand von Privatheit kaum effektiv begegnen können. Deshalb sind stattdessen die Kontextbedingungen jeweiliger Privatheitssicherungen und -gefährdungen online wie offline zu identifizieren. Dann stellt sich die Frage, welche Güter, Werte und Identitätsbilder für eine Person relevant sind und in welchem Kontext sie Vertraulichkeitsschutz besonders erforderlich erscheinen lassen.

Zwar mögen viele Informationen, die eine Person in einem bestimmten Zusammenhang über sich preisgibt, ursprünglich an diesen Kontext gebunden sein. So ist die Angabe einer häufig genutzten Stammroute beim Auto- oder Bahnfahren für den Nutzer von Navigationssystemen oder Online-Fahrausweisen sogar häufig sehr bequem. In einer digitalen Gesellschaft und speziell unter Big-Data-Bedingungen können solche zunächst kontextuell gebundenen Informationen jedoch in Echtzeit vielfach und unvorhersehbar verknüpft werden. Zugleich werden bislang noch keine effektiven und sicheren Vergessens- bzw. Löschverfahren angeboten, die verbindlich und dauerhaft wirken.

Deshalb ist es mindestens fahrlässig, gutgläubig vorauszusetzen, man habe doch nichts zu verbergen. Denn diese sprichwörtliche Unschuldsbeteuerung heißt doch nur, dass man derzeit keinen Verwendungszusammenhang kennt, in dem einem das Öffentlichwerden dieser oder jener Information zum Nachteil gereichen würde. Eine solche bislang schon problematische Gutgläubigkeit ist unter den Bedingungen von Big Data endgültig unangemessen: Niemand kann wissen, ob es Verwendungszusammenhänge bereits gibt oder geben wird, in denen die Nutzung bestimmter Informationen für den Datengeber zumindest unerwünscht und unangenehm, wenn nicht sogar schädlich ist. Denn was in einem bestimmten Zusammenhang heute unproblematisch erscheint (feuchtfröhliche Vergnügungen im Urlaub und Ähnliches), kann

[295] Vgl. Nissenbaum 2009 und Nissenbaum 2011.

durch eine retrospektive Durchleuchtung der eigenen Lebensführung (etwa im Zuge einer Stellenbewerbung) im Nachhinein Erkenntnisse ans Licht bringen, die für die weitere Lebensplanung einer Person oder ihres Umfeldes gravierende Konsequenzen nach sich ziehen können.

An den zunehmenden Möglichkeiten, intime Details digital vielfältig preiszugeben, zeigen sich zudem negative Auswirkungen der selbstinduzierten Fremdbestimmung (siehe Abschnitt 2.5.5) bzw. der informationellen Selbstgefährdung (siehe Abschnitt 3.1) besonders plastisch: Die selbst vorgenommene Bereitstellung persönlicher Informationen schlägt um in eine persönliche Lebensführung, die sich maßgeblich von äußeren Einflussfaktoren abhängig macht. Je mehr aber solche ungeahnten und vom Betroffenen nicht mehr zu kontrollierenden Verknüpfungsmöglichkeiten zunehmen, desto nötiger wird es, bewährte Vertrauensstandards, die Menschen aus ihrem Offline-Leben kennen und wertschätzen, auch online rechtlich und sozialethisch zu etablieren.[296]

Wie Privatheit unter den Bedingungen von Big Data über verschiedene Wege (Datenschutz, höhere Vertrauenserwartungen an Datenströme usw.) zu schützen und zu wahren ist, betrifft nicht nur Individuen, sondern auch Gruppen von Individuen. Deutlich wird dies im medizinischen Kontext, wenn der Anspruch eines Einzelnen, seine genetische Disposition zu erfahren, andere Personen berührt oder belastet. In diesem Zusammenhang wird immer wieder auf das Recht auf Nichtwissen hingewiesen, demzufolge etwa Familienmitglieder von möglichen genetischen Belastungen in der Familie selbst dann keine Kenntnis nehmen müssen, wenn einem anderen Familienmitglied mit diesem Wissen medizinisch geholfen werden könnte. Wie dieses Recht auf Nichtwissen juristisch durchsetzbar ist und wann es in Spannung zum Recht auf Wissen des anderen Familienmitglieds geraten kann, wird rechtlich und ethisch kontrovers diskutiert.[297]

Unter Big-Data-Bedingungen verschärft sich die Frage nach Privatheitsansprüchen nochmals, da die Möglichkeiten, durch die Analyse von Massendaten auf Merkmalskombinationen von weiteren Personen zu schließen, weit über den Familienkreis hinausgehen und auch größere Personengruppen betreffen können. Betroffene werden – für sie oft völlig undurchsichtig – von Algorithmen zu einer Gruppe zusammengefasst, der ein bestimmtes Zielmerkmal oder Etikett zugewiesen wird. Ein solches Merkmal kann im Ernstfall stigmatisierende, diskriminierende oder exkludierende Rückwirkungen zeitigen. Bestimmte Aspekte der Lebensführung, einzelne Verhaltensweisen oder auch demografische Merkmale, an denen der Betroffene wenig ändern kann, können etwa aufgrund ihrer statistischen Korrelation mit gesundheitlichen Merkmalen (siehe Kapitel 2) gesellschaftlich als negativ oder seitens bestimmter Interessenten (etwa Kranken- oder Lebensversicherungen) als risikobehaftet eingestuft werden (siehe Abschnitt 4.6).

[296] Vgl. Conley et al. 2012.
[297] Vgl. etwa Duttge 2010; Koppernock 1997, 89 ff.; Dorniok 2015 und Schroeder 2015.

Solche in Big-Data-Analysen aufgedeckten Korrelationen, auf deren Grundlage Gruppen mit möglichst genau definierten Merkmalskombinationen identifiziert werden sollen, rechtfertigen es aber noch nicht unbedingt, alle individuellen Träger solcher Merkmalskombinationen ohne Weiteres den so konstituierten Gruppen zuzurechnen, da weitere, in der Analyse nicht erfasste Merkmale die persönliche Wahrscheinlichkeit, auch das jeweilige Zielmerkmal auszubilden, erheblich beeinflussen können.

Im Gegensatz zur klassischen medizinischen Situation, in der die Erhebung solcher Merkmalskombinationen und die Kommunikation der damit verbundenen Risikoeinschätzungen streng vertraulich erfolgen kann, ist dies unter Big-Data-Bedingungen kaum mehr möglich, da zunehmend auch öffentlich verfügbare und/oder zumindest auf den ersten Blick nicht offensichtlich gesundheitsrelevante Merkmale in Analysen einfließen. Für den Einzelnen mag dies bedeuten, dass er seine Zuordnung zu einem bestimmten Risikoprofil nicht nur weder verhindern noch durchschauen kann, sondern dass auch Dritte eine entsprechende Zuordnung anhand der verfügbaren Merkmalskombination leicht vornehmen oder nachvollziehen können. Damit verschärfen sich die potenziellen Gefahren für die Privatsphäre, da sie nicht mehr nur einzelne Individuen oder kleine (Familien-)Gruppen betreffen können, sondern zunehmend auch größere Gruppen, denen aufgrund von durch Big-Data-Analysen extrahierten Merkmalskombinationen bestimmte Risikoprofile zugewiesen werden.

4.3 Souveränität und Macht

Privatheit ist für Selbstbestimmung als lebensgeschichtlichen Ort von Freiheit essenziell. Wenn Privatheit in der Kontrolle über den Zugang zu allen Orten und Vollzügen des eigenen Lebens besteht, dann gilt dies sowohl offline wie online, zumal diese Sphären angesichts der zunehmenden digitalen Vernetzung unserer Umwelt zunehmend verschmelzen.[298] Das Sammeln und in der Folge das Auswerten, Aufarbeiten und Verbreiten persönlicher Daten ist immer ein Eingriff in die Privatsphäre, der legitimationsbedürftig ist.[299] Mit diesem Verständnis von Privatheit verbindet sich ein Verständnis von Selbstbestimmung in der besonderen Form von Selbstgestaltung, für das sich auch in der Alltagssprache der Begriff der *Souveränität* etabliert hat und das zum Begriff der *Datensouveränität* weiterentwickelt werden kann (siehe Kapitel 5).[300]

Freilich ist der Begriff der Souveränität keineswegs unproblematisch und deshalb erläuterungsbedürftig. Einerseits enthält er einen emphatischen Ton: Als moralisches Hochwertwort signalisiert er einen Anspruch auf unbedingte Beachtung, der kaum infrage zu stellen sei. Persönliche

[298] Luciano Floridi benutzt das Kunstwort „onlife", um deutlich zu machen, dass sich Online- und Offline-Leben kaum noch unterscheiden lassen. Vgl. Floridi 2014.
[299] Vgl. Pillay 2014.
[300] Vgl. Friedrichsen/Bisa 2016.

Souveränität gilt vielen als Grundlage und Zieloption eines gelingenden, selbstbestimmten Lebens. So wundert es nicht, dass der Begriff in zahlreichen Lebensbereichen die verschiedensten Konkretionen und Variationen seiner Bedeutung erfährt – von Konsumenten- bzw. Kundensouveränität[301] über Patientensouveränität in der Pflege[302] und Ernährungssouveränität[303] bis hin zur Saatgut-Souveränität[304] (dem ungehinderten Zugang zur Sortenvielfalt des Saatgutes). Ähnlich wird im Bereich der digitalen Medien Souveränität über unterschiedliche Verwendungen des Begriffs der Datensouveränität[305] bis hin zur Filtersouveränität ausbuchstabiert, die das eigene Extrahieren relevanter Informationen aus der Datenflut mithilfe von selbst gewählten Filtern und Analysewerkzeugen kennzeichnen soll.[306] Andererseits kontrastiert diese Inflation von Souveränitätsansprüchen und -bereichen mit dem real erlebten Verlust an Souveränität – jedenfalls dann, wenn man die Bedeutung des klassischen Souveränitätsbegriffs zum Maßstab nimmt. Die Rede von digitaler Souveränität gleicht dann manchmal einer Beschwörungs- und Beschwichtigungsformel.

Der Begriff der Souveränität entstammt kulturhistorisch vornehmlich dem religiös-politischen Bereich. Auch wenn die ideengeschichtliche Entwicklung des Souveränitätsbegriffs keineswegs ohne gravierende Brüche ist, lässt sich aus seiner kulturhistorischen Rekonstruktion viel lernen. Souveränität galt als jene Eigenschaft des (monotheistischen) Gottes, kraft deren er absolut und unbedingt von anderen Mächten und Gewalten alles zu tun oder zu lassen imstande sei. Dieser Anspruch des Unbedingten und Unabhängigen wird zu Beginn der Neuzeit in der staatstheoretischen Konzeption Jean Bodins legitimatorisch auf die Souveränität eines absolutistischen Herrschers bezogen. Dessen Herrschaft werde durch keine andere Macht eingeschränkt, er schulde niemandem moralisch oder rechtlich Rechenschaft und verkörpere in seinem Herrschaftsgebiet (nach Gott) die höchste Autorität.[307] Solche absolute Unbedingtheit gewährt stets die Möglichkeit von Willkür.

---

[301] Sie bildet gewissermaßen den normativen Kern des zumindest methodologischen Individualismus in der (neo-) klassischen Ökonomie (vgl. Samuelson 1975).
[302] „Patientensouveränität bedeutet die Möglichkeit und Fähigkeit des Patienten, als Nachfrager nach und Verbraucher oder Nutzer von gesundheits- und krankheitsbezogenen Versorgungsleistungen verschiedene Alternativen bei (annähernd) gleicher Indikation abzuwägen und die in seinem Sinne am besten geeignete auszuwählen." Vgl. Struppek 2010, 62.
[303] Vgl. http://www.weltagrarbericht.de/themen-des-weltagrarberichts/ernaehrungssouveraenitaet/ernaehrungssouveraenitaet-volltext.html [17.10.2017]
[304] Vgl. http://www.saatgutkampagne.org [17.10.2017]
[305] Der Begriff der Datensouveränität wurde beispielsweise von Politikern genutzt, um eine Erweiterung bzw. Erneuerung von Datenschutzkonzepten zu beschreiben, insbesondere mit Blick auf Datensparsamkeit (vgl. https://www.heise.de/newsticker/meldung/IT-Gipfel-2016-Merkel-plaediert-fuer-Datensouveraenitaet-statt-Datenschutz-3490629.html [17.10.2017]). Die Bertelsmann Stiftung untersucht anhand des Begriffes, wie die „Kontrolle" der eigenen Daten ermöglicht werden kann (vgl. Bertelsmann Stiftung 2017 und vom Bundesverband Digital Wirtschaft e.V. wurde der Begriff auf einer Tagung im Juni 2017 ohne nähere inhaltliche Bestimmung als Schlagwort für aktuelle Herausforderungen der Digitalisierung verwendet (vgl. http://www.data-summit.de [17.10.2017]).
[306] Vgl. Seemann 2011, 79.
[307] Vgl. Bodin 1976, 39 ff.

Es sind vermutlich diese Momente von absoluter Unbedingtheit und potenziell despotischer Willkür, weshalb Souveränitätsansprüche jenseits der politischen Theorie und Staatswissenschaften auch in ethischen Debatten kritisiert werden.[308] Vor diesem Hintergrund sind Konzepte von Souveränität entwickelt worden, die anstelle einer vermeintlichen absoluten Ungebundenheit des souveränen Subjekts die unhintergehbaren Abhängigkeiten seiner physischen wie sozialen Leiblichkeit betonen.[309] Solche Souveränität muss keineswegs zu einem Fatalismus führen, der sich mit den Abhängigkeiten und Limitationen der persönlichen Lebensführung widerspruchslos abfindet, im Gegenteil.

Von solchen Nuancierungen eines moralphilosophischen Souveränitätsbegriffs kann seine Weiterentwicklung für den digitalen Bereich (siehe Kapitel 5) profitieren. Dass sich ein Mensch in absolut ungebundener Souveränität den Interessen am Zugriff auf seine persönlichen Daten beliebig öffnet oder verschließt, ist denkbar, aber wenig realistisch und lebensfremd. Solche absolut ungebundene Souveränität könnte dort in ein moraltheoretisch schlüssiges wie realistisches Konzept von digitaler Souveränität Eingang finden, wo sie von dem religionsgeschichtlich wirksamen Grundmotiv des Souveränitätsdenkens her gedeutet wird. Die Pointe dieses Souveränitätsverständnisses – und darin liegt die Bedeutung dieses kulturhistorischen Motivs – besteht jedenfalls einer Lesart zufolge darin, dass der absolut unbedingte Souveränitätsanspruch des monotheistischen Gottes sich religionsgeschichtlich herausgebildet hat als kritische Reaktion auf die als Anmaßung empfundenen Souveränitätsansprüche innerweltlicher, sich nicht selten als Gottheiten zelebrierender Mächte (etwa ägyptische Pharaonen oder römische Kaiser). Absolute Souveränität auf eine Gottheit zu übertragen, schloss in der Folge jeden Versuch aus, eine Souveränität (im Sinne absoluter Verfügungsgewalt) von Menschen über andere Menschen zu etablieren. Die Vorstellung der absoluten Souveränität einer Gottheit diente gleichsam als regulative Idee, mit der zwischenmenschliche Verfügungsgewalten strikt begrenzt werden konnten. Dieses herrschaftskritische Motiv des Souveränitätsverständnisses kann in ähnlicher Weise Versuchungen entgegenwirken, auf der digitalen Souveränität jedes einzelnen Menschen so zu bestehen, dass eine absolute Verfügungsmacht dessen besteht, über den Daten erhoben und in welcher Weise auch immer weiterverarbeitet werden. Folgt man dieser möglichen Deutungsvariante des Souveränitätsbegriffs, dann sind personenbezogene Daten für die Sammler und Nutzer grundsätzlich nur Leihgabe, niemals frei und willkürlich verfügbares Eigentum. Das bedeutet umgekehrt nicht, dass damit der Datengeber automatisch Eigentümer seiner Daten ist oder selbst seinen Souveränitätsanspruch unter allen Umständen realisieren kann (siehe Kapitel 3). Es muss umgekehrt aber auch keinesfalls zu einem Fatalismus führen, der sich mit den Abhängigkeiten und Limitationen der persönlichen Lebensführung widerspruchslos abfindet.[310]

[308] Vgl. Klein 2016.
[309] Vgl. Böhme 2008, 188.
[310] Vgl. ebd., 194.

Natürlich können Souveränitätsansprüche auf andere übertragen werden – etwa auf Repräsentanten oder – wie bei bislang souveränen Nationalstaaten – auf supranationale Organisationen (wie etwa die Europäische Union). Solche Übertragungen entfalten mitunter eine Eigendynamik, die zu teils erheblichen Souveränitätseinbußen des ursprünglichen Souveräns führen, die kaum noch rückgängig gemacht werden können. Auf dem Souveränitätsanspruch des ursprünglichen Souveräns dennoch gleichsam kontrafaktisch zu bestehen, schützt ihn aber vor einer totalen Vereinnahmung seitens des jeweils faktischen Souveräns und beschränkt dessen Souveränitätsausübung.

Aus diesem Souveränitätsverständnis folgen im Prinzip weitreichende Kontrollmöglichkeiten des Individuums, die zugleich in vielfältige Interaktionszusammenhänge eingebunden sind und damit regulatorisch eine Multiakteursperspektive nahelegen (siehe Abschnitt 4.7). Beispielsweise kann, was als „Recht zur Löschung" bestimmter personenbezogener, nicht anonymisierter Daten individuell noch realisierbar erscheint, im Kontext massenhaft erhobener und verarbeiteter Daten effektiv nur in kollektiver Verantwortung (siehe Kapitel 4) wahrgenommen werden: als eine Art Recht auf Inspektion datenverarbeitender Institutionen. Darüber hinaus ist zu prüfen, ob analog zum Verfahren bei Biobanken dynamische und kaskadisch strukturierte Einwilligungsmodelle zumindest teilweise auch im weiteren Gesundheitsbereich Anwendung finden können (siehe Abschnitt 4.1.2). Im Sinne der in Abschnitt 4.6 entwickelten Überlegungen zur Solidarität und zum Verständnis von Souveränität, das sich auch im Blick auf die Verletzlichkeit anderer bestimmt, ist zu bedenken, ob Souveränität nicht einschließt, den Verzicht auf bestimmte Kontrollmöglichkeiten und Widerrufmöglichkeiten zu akzeptieren. Zudem erschöpft sich Souveränität keineswegs in einem Schutzkonzept, sondern umfasst auch ein Teilhaberecht: nämlich als Recht auf offenen Zugang zur Online-Welt und insbesondere zu den Wissensbeständen und neuen Erkenntnissen, die durch Big Data generiert werden.[311]

Souveränität verwirklicht sich im Modus der Ausübung von Macht und wird umgekehrt begrenzt durch die Ausübung souveräner Macht anderer. Diese reziproke Beschränkung jeweils individueller Macht bedeutet nicht, dass souveräne Akteure wechselseitig Herrschaft ausübten. An einer in neueren Machttheorien geläufigen funktionalen Unterscheidung zwischen „Macht *zu* (etwas)" und „Macht *über* (jemanden)" lässt sich das erhellen.[312] In ihrer universal-reziproken Funktion, die Macht anderer zu begrenzen, verwirklicht sich die jeweils eigene Souveränität

[311] Vgl. Kettemann 2015.
[312] Vgl. Lovett 2007.

von Personen lediglich als Macht zu etwas, nämlich zu legitim-selbstbestimmtem Handeln, nicht dagegen als Macht über andere.[313]

Der Begriff der Macht erstreckt sich über ein weites Spektrum unterschiedlicher Bedeutungen und eine kaum überschaubare Vielfalt der damit bezeichneten Phänomene. Eine umfassende und zugleich gehaltvolle Definition dürfte nicht möglich sein.[314] Als plausibler Ausgangspunkt einer konzeptuellen Analyse individuell-personaler Macht mag aber noch immer die Definition dienen, die Thomas Hobbes, der erste neuzeitliche Theoretiker der Macht, im „Leviathan" vorschlägt: „Die Macht eines Menschen besteht, allgemein genommen, in seinen gegenwärtigen Mitteln zur Erlangung eines zukünftigen anscheinenden Guts".[315] Das erfasst sowohl Konzeptionen der *Macht zu* als auch solche der *Macht über* und ist insofern als allgemeine Definition einleuchtend. In seiner extremen Abstraktheit ist es aber nahezu inhaltsleer. Ähnliches gilt für die soziologische Definition Max Webers, die im Unterschied zur hobbesianischen den Begriff zudem nur in dessen Bedeutung als Macht über andere erfasst: als „Chance, innerhalb einer sozialen Beziehung den eigenen Willen auch gegen Widerstreben durchzusetzen, gleichviel worauf diese Chance beruht".[316]

In den vergangenen Jahrzehnten sind spezifischere Machtkonzeptionen vorgeschlagen worden. Diese nehmen nicht primär den Akteur der Machtausübung zum Ausgangspunkt, sondern richten den Blick einerseits auf das gesellschaftliche Umfeld von Machtrelationen[317] und andererseits auf die Ziele bzw. Adressaten der Machtausübung. Die letztgenannte Perspektive ist ersichtlich als die eines Blicks auf Formen der Macht über andere. Jenseits der Anwendung von Gewalt und unmittelbarem Zwang, die hier keine Rolle spielt, kann man die Modi einer Ausübung von *Macht über* weiter unterscheiden in erstens die Manipulation der Präferenzen und Überzeugungen anderer (etwa durch die Verbreitung sozialer Mythen) und zweitens über diese Einflussnahme hinaus die Manipulation der Subjekte selbst:[318] durch die subtile Formung, Veränderung und damit zuletzt die mögliche Beherrschung ihrer Charaktere.

---

[313] Freilich löst die illegitime Überschreitung jener Grenze durch Dritte Abwehrbefugnisse aus. Sie mögen im Einzelfall bis zur Selbsthilfe durch Notwehr reichen und damit eine spezifische Form von Machtausübung über jene Dritten legitimieren.

[314] Macht gehört zu jenen komplexen Begriffen, deren weitverzweigtes Bedeutungsnetz sich kaum in die Klammer einer allgemeinen Definition (also unter die Maßgabe notwendiger und hinreichender Bedingungen) zwingen lässt. Fassbar werden sie nur im Modus der Feststellung ihrer „Familienähnlichkeit" mit jeweils paradigmatischen (und konsentierten) Grundtypen des Begriffs; Vgl. Wittgensteins Philosophische Untersuchungen, §§ 65, 66 (Rhees 1969).

[315] Hobbes 1999, 66. Das etwas seltsam anmutende „anscheinend" (im Original „apparent") bezeichnet die von Hobbes für maßgeblich erklärte rein subjektive Bestimmung des erstrebten Guts durch den Inhaber der Macht.

[316] Weber 1980, Teil 1, § 16.

[317] So etwa die Konzeption von Macht als „power through control of the agenda" in Bachrach/Baratz 1962. Damit ist vor allem eine Form politischer Macht gemeint, die sich auf dem Wege der Begrenzung oder Beeinflussung dessen manifestiert, was überhaupt als derzeit verhandelbar und lösungsbedürftig zu gelten hat. Für die Zwecke unserer Analyse spielt diese Form der Macht keine erhebliche Rolle.

[318] Vgl. zu erstens Lukes 2005, 25-37 und zu zweitens Wartenberg 1990 und vor allem die zahlreichen (unsystematisch verstreuten) Analysen Michel Foucaults, hauptsächlich in Foucault 1976. Zu Foucaults Machttheorie siehe Kneer 2012.

Dies ist hier besonders bedeutsam. Denn der Einsatz von Big-Data-Algorithmen in der Interaktion zwischen Anbietern und Nutzern von Internetdiensten eröffnet den Erstgenannten neben anderen Möglichkeiten auch solche der gegebenenfalls gezielt strategischen Einflussnahme auf Denken, Fühlen, Handeln und damit zuletzt auf die Lebensführung ihrer Abnehmer. Solche Möglichkeiten manifestieren neuartige Formen interpersonaler Macht. Das allein ist kein zwingender Grund, sie für ausnahmslos unzulässig zu halten. Einige davon sind dies aber jedenfalls, und andere mögen der Notwendigkeit einer jeweils besonderen Rechtfertigung unterliegen. Im Kontext von Big Data und Gesundheit geht es weniger um offensichtliche Machtausübung, die Menschen gegen deren Willen zur Preisgabe ihrer persönlichen Daten zwingt. Solcher Zwang kann zwar Ausübung legitimen Rechtszwangs sein; dann ist er als Mittel der Gefahrenabwehr – etwa in bestimmten Bereichen des öffentlichen Gesundheitswesens – gerechtfertigt. Beschränkungen der Souveränität resultieren jedoch oft aus subtiler wirkenden, verdeckten Machtfaktoren – etwa über eine Steuerung von Präferenzen, die sich Personen unbewusst aneignen und als Determinanten unreflektiert in ihre Entscheidungen einfließen lassen.[319]

Um dies genauer zu klären, sind einige weitere Unterscheidungen erforderlich. Die Möglichkeiten der Einflussnahme auf fremde Überzeugungen und Dispositionen erstrecken sich über ein weites Spektrum denkbarer Formen. Nicht alle davon erscheinen moralisch dubios oder spezifisch rechtfertigungsbedürftig. In diesem Sinn liegen am ethisch unverdächtigen Ende des Spektrums etwa Weisen der Überzeugung durch die Kraft vernünftiger Argumente. Dicht daneben wären moderat paternalistische, aber dennoch „libertäre", also Freiheit und Selbstbestimmung achtende Versuche einzuordnen, bestimmte Verhaltensweisen mithilfe subtil motivierender mentaler Nudges zu beeinflussen. Bekannte Nudges im Gesundheitsbereich sind etwa die verhaltenspsychologisch geleitete Umgestaltung von Kantinen mit dem Ziel, Menschen zur Wahl gesünderer Mahlzeiten zu bewegen; Bonuspunkte oder Belohnungen bei Versicherungen für die Teilnahme an freiwilligen Sportprogrammen; oder die räumliche und bauliche Umgestaltung von Büros und Gebäuden, um mehr Bewegung zum Standard zu machen.[320] Für die Legitimität solcher Maßnahmen spricht unter anderem ihre Transparenz, Nutzerwohlorientierung und Sachgerechtigkeit.[321] Dagegen mögen sich Formen der Einflussnahme mittels suggestiver Strategien als moralisch suspekt darstellen. Gleichwohl sind sie im Grundsatz sozialadäquat. Über einige weitere begriffliche Unterscheidungen gelangt man schließlich an das andere,

[319] Vgl. Lukes 2005.
[320] Für eine Übersicht sowie weitere Beispiele vgl. Roberto/Kawachi 2015; Halpern 2015 oder auch Buyx 2010.
[321] Die Erforschung und Ausdifferenzierung solcher Möglichkeiten der Beeinflussung Dritter ist vor allem Gegenstand der sogenannten Verhaltensökonomie und wird deshalb primär unter wirtschaftlichen Gesichtspunkten betrieben. Längst werden Nudges aber auch als Formen legitimen politischen Handelns erwogen und angewandt. Das Institute for Government des britischen Premierministers hat in seinem Projekt MINDSPACE ein Modell moderner Governance entwickelt, das auf sechs Voraussetzungen beruht; sie implizieren ausnahmslos eine selbstbestimmte Mitwirkung der dabei adressierten Bürger: „explore, enable, encourage, engage, exemplify, evaluate" (vgl. Institute for Government 2010, 9). Vgl. auch Bröckling 2017.

sinistre Ende des Spektrums: etwa zu subliminalen, dass heißt für die Adressaten nicht wahrnehmbaren, vor allem fremdnützigen Intervention ins Ich zur Manipulation von Handlungen, Präferenzen, Überzeugungen und Charakterzügen.

All diese Formen mehr oder weniger subtiler Machtausübung durch mental invasive Verfahren sind auch im Modus des Einsatzes von Big-Data-Mechanismen denkbar. Auf der ersichtlich gleitenden Skala ihres ethischen Werts oder Unwerts bedarf es zum Zwecke ihrer normativen Beurteilung eines Kriteriums der Unterscheidung zulässiger von unzulässigen Methoden. Dafür scheint sich zunächst das Merkmal der Täuschung anzubieten. Die meisten Formen von Manipulation dürften mit einer Strategie gezielten Irreführens operieren. Allerdings implizieren keineswegs alle Täuschungen anderer auch deren Manipulation. Schlichte Lügen, etwa über eigene Lebensumstände, täuschen den Adressaten, manipulieren ihn aber nicht notwendig. Umgekehrt sind Manipulationen durchaus ohne Täuschung denkbar, etwa mittels direkter Eingriffe ins Gehirn.[322] Für den Begriff der Manipulation bezeichnet Täuschung somit weder eine notwendige noch eine hinreichende Bedingung. Erst recht nicht geeignet ist sie deshalb als Kriterium einer normativen Beurteilung des Manipulierens.

Das gesuchte Kriterium muss sich schlüssig auf den Gegenstand des Schutzes beziehen, der mit einer kategorialen Kennzeichnung unzulässiger Manipulationen beabsichtigt ist. Dieser Schutzgegenstand kann nichts anderes sein als die Selbstbestimmung des Adressaten einer manipulativen Intervention.[323] Dabei bezieht sich Selbstbestimmung in diesem Sinn nicht nur auf das jeweils konkrete Motiv einer Handlung, sondern darüber hinaus auf das gesamte Feld möglicher Handlungsmotive einer Person, auf die kognitive wie die emotionale und motivationale „Umwelt ihres Handelns“ und damit auf den „sozialen Raum“ ihrer gesellschaftlichen Existenz.[324] Jede manipulative Veränderung des sozialen Handlungsraums anderer bedeutet eine Ausübung von Macht über diese anderen und somit eine mögliche Beeinträchtigung ihrer Selbstbestimmung.

Aus dem Schutzgegenstand der Selbstbestimmung lässt sich das entscheidende Kriterium der Unzulässigkeit manipulativer Machtausübung schlüssig entwickeln. Es besteht nicht in dem (wie auch immer zu bestimmenden) schieren Gewicht der Folgen solcher Interventionen.[325] Unzulässig oder besonders rechtfertigungsbedürftig sind diese vielmehr dann, wenn sie die Möglichkeiten ihrer Adressaten zur Kontrolle der Bedingungen eigenen Handelns umgehen

[322] Vgl. Merkel et al. 2007.
[323] Vgl. dazu noch einmal Abschnitt 4.1.2 und die zum Selbstbestimmungsbegriff (und im Anschluss an Joel Feinberg) entwickelten Differenzierungen.
[324] Dazu die Begriffe „action-environment“ und „social space“ bei Wartenberg 1990, 74, 85.
[325] Auch gravierend schädliche Einwirkungen Dritter sind grundsätzlich (und bis an die Grenze der Lebensgefährdung oder dauernder schwerer Schäden) legitim, wenn sie von einer informierten Einwilligung des Betroffenen gedeckt sind.

und damit dessen Selbstbestimmtheit untergraben oder doch zweifelhaft machen. Das geschieht, wenn die Manipulation in einem *direkten* Zugriff auf das motivationale Feld möglichen Handelns erfolgt und sich somit der kognitiven Kontrolle durch den Betroffenen entzieht.[326]

Der Einsatz von Big-Data-Algorithmen in der Interaktion zwischen Anbietern und Nutzern von Internetdiensten und sozialen Medien eröffnet eine Vielzahl neuartiger Möglichkeiten zu solchen subliminalen und damit unzulässig manipulativen Einflussnahmen. Eine exemplarische Veranschaulichung dieser Möglichkeiten hat im Januar 2012 eine Facebook-Studie zu „emotionaler Ansteckung" (*emotional contagion*) geliefert, bei der knapp 700.000 nichts ahnende Facebook-Nutzer zu Forschungsobjekten wurden.[327] Mit ihr verfolgten die Wissenschaftler das (durchaus legitime) Ziel einer Klärung der Frage, ob eine Gefühlsansteckung zwischen Menschen auch außerhalb höchstpersönlicher Kommunikation, nämlich allein über digital vermittelte Kontakte möglich sei. Dies sollte sich, so die Grundidee des Projekts, ermitteln lassen, wenn man in den News Feeds der Facebook-Nutzer, also in den via Internet kommunizierten Reaktionen anderer Nutzer, die Quantität bestimmter dezidiert emotionaler Inhalte substanziell verringerte. Die Zielgruppe der ausgewählten 700.000 Nutzer wurde in einem randomisierten Verfahren zweigeteilt. 350.000 von ihnen bekamen eine Woche lang erheblich weniger positiv-emotionale Reaktionen, die anderen 350.000 dagegen erheblich weniger negativ-emotionale Reaktionen zu Gesicht, als ihnen in Wahrheit zugesandt worden waren. Über Big-Data-Verfahren wurden die jeweils gesuchten emotionalen Sprachwendungen herausgefiltert und den Nutzern hinterzogen. In den Folgetagen wurde, ebenfalls per Big Data, aus den geposteten eigenen Äußerungen jener 700.000 ahnungslosen Studienteilnehmer deren jeweiliger emotionaler Gehalt herausgefiltert und klassifiziert. Das Ergebnis: Die mit positiven Emotionen ihrer „Follower" unterversorgten 350.000 Nutzer produzierten ihrerseits in signifikant höherer Zahl emotional düster gefärbte Äußerungen als die Mitglieder der Gruppe, die man vor solchen negativen Emotionen anderer „bewahrt" hatte.

Der massiven öffentlichen wie wissenschaftlichen Kritik an dieser Studie[328] hielten deren Verfasser Facebooks vertragliche Nutzungsbedingungen und die damit verbundene Datenrichtlinie entgegen, die jeder Nutzer akzeptiert hat und nach der auch die Nutzung der Daten für „Studien" ausdrücklich vorgesehen ist.[329] Daher, so das Argument, hätten die betroffenen Nutzer in die Studie wirksam eingewilligt. Dass dies nicht richtig ist, liegt auf der Hand. Die Rechtfertigung der Autoren verweist auf den gegebenenfalls zulässigen Umgang mit den Daten der

---

[326] Den expandierenden Möglichkeiten solcher nicht autorisierten, direkten Interventionen ins „Ich" ihrer Adressaten wird deshalb zunehmend das menschenrechtliche Postulat eines „Rechts auf mentale Selbstbestimmung" entgegengehalten; grundlegend hier Bublitz/Merkel 2014.
[327] Vgl. Kramer/Guillory/Hancock 2014.
[328] Zahlreiche Nachweise auf dem wissenschaftlichen Blog „The Laboratorium" von James Grimmelmann (vgl. http://laboratorium.net/archive/2014/06/28/as_flies_to_wanton_boys [17.10.2017]).
[329] Vgl. https://de-de.facebook.com/full_data_use_policy [17.10.2017].

Nutzer zu Forschungszwecken. Sie ignoriert aber vollständig das damit verbundene – und ohne informierte Einwilligung unzulässige – Eindringen in die mentale Sphäre dieser Nutzer. Deren sensorische oder kognitive Kontrollen gegenüber externen Interventionen ins eigene Ich wurden dabei vollständig unterlaufen. Daher ging es bei der Studie nicht allein um eine (möglicherweise zulässige) Beobachtung der Nutzer mittels der Analyse ihrer Daten. Vielmehr ging es um die manifeste Veränderung ihres mentalen Befindens: eine tagelange Manipulation ihres emotionalen Zustands, und für die Hälfte der unfreiwilligen Studienteilnehmer eindeutig zum Schlechteren. Dass auf der Seite der Forscher womöglich niemand auch nur einen der dafür ausgewerteten Beiträge gelesen hat, deren Analyse ausschließlich maschinell erfolgte, ändert daran nichts.

Ähnliche Versuche, die Überzeugungen oder Präferenzen von Einzelnen oder Gruppen ohne deren Wissen und Einwilligung über die weithin akzeptierte Manipulation durch Werbung und Ähnliches hinaus zu beeinflussen und eigentlich in der Kommunikation vorausgesetzte Kontrollmechanismen in problematischer Weise zu unterlaufen, wären im Prinzip auch im gesundheitsrelevanten Bereich möglich. Bekannter und weitaus verbreiteter ist die Konsumentenbeeinflussung durch gezieltes Tracking von Online-Verhalten in der Zusammenschau mit anderen gesammelten Daten über Personen und Gruppen (siehe Abschnitt 2.4.1), sodass maßgefertigte Werbung an diese herangetragen werden kann.[330] Noch problematischer wäre eine solche Beeinflussung, wenn sie die Subjekte selbst beträfe, also eine datengestützte, subtile Formung und Veränderung der Charaktere von Subjekten darstellte.

Die genannten Beispiele demonstrieren exemplarisch Möglichkeiten, Risiken und vor allem normative Grenzen von Big-Data-basierten Verfahren, die ins mentale Innere der Datengeber eindringen. Umgehen solche Interventionen die personalen Kontrollmöglichkeiten der Adressaten, so sind sie – vorbehaltlich spezieller Rechtfertigungsgründe (wie informierte Einwilligung oder bestimmte Notstandslage) – jedenfalls moralisch dubios oder unzulässig: Ausübung einer neuen Möglichkeit illegitimer Macht. Auch nach rechtlichen Prinzipien unzulässig werden sie freilich erst dann, wenn diese subliminalen Eingriffe außerhalb der Sphäre sozialadäquaten und daher generell tolerierten Verhaltens liegen. Die meisten Formen psychologisch raffinierter Werbung, die auf unbewusste Motivationspotenziale ihrer Adressaten abzielen, oder subliminale Nudges zu moralisch unverdächtigen Zwecken fallen deshalb jedenfalls aus dem Verdikt einer Rechtsverletzung heraus. Ob sie damit auch moralisch akzeptabel sind, ist eine schwierige und im Maße der Erweiterung solcher Möglichkeiten zunehmend schwieriger werdende Frage. Sie muss für jeden Einzelfall anhand der hier entwickelten, dafür erforderlichen normativen Maßgaben geklärt werden.

---

[330] Vgl. https://ca-commercial.com/casestudies/casestudyemployeebenefitsprovider [17.10.2017] sowie Datta/Tschantz/Datta 2015.

Mit dem Konzept der Souveränität, wie es hier skizziert und mit verschiedenen Konzepten von Einwilligung für die Big-Data-Fragestellung (siehe Abschnitt 4.1.2) konkretisiert wurde, wird man solche offensichtlichen oder auch subtilen Machtausübungen nicht einfach aufheben können. Wohl aber kann man so zum Ausdruck bringen, dass ein legitimer Anspruch besteht, sie einzuhegen, damit das Individuum durch die Verwendung unzähliger seiner Daten seitens anderer nicht selbst beeinträchtigt wird.

4.4 Schadensvermeidung und Wohltätigkeit

Ein weiterer für die ethische Analyse von Big-Data-Anwendungen im Gesundheitsbereich relevanter normativer Bezugspunkt, der allerdings nicht nur aus individual-, sondern auch aus sozialethischem Blickwinkel bedeutsam ist, ergibt sich aus der moralischen Verpflichtung zur Wohltätigkeit. Zu den Grundüberzeugungen abendländischer Ethik gehört die Annahme, dass moralisch qualifiziertes Handeln in vielen Situationen über die bloße Schadensvermeidung hinaus auch einen positiven Mehrwert erbringen sollte, der insbesondere die Lebenssituation besonders hilfsbedürftiger Menschen verbessert. Je nach kulturellem Kontext ist diese elementare Intuition jedoch auf sehr unterschiedliche Weise konkretisiert und ausgestaltet worden. So hat sie etwa innerhalb der jüdisch-christlichen Tradition mit dem Gebot der Gottes- und Nächstenliebe Eingang in ein umfassendes Ethos gefunden, das auf bestimmten religiös-metaphysischen Voraussetzungen beruht und daher trotz seiner enormen Wirkmächtigkeit nicht einfach verallgemeinert werden kann.

Eine andere Gestalt nahm die Aufforderung zur Wohltätigkeit in neuzeitlichen Theorien des klassischen Utilitarismus an, mit denen versucht wurde, das sittlich richtige Handeln vermittels quantitativer bzw. qualitativer Nutzenkalküle zu bestimmen. Damit ging freilich die Gefahr einher, die moralische Bedeutung von Personengrenzen[331] nicht ernst genug zu nehmen und dem vermeintlich besseren Gesamtzustand eines aggregierten Nutzens für alle, die von den Folgen einer Handlung betroffen sind, grundlegende Rechte einzelner Individuen zu opfern.[332] Um diese grundsätzliche Schwäche des klassischen Utilitarismus zu überwinden, wurden bis in die Gegenwart neue und verschiedene handlungs- und regelutilitaristische Denkmodelle entwickelt.[333] Ohne die facettenreiche Entwicklung des Wohltätigkeitsbegriffs hier detailliert rekonstruieren zu wollen, seien lediglich drei systematische Anforderungen formuliert, die für einen ethisch überzeugenden Rückgriff auf diese Kategorie unverzichtbar erscheinen.

---

[331] Während Personen als Rechtssubjekte gerade in ihrer Individualität zu schützen sind, neigte der klassische Utilitarismus dazu, die Grenzen zwischen einzelnen Personen moralisch zu entwerten und kollektivistische Zielvorstellungen wie zum Beispiel „den größten Nutzen der größten Zahl“ zu propagieren.
[332] Vgl. Rawls 1972, 3-8.
[333] Siehe hier Smart 1961; Smart/Williams 1973; Brandt 1992; Gesang 2003 sowie Gesang 2013 und Gesang 2000.

Erstens ist deutlich zu machen, dass der Rückgriff auf den recht verstandenen Begriff der Wohltätigkeit einer medizinischen Handlungsweise keineswegs auf einen offen oder verdeckt utilitaristischen Standpunkt mit seinen gravierenden handlungstheoretischen, epistemologischen und rechtsphilosophischen Folgeproblemen hinausläuft.[334] Zweitens ist der Zielbezug der Wohltätigkeit auf die normative Leitvorstellung eines gelingenden Menschseins in einer Art und Weise offenzulegen, die aufgrund ihres abstrakten Charakters von partikularen weltanschaulichen Voraussetzungen sowohl des Handelnden als auch des Handlungsadressaten möglichst absieht. Dies kann zum Beispiel dadurch geschehen, dass die Bedeutung der jeweils wohltätigen Handlung für die Entfaltung menschlicher Grundbefähigungen nachgewiesen wird, ohne die eine gedeihliche Entwicklung des Einzelnen nicht vorstellbar erscheint.[335] Drittens ist die durch ihre Ausrichtung primär am Wohlergehen des Einzelnen charakterisierte Wohltätigkeit durch inhaltliche und prozedurale Regeln in einer Weise zu konkretisieren, die einen übertriebenen Altruismus ebenso ausschließt wie eine Verrechenbarkeit elementarer Grundrechte des Einzelnen.[336]

Für das Thema Big Data und Gesundheit dürften vor allem zwei Aspekte von Wohltätigkeit von besonderem Interesse sein: zum einen der Wissens- und Erkenntniszuwachs und zum anderen der therapeutische Mehrwert, der aus neuen Möglichkeiten der digitalen Informationsgewinnung und -verarbeitung großer Datenmengen im Gesundheitsbereich für unterschiedliche Beteiligte resultiert. Auch wenn es für moderne Wissens- und Informationsgesellschaften eigentlich selbstverständlich sein sollte, der kontinuierlichen Ausweitung des jeweiligen Erkenntnisstandes einen intrinsischen Wert beizumessen, bedarf diese Aussage doch insofern einer wichtigen Differenzierung, als hier zwei Perspektiven zu unterscheiden sind: Im Blick auf den Einzelnen ist zunächst festzustellen, dass Wissen und Erkenntnis von großer Bedeutung für die Selbstkonstitution des Individuums und seine Befähigung zu einer autonomen Lebensführung sind. Ohne ein zumindest in den Grundzügen realitätsnahes Selbst- und Weltverständnis ist weder eine gelungene Identitätsbildung noch eine autonom-verantwortliche Orientierung des eigenen Handelns möglich. Darüber hinaus kommt zweitens der kritischen Überprüfung, der Sicherung und der Ausweitung von Wissensbeständen auch eine wichtige gesamtgesellschaftliche Funktion zu. Funktional differenzierte, arbeitsteilige Gesellschaften sind in ihrer

---

[334] Der Begriff der Wohltätigkeit ist aufgrund seiner unaufgebbaren Bindung an individuelle Personenrechte keineswegs bedeutungsgleich mit demjenigen des Nutzens, vielmehr stellt die für den Utilitarismus zentrale Kategorie des Nutzens eine ganz bestimmte, ethisch hochgradig umstrittene Interpretation von Wohltätigkeit dar, die nicht einfach mit dem *benefit* einer Handlungsweise identifiziert werden darf (vgl. Beauchamp/Childress 2001, 166 f.).

[335] Eine solche an den Grundfähigkeiten orientierte (schwach essenzialistische) Bestimmung des „guten Lebens" hat den Vorteil, dass sie mit einer Vielzahl von Glücksvorstellungen kompatibel ist, wie sie in modernen freiheitlichen und pluralistischen Gesellschaften anzutreffen sind. Chancen und Risiken technologischer Innovationen können aus ethischer Perspektive entsprechend erst dadurch zum Gegenstand normativer Weisungen werden, dass ihr jeweiliger Bezug auf den Schaden bzw. den positiven Wert für die Beförderung des Ziels umfassender menschlicher Entfaltung für die betroffenen Individuen überzeugend nachgewiesen wird.

[336] Vgl. dazu zum Beispiel die Liste der „rules of beneficence" bei Beauchamp/Childress 2001, 167.

Entwicklung darauf angewiesen, dass die gestiegene Kontingenz der sozialen Beziehungen durch stabile wechselseitige Verhaltenserwartungen kompensiert wird. Die transparente Kommunikation gesicherter Erkenntnisse und die methodisch kontrollierte Erweiterung von Wissensbeständen sind dabei ein wichtiger Faktor zur Sicherung von Verlässlichkeit und sozialer Kohärenz.

Gegenteilige Effekte sind dort zu erwarten, wo durch bewusste Täuschungsmanöver (zum Beispiel im Kontext ideologisch motivierter Propaganda, gezielter Datenfälschung oder der Streuung von Falschmeldungen) das Vertrauen in die Zuverlässigkeit des anderen und seiner Mitteilungen schwindet und der soziale Zusammenhalt damit sukzessive erodiert. Um derartige Fehlentwicklungen korrigieren zu können, bedarf es des Schutzes einer der Wahrhaftigkeit verpflichteten Kommunikation, zu deren Sicherung sich insbesondere auf dem Feld der Wissenschaften differenzierte methodologische und wissenschaftstheoretische Maßgaben entwickelt haben. Aus ethischer Perspektive ist daher nicht nur zu fragen, ob und inwieweit die neuen digitalen Verfahren der Datensammlung und -auswertung mit wissenschaftsphilosophisch relevanten Verschiebungen epistemischer Standards oder Einbußen der Zuverlässigkeit der generierten Aussagen verbunden sind (siehe Abschnitt 2.2). Vielmehr ist auch zu klären, welchen Personengruppen die dabei tatsächlich erzielten Erkenntnisfortschritte jeweils primär zugute kommen, wie sich derzeit bestehende Hindernisse auf dem Wege einer effizienteren Gestaltung des Datennutzungsprozesses (zum Beispiel durch Open-Data- bzw. Open-Science-Strategien) beseitigen lassen und eine gerechte Verteilung jener positiven Effekte erreicht werden kann, die aus den zu erwartenden Wissenzuwächsen resultieren (vgl. dazu Abschnitt 4.5).

Die generelle Einschätzung, der zufolge Daten der wichtigste Rohstoff des 21. Jahrhunderts sind, dürfte jedenfalls auch und gerade im Gesundheitsbereich zutreffen. Bedenkt man, dass bislang nur für einen Teil medizinischen Handelns evidenzbasierte Maßstäbe hoher Qualität existieren und gesundheitsbezogene Entscheidungen folglich noch immer häufig unter Bedingungen hoher Unsicherheit getroffen werden müssen, ist es nicht überraschend, dass die Aussicht auf Erkenntnisgewinn und Wissenszuwachs durch die gezielte Sammlung und Interpretation großer Datenmengen eine der wichtigsten Triebfedern der dynamischen Entwicklung gesundheitsbezogener Dienstleistungen ist.

Zu bedenken ist dabei allerdings nicht nur, dass die Ansprüche an Datenqualität und Datensicherheit zur Fundierung der jeweiligen Erkenntnisgrundlagen auf den verschiedenen Feldern gesundheitsbezogenen Handelns sehr verschieden ausfallen können. Auch die Erwartungen an die dadurch ermöglichten Chancen können je nach Akteursgruppe erheblich variieren: Während die medizinische Grundlagenforschung vor allem auf ein verbessertes Verständnis bestimmter basaler gesundheitsrelevanter Wirkungsabläufe abzielt, ist der behandelnde Arzt primär an einer Optimierung des für die Patientenversorgung relevanten Zusammenhangs von

Diagnostik und Therapie interessiert. Der einzelne Patient wird den Vorteil hingegen primär an der tatsächlichen Verbesserung seines Wohlbefindens bemessen. Wieder eine andere Perspektive auf die möglichen Chancen dürfte bei denjenigen Gesunden anzutreffen sein, die zum Zweck individueller Prävention gesundheitsrelevante Informationen sammeln und auswerten. Dabei mögen sie sich diesen neuen Möglichkeiten durchaus auch im Hinblick auf einen überindividuellen, gesamtgesellschaftlichen oder einem Versicherungskollektiv zugute kommenden Nutzen annähern; sie mögen aber auch lediglich eigene ökonomische Zwecke verfolgen. Da gegenwärtig noch keine realistische Gesamtbilanzierung des umfassenden, sektorenübergreifenden Vorteils der Sammlung und Auswertung größerer Mengen gesundheitsbezogener Daten möglich ist, müssen die jeweiligen Chancen und Risiken solcher Maßnahmen für einzelne Segmente medizinischer Leistungen (wie zum Beispiel Diagnostik, Prädiktion, Therapieplanung und -durchführung), unterschiedliche Handlungskontexte und Betroffengruppen differenziert bestimmt werden.

Mithilfe der digitalen Datenverarbeitung können nicht nur die Früherkennung, die Diagnosestellung und die Therapieempfehlungen deutlich verbessert werden. Es entstehen vielmehr auch neue Möglichkeiten für die Steigerung der Lebensqualität von Patienten. So ist es etwa durch die Entwicklung neuartiger Sensoren möglich, die Therapie genauer an die Bedürfnisse des einzelnen Patienten anzupassen. Da der Arzt-Patienten-Kontakt zudem oftmals durch räumliche Distanzen erschwert ist oder erforderliche Spezialisten nur schwer erreichbar sind, kann der Einsatz der Telemedizin vor allem in ländlichen Gebieten dazu beitragen, eine schnellere ärztliche Konsultation zu ermöglichen und die generelle Kommunikation zwischen Ärzten und Patienten ohne langwierige und aufwendige Arztbesuche nachhaltig zu verbessern.

Auch im Vorfeld der Manifestation einer Erkrankung besitzt die Sammlung und Auswertung gesundheitsbezogener Daten ein erhebliches präventives Potenzial. Für Träger bestimmter Anlagen können nicht nur gruppenspezifische Risiken besser vermieden werden; vielmehr lässt sich auch die Einübung eines gesundheitsfördernden Lebensstils etwa durch die Überwachung von Körperfunktionen, Bewegungsformen oder eine Optimierung der Ernährungsgewohnheiten mit der Hilfe maßgeschneiderter Wearables fördern. Die rasante Entwicklung gerade in diesem Bereich weckt derzeit große Hoffnungen auf positive Impulse für eine deutliche Verbesserung der Gesundheitsversorgung.

Diese und ähnliche Anwendungen zeigen, dass sich die zu erwartenden Potenziale von Big-Data-Anwendungen im Gesundheitsbereich nicht darin erschöpfen, rascher neue Erkenntnisse zu generieren. Vielmehr hat der komplexe Begriff des *benefit* neben seiner rein epistemischen auch eine ökonomische und soziale Dimension, sodass sich je nach Art und Motivlage des jeweiligen Akteurs aus der Sammlung und Interpretation von Daten ganz unterschiedliche Ar-

rangements von Vorteilen ergeben können: Während der einzelne Bürger zum Beispiel bestimmte persönliche Gesundheitsdaten vor allem deswegen sammeln kann, um auf dieser Grundlage sein individuelles Präventionsverhalten zu optimieren, dienen die ungleich größeren Datenmengen bestimmter Patientenkollektive im Rahmen von Grundlagen- und klinischer Forschung primär dazu, das Verständnis von Entstehung und Entwicklung pathologischer Prozesse zu verbessern und die so gewonnenen Erkenntnisse für das diagnostische und therapeutische Handeln von Ärzten fruchtbar zu machen. Wichtige Akteure großer Datensammlungen sind inzwischen aber auch private Firmen, die das auf diesem Wege gewonnene Wissen für kommerzielle Produktentwicklungen – weit über den pharmazeutischen Bereich hinaus – nutzbar machen wollen. Auch Versicherungen dürften ein zunehmendes Interesse daran haben, über die Etablierung großer Sammlungen gesundheitsbezogener Daten das Risikoprofil ihrer Kunden besser abschätzen zu können und über gezielte finanzielle Anreize Einfluss auf deren Präventionsverhalten zu nehmen.

Schließlich sind auch staatliche Organisationen in mehrfacher Hinsicht in dieses facettenreiche Handlungsfeld involviert. Nicht nur können sie als Geldgeber wichtige forschungspolitische Akzente setzen; sie tragen auch die Verantwortung für die politischen Rahmenbedingungen unseres Gesundheitssystems und die Begrenzung der dafür anfallenden Kosten.

4.5 Gerechtigkeit

Die Sammlung und Weitergabe großer Mengen gesundheitsbezogener Daten berührt insofern auch grundlegende Fragen der Gerechtigkeit, als es hier einerseits um die Verteilung der jeweils zu erwartenden Chancen und Belastungen für unterschiedliche Personengruppen und andererseits um die Regulierung des Zugangs zu bestehenden Datensammlungen geht. Beides birgt vor dem Hintergrund der Überschneidung von öffentlich geförderten und privatwirtschaftlich organisierten Datenerhebungen ein erhebliches Konfliktpotenzial.

Konzepte der Gerechtigkeit dienen seit jeher dazu, die Grundkoordinaten eines moralisch und rechtlich begründeten Beziehungsgefüges zu bestimmen. Bereits in der antiken Philosophie vollzieht sich eine Entwicklung, die die Kategorie der Gerechtigkeit nicht nur pauschal zur Bezeichnung umfassender Tugendhaftigkeit verwendet, sondern durch verschiedene Binnendifferenzierungen mehrere Dimensionen der Gerechtigkeit terminologisch gegeneinander abgrenzt: So unterscheidet schon Aristoteles die sogenannte allgemeine Gerechtigkeit (*iustitia generalis*) im Sinne der Gesetzeskonformität des Handelns von verschiedenen Formen spezieller Gerechtigkeit (*iustitia specialis*) wie der austeilenden (*iustitia distributiva*) und der ausgleichenden Gerechtigkeit (*iustitia commutativa*). Die austeilende Gerechtigkeit, die Aristoteles zufolge die Beziehungen des Staates zu den verschiedenen das Gemeinwesen tragenden Bevölkerungsgruppen regelt und kriteriologisch auf der jeweils erbrachten Leistung bzw. Würdigkeit (*axia*)

der jeweiligen Akteure beruht, gibt dabei dem Gedanken der sogenannten Leistungsgerechtigkeit Raum. Demgegenüber bezieht sich die ausgleichende Gerechtigkeit auf das Recht im Austausch der Bürger untereinander und regelt zum einen im Sinne der Bedarfsgerechtigkeit die dem Bedürfnis nach ausreichender Güterversorgung entspringenden ökonomischen Tauschbeziehungen und zum anderen den im Falle einer unzulässigen Übervorteilung eines Akteurs erforderlichen Ausgleich durch rechtliche Kompensation.

Im Zuge dieses begrifflichen Differenzierungsprozesses sind zwei Einsichten wirkungsgeschichtlich besonders wichtig geworden: zum einen der Gedanke, dass der schon früh erkannte Bezug der Gerechtigkeit auf den anderen (*pros heteron, ad alterum*) nicht nur die Unterscheidung unterschiedlicher sozialer Rollen (etwa als Bürger und Rechtssubjekt oder als privatwirtschaftlicher Vertragspartner) und damit korrespondierender Beziehungsarten unumgänglich macht, sondern es damit auch ermöglicht, den Anspruch der Gerechtigkeit auf weite Bereiche des gesellschaftlichen Miteinanders auszudehnen. Die bis ins 19. Jahrhundert zurückreichende moderne Idee einer umfassenden sozialen Gerechtigkeit als Inbegriff einer gesamtgesellschaftlichen Wohlordnung, in die alle zuvor unterschiedenen Einzeldimensionen der Gerechtigkeit integriert werden, ist nur der Endpunkt einer langen und für das zeitgenössische politische Bewusstsein folgenreichen Entwicklung. Neben dem Handeln individueller Akteure unterstellt sie auch die institutionelle Grundstruktur eines Gemeinwesens dem Anspruch der Gerechtigkeit.[337] Zum anderen ist bereits in der antiken Gerechtigkeitsdiskussion deutlich geworden, dass die elementare Grundforderung der Gerechtigkeit, jedem das Seine zukommen zu lassen (*suum cuique*), die Ausarbeitung einer differenzierten Kriteriologie verlangt, in der neben der Eigenart der jeweils betroffenen Güter auch die Funktion und der Zweck der jeweiligen Beziehungsform angemessen berücksichtigt werden.[338] In der neueren Gerechtigkeitsdiskussion hat vor allem die sozialpolitisch besonders bedeutsame Verteilungsgerechtigkeit eine wichtige konzeptuelle Fortentwicklung erfahren, die über die traditionelle Gegenüberstellung von Leistungs- und Bedarfsgerechtigkeit hinausführt. Während lange Zeit die Vorstellung herrschte, ein der Gerechtigkeit wegen gebotene soziale Ausgleich sei ohne Weiteres durch bestimmte finanzielle Transfers zu leisten, haben Vertreter eines sogenannten Befähigungsansatzes[339] zu Recht darauf hingewiesen, dass Menschen in unterschiedlichen sozialen Lebenslagen keineswegs denselben Gebrauch von bestimmten Gütern machen können, sodass zunächst die begrifflichen Grundlagen des Gerechtigkeitsdiskurses erneut zu klären seien. Vor diesem Hintergrund wird vor allem die Debatte um die aktuellen Spielarten einer sogenannten Teilhabe-[340] sowie einer Befähigungsgerechtigkeit[341] konstruktiv weitergeführt. Teilhabegesichtspunkte spielen auch in der aktuellen

[337] Vgl. dazu vor allem Rawls 1972.
[338] Vgl. Walzer 1983.
[339] Vgl. Sen 1979; Sen 2009; Daniels 1990; Nussbaum 1998; Nussbaum 2014 sowie Nussbaum 2015.
[340] Vgl. Bormann 2006.
[341] Vgl. Dabrock 2012.

Diskussion über die politische Gerechtigkeit eine zentrale Rolle. Hier geht es nicht mehr nur um die Legitimität bestimmter Herrschaftsformen im Allgemeinen, sondern auch um die spezifischen Chancen des Zugangs zu und der Beteiligung an solchen politischen Prozessen, in denen die Regeln des gesellschaftlichen Miteinanders für jedermann ausgehandelt und definiert werden[342], insbesondere für bestimmte besonders vulnerable Personengruppen.

So facettenreich sich der Gerechtigkeitsbegriff aus einer ideengeschichtlichen Perspektive auch ausnimmt, die Grundstruktur seines Sinngehaltes lässt sich doch wie folgt bestimmen: Als normierendes Prinzip sozialer Beziehungen gebietet es die Gerechtigkeit, willkürliche Privilegierungen Einzelner oder bestimmter Gruppen dadurch zu überwinden, dass das jedem jeweils Angemessene auf rationale Weise bestimmt wird, im Handeln der anderen gleichmäßige Berücksichtigung erfährt und dass Unterschiede in seiner Behandlung einer konsensfähigen Begründung bedürfen.

Für den spezifischen Gegenstandsbereich der Big-Data-Anwendungen im Gesundheitsbereich sind aus gerechtigkeitsethischer Perspektive vor allem die folgenden vier Problemfelder besonders wichtig:

Erstens ist die Frage der Regelung von Zugangsbedingungen zu Datensammlungen für den Forschungsbereich bedeutsam. Sowohl auf nationaler wie auf internationaler Ebene gibt es verschiedene Hürden dafür, dass bereits etablierte Daten innerhalb der Wissenschaftsgemeinschaft zeitnah ausgetauscht und wechselseitig zur Verfügung gestellt werden. Nicht nur Kliniken und private Firmen, sondern auch einzelne Forschungsteams neigen dazu, die von ihnen erhobenen Daten als ihr Eigentum zu betrachten, selbst wenn für die Erhebung und Auswertung dieser Daten öffentliche Forschungsgelder oder die öffentliche Infrastruktur in Anspruch genommen wurden. Ein solcher mit Blick auf soziale und politische Gerechtigkeit problematischer Entzug von Daten aus der öffentlichen Sphäre entsteht auch dann, wenn private Anbieter von Gesundheits-Apps gesammelte Daten horten, obwohl die Entwicklung und Nutzung der Technologie öffentlich gefördert worden ist (siehe Kapitel 2).

Es ist unstrittig, dass bestimmte Daten – kontextabhängig – nicht allgemein für jedermann zugänglich sein dürfen, sondern durch geeignete Schutzmaßnahmen dem willkürlichen Zugriff Dritter entzogen bleiben müssen. Ebenso verständlich ist das Interesse privatwirtschaftlicher Initiativen, über Datensammlung und -nutzung Gewinne zu erzielen. Letzteres muss allerdings in Abwägung mit berechtigten Interessen erfolgen, die in einer gegebenenfalls vorangegangenen öffentlichen Förderung gründen. Dabei dürfte es schwierig bzw. unmöglich sein, den klassischen Eigentumsbegriff im Sinne exklusiver Verfügungsmacht auf das spezifische Feld solcher

[342] Vgl. Forst 1996 sowie Forst 2007.

gesundheitsbezogenen Daten, die ausdrücklich für bestimmte Forschungszwecke erhoben worden sind, anzuwenden. Unbegründete Zugangsbarrieren verursachen nicht nur vermeidbare zusätzliche Kosten durch unnötige Mehrfacherhebungen derselben Datenarten, sondern können auch dazu führen, dass das Potenzial schon vorhandener Daten nicht optimal genutzt wird oder öffentlich finanzierte Datenerhebungen rein privatwirtschaftlicher Gewinnerzielung dienen. Der Gerechtigkeit wegen erscheint es daher geboten, mittels geeigneter Instrumente einen möglichst ungehinderten und zeitnahen Zugang berechtigter Personen zu bereits vorhandenen forschungsrelevanten Datensammlungen sicherzustellen. Hier böten sich Instrumente wie Open-Access- bzw. Public-Science-Strategien oder spezifische Publikationspflichten etc. an.

Zweitens liegt eine ebenfalls die Teilhabegerechtigkeit berührende Problematik dort vor, wo sich schleichend monopolartige Strukturen etablieren, die es Forschern erschweren bzw. sogar gänzlich unmöglich machen, ihre Ziele zu verfolgen oder einzelnen Individuen oder Gruppen daran hindern, über die Auswirkungen der Nutzung ihrer Daten auf sie selbst souverän zu entscheiden (siehe Abschnitt 4.3). Herausforderungen für die Teilhabegerechtigkeit könnten auch mit Blick auf (die in Abschnitt 4.3 genannten) subtileren Formen der Machtausübung bestehen. Dies wäre etwa der Fall, wenn Datenmonopole intransparent genutzt würden, um einzelne Nutzer von vornherein von bestimmten Angeboten auszuschließen.

Drittens entstehen auch auf dem besonders rasant wachsenden Markt von Gesundheits-Apps und verschiedenen, der privaten Selbstvermessung dienenden Wearables insofern neue Gerechtigkeitsfragen, als diese Produkte zumeist auf die Verstärkung eines gesundheitsförderlichen Lebensstils ausgerichtet sind. Im Hinblick auf individuelles Präventionsverhalten einzelner gesundheitsbewusster Versicherter ist es für eine gerechte Gestaltung von Krankenversicherungstarifen daher erforderlich, den Gedanken der Leistungsgerechtigkeit mit den Bedingungen eines solidargemeinschaftlich finanzierten Gesundheitssystems zum Ausgleich zu bringen. Die möglichen positiven Effekte dieser Entwicklung sowohl für die Gesundheit der betroffenen Einzelnen als auch für die Kostenersparnis des Gesundheitssystems sollen nicht in Abrede gestellt, sondern durchaus betont werden. Skepsis ist aber angebracht gegenüber einem sich damit gelegentlich verbindenden Verständnis einer atomisierten Zuschreibung individueller Verantwortung, mit dem die Leistungsfähigkeit Einzelner oder bestimmter Gruppen gänzlich abstrakt definiert und von ihren sozialen, ökonomischen und kulturellen Lebensbedingungen isoliert wird.

Gewiss erfordert ein aktives und reflektiertes Präventionsverhalten immer auch ein Leistungselement in Form eigenen Engagements und individueller Anstrengung; doch darf nicht übersehen werden, dass die Fähigkeit, sie zu erbringen, stets auch von bestimmten Voraussetzungen abhängt, die teilweise außerhalb der Verfügungsmacht des Einzelnen liegen. Vor allem dann,

wenn die Nutzung solcher Präventionsmaßnahmen im Rahmen bestimmter Krankenversicherungstarife leistungsgerecht honoriert werden soll, ist aus Gründen der Chancengerechtigkeit darauf zu achten, dass auch alle Betroffenen nicht nur eine abstrakt-theoretische, sondern die real-praktische Gelegenheit erhalten, sich im Rahmen ihrer individuellen Lebensplanung bewusst für oder gegen die Nutzung entsprechender Angebote zu entscheiden.

An dieser Stelle verschränken sich Gerechtigkeits- mit Solidaritätsüberlegungen, die in Abschnitt 4.6 genauer ausgeführt werden. Bei einem allzu engen Solidaritätsverständnis könnten Diskussionen entstehen, ob in einem solidarisch finanzierten Versorgungssystem der Leistungsfähigen Ansprüche von Personen, die individuell nicht in der Lage sind, eigene präventive Leistungen zu erbringen, beschränkt oder gar ausgeschlossen werden sollten. Praktische Beispiele gibt es hier bereits in privaten Krankenversicherungen. Sowohl Befähigungs- und Chancengerechtigkeit als auch ein inklusives Verständnis von Solidarität verlangen hingegen, dass einem Einsatz von Gesundheitsdaten Grenzen gesetzt werden, soll die grundsätzliche solidarische Zugangs- und Finanzierungsstruktur zumindest der gesetzlichen Krankenversicherung erhalten bleiben (siehe Abschnitt 4.6).

Viertens haben schließlich auch die Motive der gesellschaftlichen Inklusion und einer angemessenen Berücksichtigung der vielfältigen sozialen Bedingtheit individuellen Freiheitsgebrauchs in die integrative Vorstellung einer Befähigungsgerechtigkeit Eingang gefunden, die generell dem Schutz all jener Güter und Rechte des Individuums dient, die die Bedingung der Möglichkeit persönlicher Handlungsfähigkeit ausmachen. Zu diesen Bedingungen gehört in einer zunehmend digitalisierten Umwelt auch die Entwicklung der durchaus anspruchsvollen Fähigkeit, verantwortlich mit eigenen (und fremden) gesundheitsbezogenen Daten umzugehen. Je nach Handlungskontext und funktionaler Rolle des jeweiligen Akteurs sind dazu unterschiedliche Aspekte zu beachten: Während aus Sicht des Patienten vor allem eine möglichst umfassende Aufklärung über Art, Umfang und Zweck der Datensammlung im Blick auf die Gestaltung von Diagnose und Therapie erforderlich erscheint, um eine wirklich informierte Zustimmung zu ermöglichen, dürfte der private Benutzer von digitalen Dienstleitungen zur Optimierung seines Gesundheitsverhaltens primär daran interessiert sein, den Prozess der Weitergabe seiner Daten aktiv mitgestalten und die Qualität der gelieferten Dateninterpretation realistisch beurteilen zu können. Auch einem am Ideal der Befähigungsgerechtigkeit orientierten Ansatz zufolge muss der jeweils betroffene Einzelne nicht alle technischen Details moderner Informationserfassungs- und -verarbeitungssysteme genau verstehen. Doch gilt es, jene konstitutive Fähigkeit zur informierten Urteilsbildung zu stärken, die für eine qualifizierte Zustimmung bzw. ein begründetes Verbot zur Datenerhebung, -auswertung, -aufbewahrung und -weitergabe notwendig ist. Insofern verweist der Begriff der Befähigungsgerechtigkeit auf ein Verständnis von Datensouveränität, das in Kapitel 5 dieser Stellungnahme näher entfaltet wird.

4.6 Solidarität

Der Begriff der Solidarität hat eine heterogene Begriffsgeschichte und wird heute in verschiedenen Fachdisziplinen sehr unterschiedlich verstanden; daher gibt es eine ganze Reihe parallel vertretener Solidaritätsbegriffe.[343] Doch lassen sich einige wesentliche Kernelemente ausmachen, die jedenfalls den meisten Solidaritätsverständnissen eigen sind. Solidarität bezeichnet demnach, grob gesprochen, prosoziale Handlungen, Praktiken und Dispositionen sowie institutionelle, politische und vertragliche Regelungen, die dazu dienen sollen, „andere zu unterstützen, oder zumindest […] eine Neigung auszudrücken, helfen und unterstützen zu wollen".[344] Dabei sind sich die meisten Autoren zudem einig, dass zur prosozialen Neigung oder Haltung entsprechende Handlungen oder die Übernahme von Kosten[345] hinzutreten müssen, damit von Solidarität die Rede sein kann.

Solidarität wird vielfach als komplementär – und oft auch subsidiär – zur Gerechtigkeit aufgefasst. Dies gilt im ideellen wie institutionellen Sinn. Während Gerechtigkeitsverständnisse vielen rechtsstaatlich kodifizierten Regelungen unterliegen, werden Pflichten zur Solidarität meist der Ebene der Sozialmoral zugewiesen. Die verschiedenen, oben skizzierten Gerechtigkeitsverständnisse beziehen sich dezidiert auf alle bzw. jedermann; auf der institutionellen Ebene liegt ihnen daher die staatliche Neutralität in der Bewertung von individuellen Lebensplänen, Werten, Zielvorstellungen etc. zugrunde. Genau diese Konzepte, einschließlich der mit ihnen verbundenen Bewertungen, sind es allerdings, woran die solidarische Praxis und entsprechende solidarische Unterstützungsgebote und -pflichten anknüpfen. Die zugrunde liegende Motivation für solidarische Hilfshandlungen und Kostenübernahme basiert regelmäßig „auf dem Erkennen von relevanten Gemeinsamkeiten mit einer anderen Person, oder anderen Personen".[346] Solche Gemeinsamkeiten bestehen etwa in der Wahrnehmung von Gefährdungen, existenziell bedeutsamen Risiken etc., deren kompetente Abwehr die Leistungsfähigkeit des Einzelnen überfordert. Solidarität entsteht also aus einem gemeinsamen Ziel einer solidarischen Gruppe, angesichts einer gemeinsamen Herausforderung oder Bedrohung oder aber auch aus der geteilten Vorstellung vom guten Leben in einer Solidargemeinschaft, wie sie etwa im deutschen Sozialstaat zum Ausdruck kommt. Eine Solidargemeinschaft teilt wesentliche Auffassungen davon, warum und welche gegenseitigen Hilfsangebote für die Mitglieder zur Verfügung stehen sollen; ihre Mitglieder sind deswegen bereit, Beiträge zu leisten, um die gemeinsame Vorstellung zu verwirklichen.

---

[343] Ausführliche Übersichten über Begriffsgenese und verschiedene resultierende Solidaritätsbegriffe, siehe Prainsack/Buyx 2017.
[344] Bayertz 1996, 308.
[345] „Kosten" werden hier in einem weiten Sinne verstanden (unter anderem finanziell, sozial, emotional, zeitlich etc.).
[346] Prainsack/Buyx 2016, 82. Aufgrund dieses Erkennens von Gemeinsamkeiten sind solidarische Beziehungen in relevanter Hinsicht symmetrisch und es zeichnet sie, insbesondere wenn sie institutionalisiert werden, ein Element der zumeist indirekten Gegenseitigkeit aus (ebd.).

Solidarität ist also nicht neutral, sondern immer auf ein bestimmtes, substanzielles Ziel bzw. eine geteilte Auffassung davon, wie eine Gemeinschaft verfasst sein sollte, ausgerichtet. Anders formuliert, könnte man sie als zwischen dem Gerechten und dem Guten stehend beschreiben. Solidarität ist ferner nicht immer automatisch im normativen Sinne gut; eine solche Einordnung hängt von der Bewertung der jeweiligen solidaritätsstiftenden Gemeinsamkeiten und Ziele ab.[347] Sie ist partikular; der jeweilige Gehalt und das jeweilige Ziel von Solidarität müssen in jedem Kontext neu bestimmt werden. Dies gilt für die konkrete Solidarpraxis in umschriebenen sozialen Gruppen wie etwa Nachbarschaften, Vereinen, Selbsthilfegruppen etc. ebenso wie für die große Solidargemeinschaft der gesetzlichen Krankenversicherung (§ 1 SGB V). Dabei können solche Zielbestimmungen, zumal auf den höheren Ebenen der Institutionalisierung, durchaus abstrakt sein; in der GKV etwa „die Aufgabe, die Gesundheit der Versicherten zu erhalten, wiederherzustellen oder ihren Gesundheitszustand zu bessern" (ebd.). Die Solidargemeinschaft der GKV etabliert über das Bedarfsprinzip auf der einen und das Prinzip individueller finanzieller Leistungsfähigkeit auf der anderen Seite, dass Erkrankungsrisiken von allen Mitgliedern – also von allen Versicherten – gemeinsam getragen werden; man könnte dies als ein fallgruppenübergreifendes solidarisches Eintreten der jeweils aktuell Starken für die jeweils aktuell Schwächeren beschreiben.

Die Bereitschaft zur Solidarität ist, wie es die Logik der Sozialversicherungen eindrücklich vor Augen führt, nicht selten abhängig von Reziprozitätserwartungen. Das bedeutet nicht, dass Solidarität übende Menschen ihre Unterstützung anderer unmittelbar an eine direkte Gegenleistung knüpfen. Dies wäre gerade keine Solidarität, sondern das Geschäft von wechselseitiger Leistung und Gegenleistung. Der Zusammenhang von Solidarität und Reziprozität ist grundsätzlicher: Er besteht in der Erwartung, dann von anderen her Solidarität zu erfahren, wenn man selbst in einem analogen Schadensfall (Krankheit usw.) fremder Unterstützung bedarf. Das ist das Prinzip der Krankenversicherung, der Gesetzlichen Unfallversicherung, aber auch der Hausrats-, Rechtshilfe- oder Reiserücktrittsversicherung.

Die Gleichsetzung von Reziprozität und Solidarität ist allerdings unzulässig. Wenn Hilfe und Unterstützung ausschließlich oder primär geleistet werden, weil es die Erwartung einer (konkreten) Gegenleistung gibt, so wäre dies eben *keine* solidarische Praxis. Dennoch ist die Solidarität in der Praxis, und zumal die institutionalisierte und sanktionierte Solidarität, mit Reziprozitätselementen in wichtiger Hinsicht verknüpft. Wo es etwa rechtlich verfasste Solidaritätsregeln gibt, die Pflichten statuieren, wie es in der Krankenversicherung der Fall ist, hat eine (indirekte) Reziprozität eine wichtige solidaritätsstabilisierende Funktion. Bei allem gemeinsamen

[347] Es sind etwa zahlreiche Beispiele von strukturell solidarischen Gruppen in kriminellen Vereinigungen bekannt, zum Beispiel in der Mafia oder in terroristischen Vereinigungen.

Bekenntnis zur Notwendigkeit einer Krankenversorgung für alle würden die Kosten eines solchen umfassenden und unpersönlichen Solidaritätsgefüges von den Mitgliedern nicht getragen, wenn nicht gewährleistet wäre, dass jeder, der beiträgt, im eigenen Schadensfall ebenfalls Hilfe erwarten kann. Die Bereitschaft zur Solidarität kann also durchaus nachlassen, wenn auf Dauer der Eindruck entsteht, die Hilfs- und Unterstützungsbedürftigkeit werde vom anderen etwa durch fahrlässige Selbstschädigung oder mangelnde Eigeninitiative verursacht und das Solidaritätsgefüge damit überstrapaziert.

Aus dieser knappen Beschreibung ergibt sich bereits, dass Solidargemeinschaften, insbesondere jene, die gesetzliche oder vertragliche Rechtspflichten institutionalisieren, sowohl hinsichtlich der jeweiligen Zielbestimmung als auch bezüglich ihrer Mitgliedschaft kontinuierlich begründungspflichtig sind.[348] Während in Deutschland die Zielsetzung der grundsätzlichen gegenseitigen sozialen Sicherungsfunktion der GKV nur selten gänzlich infrage gestellt wird, sind die Bestimmung und die Begrenzung der Mitgliedschaft durchaus Gegenstand von Kontroversen, ebenso wie Art und Umfang der Leistungen, die durch Solidarbeiträge finanziert werden sollen. Lange bereits wird debattiert, ob die Ausgliederung von Bevölkerungsteilen durch Mitgliedschaft in der PKV das grundsätzliche, nämlich gesamtgesellschaftliche Solidarprinzip verletzt. Aktuell wird vor allem erörtert, ob und wie der volle Mitgliedstatus in der GKV neuen Bevölkerungsgruppen zugestanden werden kann. Solidargemeinschaften haben also eine gewisse Tendenz zur Inklusion lediglich derer, die bereits in sie einbezogen sind, und deshalb zur Exklusion anderer, denen man gegebenenfalls (bestimmte) Solidaransprüche verweigern will oder muss.

4.6.1 Solidarität in der gesetzlichen und privaten Krankenversicherung

Es stellt sich also die Frage, ob und inwiefern Big Data im Gesundheitssystem zu Veränderungen der bestehenden Solidaritätsmuster in der GKV und darüber hinaus führt bzw. führen könnte. Wie in Kapitel 2 dieser Stellungnahme dargelegt, ergeben sich aus der Auswertung von Lebensstil- und gesundheitsrelevanten Daten in großem Umfang neue Möglichkeiten der Risikostratifizierung. Big-Data-Prozeduren können sehr viel genauere Risikoprofile von Menschen liefern. Damit entwickelt sich ein Trend weiter, der mit klinischen Risiko-Scores (klassisch: kardiovaskuläre Risikofaktoren) begann und spätestens seit dem Aufkommen der ersten präzisen genetischen Tests (vor allem zu monogenetischen Erkrankungen wie der Huntington-Krankheit[349]) kontrovers diskutiert wird. Bereits damals wurde befürchtet, dass angesichts der neuen Möglichkeit, Gruppen mit evidenzbasiert niedrigen und solche mit hohen (oder, im Fall von

[348] Dazu bereits Rorty 1989, 189-191.
[349] Die Huntington-Krankheit (auch Chorea Huntington) ist eine sehr seltene, durch Genmutation verursachte, autosomal-dominant erbliche Erkrankung des Gehirns, bei der sich die Nukleinsäurenfolge Cytosin-Adenin-Guanin auf dem vierten Chromosom mehr als 36 mal wiederholt. Typische Symptome sind Bewegungsstörungen und Wesensänderungen bis hin zur Demenz.

monogenetischen autosomalen Erkrankungen, sicheren) Erkrankungsrisiken zu bestimmen, einige, nämlich Niedrigrisikogruppen, die Solidargemeinschaft verlassen bzw. ihre Solidarbeiträge aufkündigen wollen könnten. Jedenfalls erwartete man, dass andere, nämlich Hochrisikogruppen, diskriminiert würden.[350] Diese Befürchtungen haben sich wohl wegen der vielen anderen, noch nicht in dieser Weise explizierten Krankheitsrisiken sowie weitgehend effektiver Antidiskriminierungsgesetze nicht bzw. nur sehr selten bewahrheitet.[351]

Big Data hingegen bietet so viele weitere Möglichkeiten der Risikodifferenzierung, dass diese Sorge nun verstärkt in den Vordergrund tritt – insbesondere, wie oben erwähnt, wo es um die präventive Leistung bzw. Leistungsfähigkeit Einzelner geht, die über datenreiche Methodologien immer augenscheinlicher nachvollziehbar wird. Wenn von der Gesamtgenomsequenzierung, dem Mikrobiom und Proteom über sonstige gesundheitliche und Lebensstil-Daten bis hin zu solchen zu Freizeit- und Einkaufsverhalten, Sozialkontakten und Browserbenutzung alle Daten integriert auswertbar sind, so die berechtigte Annahme, werden Risikoprofile von ganz anderer Differenziertheit und vor allem ganz anderer, sehr viel höherer, Prädiktivität möglich.

Grundlage der Solidargemeinschaft in der GKV ist die geteilte Vulnerabilität *aller* gegenüber Krankheitsrisiken, die nicht sicher antizipierbar und quantifizierbar sind, und derentwegen die Mitglieder der solidarischen Pflicht unterliegen, über die kollektive Finanzierung eines gesetzlichen Fonds in allen Fällen der Erkrankung einzelner Mitglieder für die anfallenden Behandlungskosten gemeinschaftlich einzustehen. Falls in Zukunft, von Unfallrisiken (und einigen Infektionskrankheiten) abgesehen, Krankheitsrisiken so präzise, mehrdimensional und sicher bestimmt werden können, wie sich dies abzuzeichnen beginnt, dann könnte diese Grundlage in einer Art fragmentierender Logik exklusiver Solidaritäten infrage gestellt werden.

Es gibt bereits eine Reihe von Beispielen für die stärkere Stratifizierung von Versicherten im Ausland, und auch in Deutschland werden strukturelle Veränderungen des Versicherungssystems anhand der neuen Möglichkeiten der Risikostratifizierung regelmäßig diskutiert.[352] Zwar sind deren Nutzung etwa für den Leistungsausschluss in der GKV in Deutschland sehr enge Grenzen gesetzt – was anderswo nicht immer der Fall ist. Dennoch entwickeln sich akute Herausforderungen an die solidarische Verfasstheit auch des deutschen Gesundheitssystems.

Im dualen deutschen Krankenversicherungssystem mit seinem Nebeneinander von GKV und PKV birgt eine durch Big Data präzisierte und verfeinerte Auswahl potenzieller Versicherter zunächst ganz grundsätzlich die Gefahr einer verstärkten Selektion von Personen mit günsti-

[350] Vgl. Chadwick/Berg 2001.
[351] Ein weithin bekannter Präzedenzfall ist der Prozess um das Verbeamtungsverfahren einer von Chorea Huntington betroffenen Lehrerin aus dem Jahr 2004, die ihre Verbeamtung schließlich vor Gericht erstritt. Siehe das Urteil des Verwaltungsgerichts Darmstadt in NVwZ-RR 2006, 566.
[352] Vgl. Andelfinger 2016; Deutscher Bundestag 2014; Radic et al. 2016; Heuvel 2016.

gem/niedrigem Risikoprofil durch private Versicherer auf Kosten der gesetzlichen Solidarsysteme, denen dadurch Mehrbelastungen entstünden. Das dürfte die gesundheitspolitischen Debatten zur Berechtigung der PKV befördern. Jedenfalls wird kontrovers diskutiert, ob eine zunehmende Funktion der GKV als subsidiäres Auffangbecken der „schlechten" Risiken aus der PKV das grundlegende solidarische Gefüge innerhalb eines nationalen Gesundheitssystems infrage stellt. Umgekehrt wird auch die Auffassung vertreten, dass bei dieser Stratifizierung der Risiken ein umfassendes GKV-System neue Attraktivität gewinnt.

Herausforderungen für die Solidarität ergeben sich jedoch nicht nur im Verhältnis von GKV und PKV, sondern auch innerhalb der PKV selbst. Die PKV arbeitet mit risikoäquivalenten Prämien. Vor Abschluss der Versicherung wird anhand einer umfangreichen Datenerhebung samt angeforderten medizinischen Vorbefunden über das künftige Mitglied festgelegt, in welche Tarifgruppe es gehört. Könnte man in dem Zusammenhang auch Daten etwa über die sportlichen Aktivitäten und das Ernährungsverhalten einbeziehen, könnte das zu einer noch genaueren Risikoprofilbildung beitragen – so jedenfalls die Befürchtung. Denkbar wäre zudem die individuelle Anpassung der Prämie nicht nur nach der Entwicklung in der gesamten Tarifgruppe, sondern auf der Grundlage kontinuierlich erhobener individueller Daten nach Abschluss der Versicherung. Damit würde das Versicherungsprinzip, dass Risiken einer größeren Gruppe gemeinsam getragen werden und Tarife auch nicht individualisiert angepasst werden dürfen, aufgegeben. Es könnten zunehmend kleine Tarifgruppen entstehen, bei denen Schadensfälle dann umso schneller zu Beitragserhöhungen führen.

Für die PKV hat etwa die Generali-Gruppe 2014 erstmals angekündigt, in Europa gemeinsam mit dem südafrikanischen Finanzdienstleister Discovery mit dem Produkt Vitality ein „verhaltensbasiertes Versicherungsmodell [anzubieten], das Kunden zur Verbesserung der Gesundheit und des Wohlbefindens motivieren und entsprechend belohnen soll".[353] Es basiert laut eigener Pressemitteilung auf drei Grundsätzen: der Belohnung eines gesundheitsbewussten Verhaltens, regelmäßigen Kontakten mit der Versicherung und maßgeschneiderten Programmen zur Verbesserung der Lebensgewohnheiten, sowie einem innovativen Verhältnis zwischen Versicherer und Kunden. Aus der Pressemitteilung geht allerdings nicht hervor, dass die regelmäßigen Kontakte und maßgeschneiderten Programme mit einer Sammlung von Daten des Kunden verbunden sind, die er der Versicherung zur Verfügung stellt. Die Daten über sportliche Aktivitäten und Ernährungsverhalten stammen von Wearables oder Gesundheits-Apps auf dem Smartphone oder der Smartwatch; die Belohnung besteht unter anderem in einem Prämienvorteil für den Kunden.

[353] http://presseservice.pressrelations.de/standard/result_main.cfm?r=581548&sid=&aktion=jour_pm&print=1 [19.09.2017].

Damit stellt sich die Frage, was geschieht, wenn Versicherte nicht bereit sind, in einem verhaltensbasierten Versicherungsmodell mitzuwirken. Die nächstliegende Vermutung ist, dass ihnen die Vorteile günstigerer Prämien sowie Geschenke oder Gutscheine vorenthalten werden, was auf lange Sicht zu Prämiennachteilen führen muss. Unabhängig davon, ob sie sich gesundheitsförderlich verhalten oder nicht, werden sie dafür bestraft, dass sie ihre Daten nicht der Versicherung überlassen wollen. Die Ausübung ihres Rechts auf informationelle Selbstbestimmung würde also zu einem konkreten Nachteil führen. Darüber hinaus würden in solchen Modellen diejenigen von Vorteilen ausgeschlossen, die etwa durch Schicksalsschläge ohnehin schon Nachteile erleiden mussten und kein günstiges Risikoprofil aufweisen.

Schließlich und besonders gravierend wären die Auswirkungen verhaltensdatenbasierter Versicherungsmodelle in der GKV. Das Solidarsystem der GKV zeichnet sich dadurch aus, dass der Versicherte ganz unabhängig von den Risiken eine einkommensabhängige Prämie zahlt und nach den Maßgaben des SGB V die Leistungen erhält, die er benötigt. Datengestützte Risikoprofilbildung ist der GKV bei der Beitragsfestsetzung, von einigen finanziellen Boni für präventives Verhalten einmal abgesehen, weitgehend fremd. Sie widerspricht dem Solidargedanken, der ja, wie oben beschrieben, die Absicherung gegen krankheitsbedingte Vulnerabilität gerade ohne Ansicht individueller (Verhaltens-)Risiken fordert. Schon die intensive Diskussion um das mutwillige oder zumindest fahrlässige Eigenverschulden von Krankheiten (etwa Ski-Unfällen, Nikotin-assoziierten Erkrankungen etc.) hat gezeigt, wie komplex die Zurechnung von Verantwortung für bestimmte Krankheitszustände beim Individuum sein kann (Stichwort gesundheitliche Eigenverantwortung[355]).

Gesundheit wird durch viele Faktoren beeinträchtigt, die außerhalb der Kontrolle der einzelnen Person liegen. Schon heutzutage ist es nur selten möglich, ein klares, personal zurechenbares Eigenverschulden nachzuweisen, das als Verletzung der indirekten Reziprozitätserwartung in solidarischen Systemen qualifizierbar wäre und deshalb einen Leistungsausschluss oder andere Sanktionen rechtfertigen könnte.

Im Übrigen gilt auch unter dem Solidaritätsprinzip der GKV ein Vorrang der Freiheit zur Lebensgestaltung und Selbstentfaltung vor einer strikten und permanenten Pflicht zur Vermeidung aller Gesundheitsrisiken. Gewiss gilt dies nicht unbegrenzt; der Passus über die Mitverantwortung der Versicherten für ihre Gesundheit bringt das zum Ausdruck (§ 1 SGB V). Doch könnte die dauernde gezielte Sammlung von Daten über die individuelle Lebensführung und die Nutzung Big-Data-gespeister Risikoprofile, die alle Lebensbereiche umfassen, schwerlich als zumutbare Erwartung an die Mitverantwortung für die eigene Gesundheit qualifiziert werden.

[355] Vgl. Langanke et al. 2013.

Es steht daher zu überlegen, ob die Erstellung Big-Data-getriebener, hochprädiktiver Risikoprofile für die Einstufung von Patienten verboten werden sollte.

Eine seit Langem geführte Debatte spitzt sich also angesichts neuer Big-Data-Möglichkeiten erheblich zu: die um die gesundheitliche Eigenverantwortung der Versicherten und um die Frage, ob und wie diese bei der Fortentwicklung der Krankenversicherungssysteme unter dem unausweichlichen Zwang knapper Mittel berücksichtigt werden darf. Zu dieser Debatte gehört die Frage, ob es eine zentrale Aufgabe der GKV sein darf, das Gesundheitsverhalten ihrer Versicherten zu beeinflussen und gleichsam Gesundheitserziehung zu betreiben. Im Präventionsgesetz von 2015 wurden die gesetzlichen Krankenkassen dazu aufgefordert, „qualitätsgesicherte Angebote zur Förderung eines gesundheitsbewussten Verhaltens" einzurichten (Neufassung des § 65a SGB V). Nach Auffassung der Verbraucherzentrale Nordrhein-Westfalen sollte damit der Schwerpunkt der schon seit 2003 existierenden Bonusprogramme „von einem Kundenbindungsinstrument hin zu einem Steuerungsinstrument für das Verhalten der Versicherten" werden.[356] Hierzu gibt es bereits zahlreiche Beispiele, etwa das Sammeln von Bonuspunkten, die dann in einem Prämienshop eingelöst werden können. Belohnt wird unter anderem der Nachweis gesundheitsförderlicher körperlicher Bewegung, teils aber auch schon reine Zusicherungen wie etwa die, Nichtraucher zu sein. Damit beeinflussen Versicherungen in bestimmtem Sinn gezielt die Lebensführung ihrer Versicherten und legen ihnen hierfür über Belohnungen gewisse Standards nahe. Belohnungen implizieren Wertaussagen. Vor diesem Hintergrund ist gut zu überlegen, welchem Verhalten und welchen Eigenschaften warum welcher Wert zugeschrieben wird und was bestimmte Verhaltensweisen jeweils als belohnenswert erscheinen lässt. In den klassischen Debatten zur gesundheitlichen Eigenverantwortung ist längst herausgearbeitet worden, dass es sowohl praktisch als auch in normativer Hinsicht sehr problematisch ist, einzelne Verhaltensweisen oder auch komplexere Verhaltensmuster derart zu klassifizieren und zu vergleichen, also etwa die Bewegung unter dem Gesichtspunkt ihrer gesundheitsfördernden Effekte in einen wertenden Vergleich mit (zum Beispiel) so etwas wie sozialem Engagement zu bringen. Datenbasierte Anreizsysteme könnten eine sehr intensive und invasiv-überwachende Wirksamkeit entfalten, womit allerdings nur die eine – negative – Seite einer Entwicklung angesprochen ist, die aus der verstärkten Anwendung von Big Data im Gesundheitskontext resultiert.

4.6.2 Neue solidarische Praktiken

Abschließend seien einige mögliche Entwicklungen mit einem positiven, solidaritätsstützenden Potenzial angedeutet. Die differenzierte Offenlegung von Risikofaktoren über Big-Data-Analysen, die Daten aus allen Lebensbereichen integrieren, könnte künftig ergeben, dass der weitaus überwiegende Teil der Bevölkerung gemischte Risikoprofile hat, die protektive und günstige

[356] Verbraucherzentrale Nordrhein-Westfalen 2015, 18.

Faktoren ebenso einschließen wie negative Krankheitsrisiken (körperlicher, mentaler, verhaltensbedingter und anderer Art), und dass nur bei sehr wenigen Gruppen, also jenen mit genetisch determinierten, insbesondere monogenetischen Erkrankungen, das Eintreten schwerer Erkrankungen sicher ist. Der Ausschluss oder die Schlechterstellung solcher Gruppen dürfte sich schon aus bekannten, ethisch gut fundierten und gesetzlich verbrieften Antidiskriminierungsgründen verbieten.

Dem Begriff der Solidarität können aber noch andere Phänomene zugeordnet werden, die bereits gegenwärtig sind und die als beachtens- und gegebenenfalls förderungswürdig qualifiziert werden können. In verschiedenen Bereichen der Medizin hat der Einsatz von Big-Data-Technologien zur Entwicklung neuer prosozialer Unterstützungspraktiken geführt. Dies deutet darauf hin, dass neben dem oben skizzierten, solidaritätszersetzenden Potenzial auch Solidaritätsgewinne durch Big Data möglich sind. Ähnlich wie im Zusammenhang der Fragen zur Selbstbestimmung hängt es vom jeweiligen Kontext, den jeweiligen Technologien und den involvierten Absichten ab, ob solche Gewinne realistisch erwartet werden können.

Beispiele für neue solidarische Praktiken sind etwa die Bildung kleinerer Gruppen von Patienten, die – insbesondere seltene – Krankheitsrisiken oder bereits konkrete Krankheitserfahrung teilen, und die ihre Daten und Bioproben in gemeinschaftlichen Pools zusammenführen und für die Forschung an ihrem Krankheitsbild zur Verfügung stellen. Sozialwissenschaftliche Befragungen zu solchen Phänomenen zeigen, dass die Teilnehmer explizit auf die geteilte Erfahrung und die sich daraus entwickelnde Motivation, anderen mit ähnlichen Erfahrungen zu helfen, hinweisen.[357] Solche Initiativen entstehen oft im Rahmen solidarischer Unterstützungsgruppen, etwa als Patienten-Selbsthilfegruppen, die angesichts der neuen Möglichkeiten digitalisierter Datenspeicherung durch Laien verstärkten Zulauf erfahren. Andere Solidaritätsgewinne sind gegenwärtig in Online-Foren zu beobachten, in die Patienten ihre Erfahrungen und Krankheitsdaten (aus Klinik und Selbstvermessung) einspeisen, sie dort austauschen, gemeinsam diskutieren und für das individuelle Krankheitsmanagement nutzen. Mit der zunehmenden Entwicklung von online vernetzten Instrumenten für die Patienten-Selbsthilfe[358] steht zu erwarten, dass diese Praktiken zunehmen werden. Ob sie nachhaltige Gewinne für Patienten darstellen oder ihrerseits wieder die Gefahr bergen, mit zunehmender Differenziertheit und Prädiktivität der Daten exklusiv zu werden, ist noch unklar und dürfte von verschiedenen Fak-

[357] Verschiedene Verweise in Prainsack/Buyx 2017.
[358] Ein Beispiel für dieses Phänomen ist die Online-Plattform iManageCancer, die durch ein EU-H2020-Projekt entwickelt wird (vgl. http://imanagecancer.eu [17.10.2017]). Neben verschiedenen Apps zur Gesundheitsinformation gibt es hier die Möglichkeit der Datenspeicherung und Vernetzung sowohl zwischen Patienten als auch mit Behandlungsalgorithmen; Ziel ist es, dass Patienten besser mit ihrer Krebserkrankung leben und Teile der Behandlung – inklusive der ambulanten Chemotherapie (Dosisanpassung, Behandlung von Nebenwirkungen etc.) – in Abstimmung mit ihren behandelnden Ärzten selbst übernehmen.

toren abhängen, etwa den Motiven des jeweiligen Datenaustauschs sowie der Qualität der Datenaufbereitung. Aus Solidaritätsperspektive gibt es vorerst gute Gründe, solche sich entwickelnden Praktiken zu fördern, allerdings transparent und mit Augenmerk auf mögliche Risiken.

4.7 Verantwortung

Angesichts der vielfachen individuellen und kollektiven, staatlichen und nicht staatlichen Handlungs- und Entscheidungsoptionen sowie institutioneller und rechtlicher Gestaltungsmöglichkeiten ist zu prüfen, wer in welchem Maße als das Subjekt dieser Handlungsmöglichkeiten gedacht werden kann. Traditionell wird eine solche ethische Prüfung mit dem Verantwortungsbegriff verbunden. Dabei herrscht ein breiter Konsens, dass Verantwortung eine moralische Kategorie der Orientierung wie der Beurteilung von Handlungen und Unterlassungen ist.[359] Sie lässt sich nach Handlungs- und Entscheidungstypen, aber auch nach der Ausgestaltung institutioneller Strukturen differenzieren.

Grad und Reichweite der Verantwortungsfähigkeit hängen ab von äußeren und inneren Fähigkeiten, fachlichen Kompetenzen, Macht oder räumlicher, zeitlicher oder sozialer Nähe zum Geschehen, das nach Verantwortung verlangt. Gemäß diesen differenzierenden Kriterien wird versucht, die generelle Frage der Verantwortung zu beantworten – wer was wofür weswegen wovor wann und wie verantwortet – also die Frage nach Subjekt, Gegenstand, Zweck, Grund, Instanz, Zeit und weiteren Umständen der Verantwortung.

Die entsprechenden Antworten mögen in unterschiedlichen Sphären zu finden sein. Verantwortung kann moralisch, rechtlich, politisch und vertraglich eingefordert und übernommen werden. Für alle Sphären gilt dies retro- wie prospektiv, vor und nach einer Handlung oder Entscheidung. Dabei stehen die unterschiedlichen Typen von Verantwortung oft in einem sachlichen Wechselverhältnis: Man erwartet genau von demjenigen die Übernahme von Verantwortung für die Zukunft, den man in einem geschehenen Schadensfalle zur Rechenschaft ziehen würde. Verantwortung kann sowohl für Tun als auch für Unterlassen bestehen. Grund und Folgen dieser ethisch wie rechtlich bedeutsamen Unterscheidung sind überaus kontrovers. Strittig ist auch, ob in arbeitsteiligen, vollständig durchtechnisierten Arbeits- oder Entscheidungsprozessen neben einzelnen Individuen auch Kollektive, Organisationen, Institutionen oder gar Maschinen, insbesondere selbstlernende Maschinen, Verantwortungsträger sein können. Das komplexe Zusammenspiel zwischen Einzelnen (als Individuen, aber auch als Rollen- und Funktionsträgern), Institutionen und Technik gewinnt beim Einsatz von Big Data im ge-

[359] Vgl. Heidbrink/Langbehn/Loh 2017.

sundheitsrelevanten Bereich besondere Bedeutung und verdient entsprechend Aufmerksamkeit. Vermieden werden sollte eine undurchsichtige Diffusion von Verantwortung, die insbesondere dort droht, wo viele Akteure und hoch technisierte Prozesse zusammenwirken.

4.7.1 Verantwortung des Einzelnen bezüglich der Weitergabe gesundheitsbezogener Daten in unterschiedlichen Rollen und Kontexten

Es gehört zu den Gemeinplätzen des klassischen Datenschutzrechts und seiner moralischen Grundlagen, dass die Erhebung, Speicherung, Verarbeitung und Weitergabe von personenbezogenen Daten nach informierter Einwilligung und unter Berücksichtigung der weiteren Datenschutzprinzipien (Zweckbindung, Verhältnismäßigkeit, Datensparsamkeit, Transparenz, Verbot mit Erlaubnisvorbehalt) erlaubt ist (siehe Kapitel 3). Im englischen Sprachraum hat sich dafür der Begriff *notice and consent* eingebürgert. Er besagt, dass nach der zumindest formellen Kenntnisnahme der Bedingungen der Datenerhebung, -verarbeitung und der damit als gegeben unterstellten Informiertheit die Verantwortung für die Folgen dieses Prozesses im Wesentlichen bei dem Individuum liegt, das die Daten zur Verfügung stellt. Technische Komplexität die Beschreibung möglicher Konsequenzen, wie sie in typischen Geschäftsbedingungen als rechtliche Bedingungen der vertraglichen Wirksamkeit formuliert sind, dürften jedoch die typischen Nutzer datenintensiver Verfahren oder Geräte weit überfordern. Da der Normalnutzer regelmäßig nur die Alternative hat, solchen schwer verständlichen und in ihren Folgen nicht absehbaren AGB zuzustimmen oder auf die Nutzung des jeweils angebotenen Verfahrens oder Gerätes zu verzichten, besteht für die meisten Menschen oft keine wirkliche Alternative. Von einer sachlich gehaltvollen und nicht nur formellen Verantwortung – jedenfalls im Sinne eines Verantwortungsbegriffs, der hinreichende Fähigkeiten und Handlungsalternativen voraussetzt – kann deshalb vielfach nicht die Rede sein.

Dennoch entbindet die Kritik am geläufigen Verständnis des Notice-and-Consent-Ansatzes nicht von der Notwendigkeit, auch im Big-Data-Zeitalter dem Individuum bestimmte Verantwortlichkeiten zuzuschreiben und entsprechende Handlungen und Einstellungen von ihm zu erwarten. Die idealerweise vorauszusetzende, aber durch ausschweifende und undurchsichtige AGB oft überstrapazierte Sorgfalt sollte sich deshalb auf die Beachtung längst bekannter Ratschläge richten – etwa: vorsichtig mit Identitätsdaten umzugehen; „kostenlose" Angebote genau zu prüfen; den Ort der Datenverarbeitung sowie die Datenschutzeinstellungen zu beachten; bei Nutzung durch mehrere Anbieter die Daten zu verschlüsseln; ein sicheres WLAN einzurichten; Vorsicht bei der Nutzung von Smartphones walten zu lassen, die besonders viele Daten (Orts- und Kontaktdaten) sammeln und weitergeben; die Software auf dem neuesten Stand zu erhalten; das Gerät regelmäßig abzuschalten etc.[360] Solche Maximen können jedoch im Big-Data-

[360] Vgl. Schaar 2014, 265-276.

Zeitalter umso besser umgesetzt werden, je mehr der Einzelne dafür Rahmenbedingungen vorfindet, die technisch wie organisatorisch eine leichte und effektive Nutzung ermöglichen.

Erfolgt die Nutzung von Big-Data-Anwendungen überdies aus einer bestimmten sozialen Rolle, mit der wie zum Beispiel bei Ärzten oder Wissenschaftlern ein spezifisches Ethos verbunden ist, dem andere vertrauen dürfen, so besteht eine besondere Verantwortung, die erhöhte Sorgfaltspflichten impliziert. Der Arzt hat dafür Sorge zu tragen, dass in seinem Aufsichtsbereich die Praxissoftware, das Praxisnetz, dem er gegebenenfalls angehört, oder die Weitergabe der Daten an Abrechnungsstellen etc. sicher funktioniert und nicht für Zwecke, verwendet wird, denen nicht zugestimmt wurde. Gegenüber Wissenschaftlern bestehen gerade wegen der bereits skizzierten Unmöglichkeit einer verlässlichen Anonymisierung besondere Sorgfaltserwartungen hinsichtlich der Genauigkeit der Datensammlung, der Verarbeitung und Weitergabe von Daten sowie gegebenenfalls der Programmierung von Algorithmen.

4.7.2 Verantwortung institutioneller Akteure

Die für Laien kaum mehr durchschaubare und selbst für Experten oft schwer nachvollziehbare Komplexität von Big-Data-Prozessen begründet für die involvierten Unternehmen und Institutionen moralisch und rechtlich eine besondere Verantwortung. Sie erfüllen oft am ehesten die Voraussetzungen einer plausiblen Zuschreibung von Verantwortung: Sie verfügen über die Fähigkeiten, Kompetenzen und die Macht, nötige Schritte der Kontrolle zu vollziehen oder andere, die Kontrolle unterlaufen, zu unterlassen und beides mit entsprechenden Mechanismen zu garantieren. Ob diese moralische Verantwortung rechtlich im Binnenverhältnis einer datenverarbeitenden Organisation (Behörde, Unternehmen, Wissenschaftsbetrieb) noch jeweils personalisiert wird, ist wegen der erwähnten Fähigkeiten im Außenverhältnis für die moralische Verantwortlichkeit solcher Organisationen ohne Belang.

Zu den Möglichkeiten von Unternehmen, Big-Data-Prozesse verantwortlich zu gestalten, gehört es vor allem, Bedingungen dafür zu schaffen, gegebene Zustimmungen widerrufbar zu machen. Zwar wird man einem Datengeber theoretisch immer zugestehen, die zukünftige oder fortgesetzte Datenerhebung durch eine Nichtteilnahme, das Abwählen bestimmter Optionen oder die Auflösung seines Accounts zu untersagen. Praktisch würde es seine Position aber deutlich stärken, wäre auch die nachträgliche Löschung von Daten möglich und garantiert. Technisch und moralisch ist es vorstellbar, die Verwaltung von Daten auf Abruf zu gestalten. Von der Löschung ausnehmen könnte man hinreichend aggregierte Daten, abgeleitete Daten oder Modelle, die nachweislich keinen Rückschluss auf den Einzelnen erlauben. Mit solchen Ansätzen De- und Rekontextualisierungen bei gleichzeitiger Wahrung hoher Anonymisierungsstandards zu ermöglichen und Institutionsvertrauen zu schaffen, dürfte eine der entscheidenden Aufgaben medizinischer Forschung in der Zukunft darstellen.

Eine weitere Möglichkeit, Verantwortung für die Rechte des Individuums zu übernehmen und dabei dennoch eigene legitime Geschäftsinteressen zu wahren, wären Stellvertretersysteme an den programmatischen Schnittstellen in Datennetzwerken, über die ein Großteil des Datenaustauschs erfolgt (siehe Kapitel 2). Da diese Datennetzwerke zumeist keine Schnittstellen bieten, die direkt für individuelle Datengeber nutzbar wären, könnten Software-Systeme, also algorithmische Akteure, die den Menschen und seine Interessen als *Daten-Agenten* in der Welt der Maschinen verträten, ein praktikabler Ausweg sein. Solche Software-Tools könnten vordefinierte Regeln für die Datenhandhabung beinhalten, die von Dritten (Verbraucherverbänden usw.) entwickelt werden, und würden mit den Systemen der Daten sammelnden und Daten nutzenden Unternehmen und Institutionen automatisch interagieren.

Hierdurch würde die im Allgemeinen unzumutbare Tätigkeit einer detaillierten händischen Verwaltung von Daten durch den Datengeber selbst durch eine programmatische Verwaltung von Daten ersetzt. Der Vorteil dieses Modells besteht darin, dass der Einzelne nicht darauf angewiesen ist, spezielle Werkzeuge zu schaffen, um seine Daten damit zu verwalten. Es folgt dem Prinzip der Repräsentanz eines Menschen durch ein in seinem Sinne agierendes Softwaresystem und gäbe dem Einzelnen eine technisch niedrigschwellige Möglichkeit, Verantwortung für die Wahl einer eigenen nicht nur kurz-, sondern auch mittel- bis langfristigen Strategie der Datenhandhabung zu übernehmen, ohne dass er dabei jede Einzelfrage selbst entscheiden müsste.

Hilfreich wäre für das Individuum zudem, wenn Daten sammelnde, verarbeitende und weitergebende Organisationen technische Möglichkeiten bereitstellten, um den unaufhörlichen Datenfluss für den Einzelnen an Stellen, die ihn betreffen, transparenter zu gestalten. Zuvor aber muss geklärt sein, was mit Transparenz gemeint ist und wie sie sich herstellen lässt. Die Forderung, Algorithmen offenzulegen, erscheint angesichts ihrer Komplexität und Dynamik als zu hohe Erwartung. Eine solche völlige Offenlegung würde auch die Wahrung von Geschäftsgeheimnissen, die für marktgängige Unternehmen unentbehrlich sind, erschweren oder sogar unmöglich machen. Wohl aber könnten Unternehmen in die Pflicht genommen werden, ihre Verfahren überprüfbar und somit verantwortbar zu machen. Man kann technisch beispielsweise überprüfen, ob ein Modell oder Algorithmus eine angestrebte oder behauptete Genauigkeit erreicht. Entsprechend ist es nur legitim zu verlangen und sicherstellen zu lassen, dass Verfahren bestimmte Bevölkerungsgruppen nicht systematisch benachteiligen. Man könnte über Auditing auf einer System-Ebene gewährleisten, dass Regeln zur Datenaufbewahrung, -anonymisierung oder -löschung eingehalten werden. Diese Überprüfung könnte in vielen Fällen von externen, unabhängigen Prüfstellen vorgenommen werden oder den Firmen selbst überlassen werden mit einer indirekten, prozessorientierten Prüfung.

Immer wieder wird von Daten sammelnden, verarbeitenden und weitergebenden Organisationen auch deshalb Transparenz verlangt, damit Individuen die Möglichkeit erhalten, die eigenen Datensätze selbst, aber auch ihre Weitergabe an Dritte überprüfen zu können. Um diesem ethisch legitimen Begehren zu entsprechen, müsste eine technische Infrastruktur geschaffen werden, die Herkunft, Verarbeitung, Verwendung und Austausch von Daten lückenlos und manipulationssicher protokolliert (siehe Abschnitt 2.2) und den Datengeber in die Lage versetzt, sich ein klares Bild von der Nutzung seiner Daten zu machen. Überlegungen, wie solche Organisationen mithilfe technischer Systeme und Transparenzverfahren ihrer Verantwortung genügen können, dass Datengeber im Datennetz möglichst souverän bleiben, agieren können und sich nicht aussichtslos verfangen, spielen im Gesundheitsbereich schon deswegen eine besonders große Rolle, weil viele der hier relevanten Daten von vornherein höchst sensibel sind. In der klinischen Praxis Tätige, Hersteller von Medizinprodukten, aber auch Versicherer, Kreditinstitute oder Auskunfteien müssten sich deshalb ihrer Verantwortung bewusst sein, die Datenqualität sowie Verfahren der Datensicherheit auf höchstem Niveau zu halten. Denn mit den jeweiligen Dateneingaben schreiben sich – im Einzelnen oft nicht nachvollziehbar – selbstlernende Algorithmen weiter.

Um die Qualität von Daten und Prozessen möglichst hoch zu halten, wären auch Regulierungen auf staatlicher Ebene denkbar. Eine andere Möglichkeit, inzwischen vielfach erwogen, besteht darin, über Zertifizierungen, Qualitätssiegel oder Selbstverpflichtungen, die von Interessen- oder Berufsverbänden bereitgestellt und überprüft werden, das Vertrauen in die jeweiligen Organisationen und Prozesse zu stärken. Inwieweit dies funktioniert, kann man an vergleichbaren Erfolgs- und Misserfolgsgeschichten in diversen Branchen ablesen.[361] Wenn es gelänge, ein möglichst einheitliches Qualitätssiegel zu etablieren, das für Datengeber zudem eine effiziente Berücksichtigung ihrer Bedürfnisse erkennen lässt, dürften sich die entsprechenden Institutionen auf einem verantwortungsbewussten Weg befinden. Voraussetzungen dafür wären zum Beispiel die zügige Beachtung und Prüfung von Beschwerden sowie ebenso zügige Korrekturen oder die Bereitstellung weiterer Beschwerdewege. Sie gäben dem Einzelnen proaktiv die Möglichkeit, das im folgenden Kapitel formulierte Ziel einer Datensouveränität zu erreichen und zu verteidigen.

Neben den Fragen des Zugangs, der Transparenz und der Handhabung von Daten stellt sich für Daten sammelnde, verarbeitende und weitergebende Organisationen eine weitere Verantwortungsfrage: ob, wann und inwieweit sie in bestimmten Situationen bereit sind, in die persönliche Kommunikation einzugreifen, die sie zwischen ihren Nutzern ermöglichen und gegebenenfalls auch auswerten (sofern die Geschäftsbedingungen das zulassen). Aufsehen erregte

[361] Vgl. https://www.splendid-research.com/Marktforschung/Guetesiegel-in-Deutschland-2013.pdf [27.10.2017].

Mitte 2017 die Nachricht, dass Facebook einen Algorithmus entwickelt hat, mit dessen Hilfe der Betreiber des sozialen Netzwerks meint, Personen mit Suizidabsicht identifizieren zu können und ansprechen zu sollen.[362] Offensichtlich kollidieren hier unterschiedliche Rechts- und ethische Verantwortungsmaximen. Einerseits spricht gegen diese Praxis die Ablehnung offensichtlicher Eingriffe in die Privat- oder gar Intimsphäre der betroffenen Person, wenn diese ungefragt mit einem Problem konfrontiert wird, das besser in persönlichen Gesprächen mit Personen aus dem näheren Umfeld oder mit professioneller Hilfe gelöst werden sollte. Auch mutet ein derartiger Übergriff des sozialen Netzwerks prima facie recht paternalistisch an. Wenn die Funktionssicherheit eines solchen Algorithmus aber wissenschaftlich gut belegt sein sollte, müsste man umgekehrt aus ethischer Perspektive auch berücksichtigen, dass es in einem solchen Fall buchstäblich um Leben und Tod gehen könnte. Suizid stellt zum Beispiel bei Jugendlichen die zweithäufigste Todesursache dar, und gerade junge Menschen nutzen soziale Netze besonders intensiv.[363]

Während bei Suizidgefahr die Abwägung der in Spannung zueinander stehenden Güter noch zugunsten eines Hilfsangebots oder gar einer Hilfspflicht ausfallen mag, verkompliziert sich diese Einschätzung bei weniger dramatischen Gesundheitsproblemen: Soll man ungefragt Hilfsangebote an Mitglieder sozialer Netzwerke oder identifizierbare Nutzer von Suchmaschinen verschicken, deren Aktivitäten im Netz sie als krankheitsgefährdet erscheinen lassen, zum Beispiel den Hinweis, einen bestimmten Arzt aufzusuchen? Und fällt die Antwort auf diese Frage anders aus, je nachdem, wie schwer die Erkrankung ist, auf die die Datenspuren des Nutzers hindeuten? Zweifellos wird man – auch eingedenk der in Abschnitt 4.1.2 skizzierten Herausforderungen von Einwilligungsmodellen im Big-Data-Zeitalter – tendenziell auf der sicheren Seite stehen, wenn man die Zustimmung eines Datengebers für solche möglichen gesundheitsbezogenen Interventionen vorab eingeholt hat. Gewiss wird es Grauzonen für mögliche Lösungen in solchen Fragen geben, etwa wenn ein Hinweis auf ein sich möglicherweise verschärfendes Gesundheitsproblem gegeben werden könnte, das sich gegenwärtig noch ohne größere Mühen und ohne erhebliche Risiken einer gesundheitlichen Stigmatisierung beheben ließe. Auch hier kommt es auf die Verantwortung von institutionell organisierten Akteuren an.

4.7.3 Verantwortung der staatlichen Organe

In einem Feld voller Akteure mit unterschiedlichen Verantwortungsmöglichkeiten wie im Bereich von Big Data und Gesundheit wird schnell der Ruf nach staatlicher Intervention laut. Selbstverständlich bewegen wir uns nicht in einem rechtsfreien Raum. Allerdings greifen ver-

[362] Vgl. hierzu etwa http://www.bbc.com/news/technology-39126027 [17.10.2017].
[363] Vgl. das Interview mit Jakob Henschel im Spiegel vom 11. März 2017: https://magazin.spiegel.de/SP/2017/11/149997421/index.html?utm_source=spon&utm_campaign=centerpage [17.10.2017].

schiedene rechtliche Regelungen, die sich lange Zeit bewährt haben, aufgrund der beschriebenen technischen Veränderungen nicht mehr so, wie das zur Zeit ihres Inkrafttretens der Fall war. Richtig ist auch, dass diese fehlende Passung insbesondere dadurch verursacht ist, dass viele nationale und EU-Regelungen in einer globalisierten Welt schwer umgesetzt bzw. leicht umgangen werden können. Daher ist zu fragen, wie der Staat auf nationaler Ebene, im Verbund der EU, aber auch als völkerrechtlicher Akteur Verantwortung übernehmen kann. Dazu sind im Rechtskapitel die wesentlichen Analysen dargelegt worden (siehe Kapitel 3). Darüber hinaus soll hier nur an einige rechtsethische Grundsätze für die Orientierung staatlicher Verantwortung erinnert werden. Angesichts der angedeuteten Problematik der Rechtsumsetzung sollte ein regulatorischer Subsidiaritätsgrundsatz gelten: Für das, was auf der Ebene von Selbstverpflichtungen und Zertifikaten effektiv gewährleistet werden kann, sollte man (bis zum nachweislichen Auftreten eines auf diesem Weg nicht lösbaren signifikanten Problems) keine detaillierten rechtlichen Regelungen erlassen. Die Beschränkung staatlicher Interventionen auf die Gewährleistung oder Prüfung eines Handlungsrahmens wäre umfassenden Detailregulierungen stets vorzuziehen. Ziel sollte sein, rechtliche Klarheit, Umsetzbarkeit und Flexibilität möglichst kohärent und produktiv aufeinander beziehen zu können.

Auch Gütesiegel, die auf Initiative von Daten sammelnden, verarbeitenden und weitergebenden Organisationen oder – falls dies nicht gelingt – mit staatlicher Unterstützung etabliert würden, könnten Individuen dabei unterstützen, souverän mit ihren Daten umzugehen. Das gilt insbesondere dann, wenn sie (im Idealfall einer rechtlich durchsetzbaren Regulierung) von allen oder (im Fall der Selbstverpflichtung einer Branche) von möglichst vielen Anbietern von Apps, Dienstleistungen, Studien usw. respektiert würden. Zu prüfen wäre auch, ob es gelingen könnte, mithilfe differenzierter Angebote zur maximalen Speicherdauer, Datensparsamkeit, Möglichkeiten der Löschung, sowie zur Datenmigration etc. dem einzelnen Nutzer nicht nur die Alternative einer Zustimmung oder Ablehnung des angebotenen Gesamtpaketes zu eröffnen. Ebenso sollte zumindest eine grobe Abwägung von Nutzen und Risiken bei der Auswahl einzelner Angebote des jeweiligen Service ermöglicht werden.

4.7.4 Fazit: Multiakteursverantwortung

Betrachtet man die drei Ebenen möglicher Verantwortungszuschreibung im Bereich gesundheitsbezogener Big-Data-Anwendungen (Individuen, Organisationen, Staat) wird man nach dem entscheidenden Kriterium der Fähigkeit, eine Maßnahme technisch einzuführen, umzusetzen und zu etablieren, als erstes die Daten sammelnden, verarbeitenden und weitergebenden Organisationen in der Pflicht sehen. Diese sollten Rahmenbedingungen für die verantwortliche informationelle Freiheitsgestaltung von Datengebern gewährleisten. Selbstverständlich befreit die Primärverantwortung solcher Organisationen den Einzelnen nicht davor, als Datengeber ebenfalls Verantwortung für die Nutzung seiner Daten zu übernehmen. Aber es reicht nicht aus, auf den rationalen Umgang von Nutzern mit Geschäftsbedingungen zu vertrauen oder zu

glauben, die vorsichtige und sparsame Verwendung von digitalen Anwendungen reiche aus, den Einzelnen im Datennetz vor ungewollten Nutzungen seiner Daten zu schützen. Hier bedarf es vielmehr der Unterstützung durch die Daten sammelnden, verarbeitenden und weitergebenden Akteure. Je weniger diese willens oder fähig sind, technische Möglichkeiten bereitzustellen, die dem Einzelnen die Kontrolle über seine Daten erleichtern, desto mehr drängt sich aus verantwortungsethischer Perspektive die Notwendigkeit für den Staat auf, gewährleistend, überwachend und gegebenenfalls auch regulierend und sanktionierend einzugreifen. Deutlich dürfte sein, dass das Ziel, dem Einzelnen die Möglichkeit zum souveränen Umgang mit seinen Daten zu geben, nur erreichbar ist, wenn dazu auf allen Seiten die jeweils gebotene Verantwortung übernommen wird.[364]

[364] Vgl. Braun/Dabrock 2016.

5 Datensouveränität als informationelle Freiheitsgestaltung

Mit den bereits existierenden sowie den sich abzeichnenden Einsatzformen von Big Data gehen enorme Chancen, aber auch ernst zu nehmende Risiken einher. Das gilt, wie oben in Kapitel 2 eingehend dargelegt, besonders für den Gesundheitsbereich. So zeichnen sich für Forschung, Diagnose- und Therapieansätze in der Medizin und der öffentlichen Gesundheitsversorgung, aber auch für die persönliche Gesundheitsgestaltung des Einzelnen bisher ungeahnte Anwendungsmöglichkeiten ab. Daneben und in Verbindung damit entstehen zudem ganz neue, wirtschaftlich hochinteressante Geschäftsfelder. Weil solche auf Big-Data-Bedingungen aufruhenden Chancen sich nur realisieren lassen, wenn Daten in großer Menge erfasst, analysiert und neu verknüpft werden, verbinden sich mit ihnen aber auch neue Risiken – speziell für Freiheit, Privatheit, Souveränität, Sicherheit und Wohlergehen sowie Herausforderungen für die gesellschaftliche Gerechtigkeit und Solidarität und die angemessene Übernahme von Verantwortung durch komplex interagierende Akteure (siehe Abschnitt 4.7).

Es zählt zu den Kerncharakteristika von Big Data, dass sich bestimmten Datenarten kaum noch wesensgemäß und a priori gegebene Sensibilitäten zuordnen lassen. Auch aus Daten, die aus für sich genommen neutralen Verwendungszusammenhängen stammen, können durch Rekontextualisierungen und Rekombinationen neue und unter Umständen hochproblematische Aussagen gewonnen werden. Mit Blick auf diese denkbaren Konsequenzen liegt auf der Hand, dass insbesondere im Bereich der Verarbeitung von bzw. zu gesundheitsbezogenen und medizinisch relevanten Daten eine besondere Aufmerksamkeit und Sorgfalt geboten ist. Zwar ist nicht ausgeschlossen, dass in bestimmten Handlungskontexten (wie zum Beispiel der klinischen Forschung) eine wachsende Bereitschaft zur Weitergabe bestimmter Daten durchaus positive Effekte haben könnte und deswegen unter bestimmten Voraussetzungen durchaus zu empfehlen sein kann. Doch gibt es daneben auch anders gelagerte Handlungsfelder, auf denen ein allzu freigebiger Umgang mit personenbezogenen Daten verheerende Folgen haben könnte und daher nur unter hohen Schutzstandards erfolgen darf.

Die grundlegende Zielsetzung, den Chancen und Risiken von Big Data gleichermaßen gerecht zu werden und die damit verbundenen faktischen und normativen Herausforderungen anzunehmen und zu bewältigen, wird nachfolgend mit dem Begriff der Datensouveränität bezeichnet.

5.1 Datensouveränität als Leitkonzept

Datensouveränität, verstanden als eine den Chancen und Risiken von Big Data angemessene verantwortliche informationelle Freiheitsgestaltung, sollte das zentrale ethische und rechtliche Ziel im Umgang mit Big Data sein.

Der Begriff der informationellen Freiheitsgestaltung knüpft an das Konzept der informationellen Selbstbestimmung an, entwickelt dieses aber weiter. Eine solche Freiheitsgestaltung gründet nicht in einem eigentumsanalogen Ausschlussrecht; vielmehr geht es wesentlich um die Befugnis, selbst zu bestimmen, mit welchen Inhalten jemand in Beziehung zu seiner Umwelt tritt und sich dadurch kommunikativ entfaltet. Informationelle Freiheitsgestaltung in diesem Sinne meint interaktive Persönlichkeitsentfaltung unter Wahrung von Privatheit in einer vernetzten Welt. Privatheit (siehe Abschnitt 4.2) darf dabei nicht allzu starr und statisch im Sinne eines (dauerhaft) abgeschlossenen Raumes begriffen werden, sondern muss flexibel und kontextbezogen gedacht und entsprechend strukturiert werden. Eine so verstandene informationelle Freiheitsgestaltung ist gekennzeichnet durch die Möglichkeit, auf Basis persönlicher Präferenzen effektiv in den Strom persönlich relevanter Daten eingreifen zu können. Verantwortlich ist sie dann, wenn sie sich dabei gleichzeitig an den gesellschaftlichen Anforderungen von Solidarität und Gerechtigkeit orientiert.

Der so gefasste Begriff der Datensouveränität als verantwortliche informationelle Freiheitsgestaltung ist von anderen möglichen Interpretationen abzugrenzen. Mit Datensouveränität im hier vertretenen Sinne werden weder die tradierten, letztlich kaum veränderten Regulierungsansätze des Datenschutzes nur unter neuem Namen fortgeschrieben, noch wird damit eine vollständige Neuorientierung oder gar eine Aufgabe des herkömmlichen Datenschutzgedankens oder die generelle Absenkung des bestehenden Schutzniveaus gefordert. Vielmehr geht es dem Deutschen Ethikrat darum, die benannten normativen Grundanforderungen, einschließlich der ethisch wie grundrechtlich fundierten informationellen Selbstbestimmung und damit auch des Datenschutzes, unter den Bedingungen von Big Data zur Geltung zu bringen.

Datenschutz war und ist kein Selbstzweck. Der Schutz personenbezogener Daten dient dem Schutz der Person: ihrer Privatsphäre ebenso wie der nicht kontrollierten, freien Entfaltung der Persönlichkeit in der Öffentlichkeit – sofern dem nicht Rechte anderer zwingend entgegenstehen. Jeder muss selbst bestimmen können, wie und wem er über die Freigabe seiner personenbezogenen Daten Zugang zur eigenen Person gewährt. Mit dem Begriff der Datensouveränität wird aber zugleich die Absicht betont, den souveränen, also selbstbestimmten und verantwortlichen Umgang des Einzelnen mit seinen eigenen personenbezogenen Daten mit einer Realisierung der Potenziale zu verknüpfen, die Big Data sowohl gesellschaftlich als auch für die individuelle Lebensgestaltung eröffnet. Daten müssen nicht allein als wichtiges individuelles Gut verstanden, sondern auch in ihrer kollektiven Dimension verstanden werden. Der Einzelne bleibt maßgeblicher Bezugspunkt von Datensouveränität; darüber hinaus ist aber die Relevanz von Daten als soziale Ressource ebenfalls zu berücksichtigen. Das hier entwickelte Leitkonzept der Datensouveränität soll also ein hohes Schutzniveau für den Einzelnen gewährleisten und es zugleich ermöglichen, die Chancen auf der kollektiven Ebene zu nutzen, etwa in der klinischen Praxis und gesundheitsbezogenen Forschung.

5.2 Datensouveränität im Gesundheitsbereich

Datensouveränität im vorgenannten Sinne ist ersichtlich nicht auf den Bereich des klassischen Gesundheitswesens beschränkt. Im Gegenteil unterscheiden sich die dort vorfindlichen Herausforderungen nicht kategorisch, sondern nur graduell von denen anderer Sektoren. Gleichwohl lassen sich im Umgang mit medizin- und gesundheitsrelevanten Daten bestimmte Eigenarten verdeutlichen: Verantwortliche informationelle Freiheitsgestaltung speziell im Gesundheitsbereich hat zum Ziel, die Big-Data-spezifischen Potenziale für die medizinbezogene Forschung, die klinische Anwendung und das individuelle Gesundheitsverhalten zu nutzen und die damit einhergehenden Risiken auf ein Minimum zu reduzieren.

Mithilfe der neuen technischen Verfahren zur Erfassung, Analyse und neuen Verknüpfung großer Datenmengen sind Wissenszuwächse im Verständnis der Entstehung und Entwicklung komplexer gesundheitlicher Beeinträchtigungen möglich, die mittel- und langfristig für innovative Behandlungskonzepte fruchtbar gemacht werden können. Von derartigen Konzepten können Patienten in der klinischen Versorgung in zweifacher Hinsicht direkt profitieren. Zum einen eröffnet eine genauere Stratifizierung von Patienten in medizinisch relevante Subgruppen im Sinne des Gebotes der Schadensvermeidung Chancen, Patienten vor unnötigen Belastungen durch erfolglose Behandlungsversuche zu schützen. Zum anderen soll Big Data im Sinne des Gebotes der Wohltätigkeit medizinischen Handelns eine verbesserte Diagnostik, Prädiktion und Therapieplanung ermöglichen, die selbst dann von enormer Bedeutung für die Patientenversorgung sein dürfte, wenn ein umfassendes Verständnis vieler multifaktoriell bedingter Erkrankungen noch in weiter Ferne liegt. Auch das Potenzial von Big-Data-Anwendungen für den Bereich der individuellen gesundheitsbewussten Lebensgestaltung eröffnet neue Dimensionen der Prävention, die nicht nur dem Einzelnen, sondern auch der Solidargemeinschaft der Krankenversicherten zugute kommt.

Diesen Chancen stehen jedoch auch bestimmte Risiken gegenüber, die unter anderem wichtige Solidaritäts- und Gerechtigkeitsaspekte betreffen. Für Einzelne bestehen Risiken von Datenschutzverletzungen und einer informationellen Selbstgefährdung, die unter anderem zu Diskriminierung, Stigmatisierung und damit zu negativen Folgen im Sinne der Teilhabe- und Befähigungsgerechtigkeit führen könnten. Verstärkte, Big-Data-basierte Risikostratifizierung könnte zu einer schleichenden oder auch, wenn diese sich beispielsweise im privaten Krankenversicherungssektor tariflich niederschlüge, expliziten Entsolidarisierung in der Gesellschaft und den Versicherungssystemen führen und besonders vulnerable Gruppen betreffen. Die Wissenschaftsgemeinschaft wiederum könnte zum Beispiel durch eine übermäßige Zunahme von Daten zweifelhafter Qualität oder durch Probleme beim Datenaustausch und der Datenverknüpfung herausgefordert werden. Solche Risiken von Big Data dürfen nicht ignoriert werden. Sie bedürfen stets der kritischen Abwägung und sollten nach Möglichkeit minimiert werden.

Sowohl Einzelpersonen als auch mit Big Data arbeitende Organisationen müssen daher befähigt werden, in Kenntnis wichtiger Risiken verantwortlich zu handeln, um die Chancen von Big Data nutzen zu können. Zugleich sind die strukturellen Voraussetzungen, Sicherungen und Schutzmechanismen so einzurichten, dass sie bei optimaler Risikominimierung dem Individuum maximale Freiheit im Handeln erlauben. Gerade wenn man auf das schon in sich komplexe deutsche Gesundheitssystem und auf die vielfältige, zudem oft global vernetzte Forschung und auf die ebenso global agierende Datenindustrie blickt, liegt zudem die Bedeutung einer differenzierten Multiakteursperspektive auf der Hand, um Datensouveränität in den vielfältigen gesundheits- und medizinbezogenen Handlungskontexten effektiv, effizient, fair und verantwortlich umzusetzen.

Vor diesem Hintergrund lassen sich bei der Wahrnehmung und Gestaltung von Datensouveränität zwei, einander zunehmend annähernde und bereits jetzt teilweise überschneidende Sphären unterscheiden: erstens die Sphäre der bislang schon durch vergleichsweise klare und strikte Datenschutz-, Qualität- und Sicherheitsstandards gekennzeichneten Datennutzung in der medizinbezogenen Forschung und klinischen Praxis, und zweitens die Sphäre der zunehmend den Gesundheitssektor mitbestimmenden, allerdings sehr heterogenen Angebote des freien Marktes. Letztere reichen von Anwendungskonzepten, die sehr nahe an der ersten Sphäre und den mit ihr verbundenen Standards liegen, bis hin zu ersichtlich unseriösen, nicht auf nachhaltige Gesundheitsförderung angelegten Angeboten.

5.3 Grundzüge eines an Datensouveränität orientierten Gestaltungs- und Regelungskonzepts

Der Begriff der Datensouveränität erhält damit im gesundheitsbezogenen Kontext eine spezifische Prägung. Er berücksichtigt die bisherigen Überlegungen, also die Analyse der bisherigen wissenschaftlichen, technischen und gesellschaftlichen Entwicklungen (siehe Kapitel 2) und rechtlichen Rahmenbedingungen (siehe Kapitel 3) sowie die Untersuchung der grundlegenden normativen Grundbedingungen von Big Data (siehe Kapitel 4). Gleichzeitig bildet er das maßgebende Kriterium für die folgenden Gestaltungs- und Regelungsüberlegungen, die in ihrer Gesamtheit eine ausgewogene und angemessene Behandlung der mit Big Data verbundenen Chancen und Risiken sicherstellen können und in die in Kapitel 6 ausgesprochenen Empfehlungen münden.

Am Anfang eines an Datensouveränität orientierten Gestaltungs- und Regelungskonzepts steht dabei die Einsicht in die Sinnlosigkeit einer reinen Blockadehaltung. Die schon aktuell vorhandenen – und sich in Zukunft noch intensivierenden – Big-Data-Entwicklungen können erkennbar nicht aufgehalten, sehr wohl aber gestaltet werden. Allerdings kann den diesbezüglich bestehenden, berechtigten Sorgen, wie in der rechtlichen Analyse gezeigt wurde, mit den Handlungsformen und Schutzmechanismen des traditionellen Datenschutzrechts nur unzureichend

begegnet werden. Damit die verfassungsrechtlich garantierte informationelle Selbstbestimmung tatsächlich wirksam ausgeübt werden kann, verwendet das klassische Datenschutzkonzept Instrumente wie die Datensparsamkeit bzw. -minimierung und die unmittelbare Zweckbindung. Derartige Instrumente stoßen jedoch bei Big Data klar an Grenzen. Sie konsequent anzuwenden, ist im Zeitalter von Big Data weder durchgängig möglich noch sinnvoll. Praktisch erwüchse hieraus deshalb kein zusätzlicher Schutz; entsprechende Forderungen sind somit also irreführend. Zudem würden innovative Potenziale von Big Data stark eingeschränkt oder verloren gehen.

Auch eingedenk der Kritik, die das Konzept des Rechts auf informationelle Selbstbestimmung in seiner spezifischen Ausgestaltung durch die verfassungsgerichtliche Rechtsprechung und die einfachgesetzliche Umsetzung erfahren hat, und vor dem Hintergrund der neuen europäischen Datenschutz-Grundverordnung (DSGVO) ist aber auf Basis der in dieser Stellungnahme vorgelegten grundlegenden normativen Erwägungen daran festzuhalten, dass es hinreichender rechtlicher und außerrechtlicher Schutzmechanismen bedarf, um auch unter den Bedingungen von Big Data, gerade im sensiblen Gesundheitssektor, den Gefahren von Machtasymmetrien und dadurch bedingten Datensouveränitätsverlusten effektiv entgegenzuwirken. Um weiterhin ein angemessen hohes Schutzniveau zu gewährleisten, ist daher ein verändertes, die Komplexität und Entwicklungsdynamik von Big Data stärker spiegelndes Gestaltungs- und Regelungsmodell zu erarbeiten. Dieses soll den Datensouveränitätsgedanken, das heißt die verantwortliche informationelle Freiheitsgestaltung, multidimensional und mit Blick auf unterschiedliche Akteursgruppen und Handlungskontexte reflektieren und dabei deren zuvor skizzierte Verantwortungsmöglichkeiten und -zuschreibungen (siehe Abschnitt 4.7) aufgreifen.

Eine derartige, gleichermaßen multidimensionale wie multiakteursbezogene Perspektive muss sich unter den Bedingungen von Big Data von überholten Vorstellungen einer spezifischen, vorgegebenen Sensibilität bestimmter Daten und hierauf rekurrierender besonderer Schutzmechanismen lösen. Datenschutz kann nicht mehr statisch an bestimmten Daten und Datennutzungskategorien ansetzen, sondern muss sich auf ständige Rekombinationen und Rekontextualisierungen einstellen. Regulatorisch betrachtet, ergibt sich aus dieser Einsicht die Konsequenz, sich nicht länger ausschließlich oder vorwiegend auf das frühe Stadium der Datenerhebung und eine dort schon feststehende, weitgehend kontextunabhängige Datensensibilität zu konzentrieren. Stattdessen ist stärker auf die jeweiligen Datenverwendungszusammenhänge zu achten. Deshalb sind in den unterschiedlichen Phasen der Erfassung, Analyse und neuen Verknüpfung von Daten die einschlägigen Prozesse, Akteure und ihre jeweiligen Verantwortlichkeiten zu fokussieren.

Das auf Datensouveränität ausgerichtete Gestaltungs- und Regelungsmodell nimmt dabei vor allem die Datengeber als den entscheidend zu schützenden und zu achtenden Zweck in den

Blick. Ziel ist es, über eine gleichermaßen kontextsensible wie falladäquate Regulierung und Institutionengestaltung diese Subjekte, aber auch die mit ihnen in Verbindung stehenden Organisationen, zu einem souveränen Umgang mit ihren Daten zu befähigen. Angesichts der im gesundheitsrelevanten Bereich besonders vielfältigen und höchst selbstbestimmungs- und persönlichkeitsrelevanten Handlungskontexte für die Erfassung, Analyse und Neuverknüpfung von Daten ist dieses Ziel dort besonders relevant. Die normative Erwartung, die im Begriff der Datensouveränität implizierte Spannung von Big-Data-Chancen und -Risiken verantwortlich zu gestalten, kann insbesondere in kaskadisch strukturierten Einwilligungskonzepten umgesetzt werden (siehe Abschnitt 4.1.2), die bis hin zu den Möglichkeiten einer advokatorischen oder repräsentativen Einwilligung reichen. Ihre Umsetzung kann durch technische und regulatorische Maßnahmen in ihrer Wirksamkeit verstärkt werden, wie beispielsweise durch die Öffnung programmatischer Schnittstellen für Datengeber, Datentreuhändlermodelle oder Mechanismen zur besseren Nachverfolgbarkeit von Datennutzungen.

Allerdings obliegt die Ermöglichung und Gestaltung von Datensouveränität keineswegs allein den Betroffenen, sondern setzt neben der Verständigung über die normativ maßgeblichen und regulativ sicherzustellenden Grundparameter eine bereichsbezogene, diversifizierte Analyse der unterschiedlichen Verantwortungssphären und -fähigkeiten voraus. Deren Ausgestaltung lässt sich nicht durch bloße Teilmodifikationen einzelner Instrumente oder gar allein durch zusätzliche Ressourcenzuweisungen und erhöhte Durchsetzungskapazitäten, etwa durch eine Aufwertung und bessere Ausstattung der Datenschutzbeauftragten, erfüllen. Die Aufstockung finanzieller, sächlicher und personeller Mittel ist zwar zweifellos notwendig, aber zumindest im Rahmen der vorhandenen Regulierungsansätze als nicht hinreichend zu betrachten. Es bedarf vielmehr einer umfassenden gesamtgesellschaftlichen, vermutlich schlussendlich nur global zu stemmenden Anstrengung, die rechtliche wie außerrechtliche Mechanismen einbezieht und auch technische Weiterentwicklungen berücksichtigt.

Dafür sollten vereinfachende Pauschallösungen aufgegeben werden zugunsten komplizierterer, aber auch flexiblerer und problemadäquater, institutionell diversifizierter Kombinationsmodelle. Erforderlich ist dabei insbesondere eine vorstrukturierende normative Begleitung, die früh ansetzt und etwa die technische Ausgestaltung datennutzungsintensiver Instrumente mit berücksichtigt, insgesamt aber darauf ausgerichtet ist, eine hohe Vertraulichkeit der Datenverarbeitungsprozesse sicherzustellen – insbesondere in den Bereichen, die der klinischen Praxis und medizinbezogenen Forschung zugeordnet werden oder ihnen nahestehen. Insoweit kann an den oben für den gesundheitsrelevanten Bereich entfalteten Zwei-Sphären-Gedanken angeknüpft werden. Hier gilt es, die heterogene zweite Sphäre nach folgender Grundregel zu gestalten: Je näher einzelne Anwendungen an die erste Sphäre heranreichen, desto mehr besteht ethisch und rechtlich die Aufgabe, ihre Gestaltung multiakteursbezogen in die Richtung der

dort generell vorherrschenden Qualitäts-, Schutz- und Vertraulichkeitsstandards zu entwickeln. Im Sinne der regulatorischen Subsidiarität und um überschießende Regulierung zu vermeiden, sollte dabei jedoch so lange wie möglich eher mit Anreizen, Selbstverpflichtungen, Zertifizierungen und sonstigen weicheren Lösungen gearbeitet werden als mit harten, sanktionsbewehrten rechtlichen Vorgaben. Ebenso gilt es bei den beteiligten Unternehmen die Einsicht zu befördern, dass solch ein Konzept zumindest mittel- und langfristig auch wettbewerbliche Vorteile gegenüber Anwendungsmodellen bietet, in denen die Datensouveränität angesichts kurzfristig zu verwirklichender kommerzieller Interessen hintansteht.

Zusammenfassend geht es also darum, bei allen Akteuren eine Sensibilität für die konstitutive Bedeutung der Souveränität des Datengebers in allen (gesundheitsbezogenen) Datenprozessen zu wecken, zu steigern oder zu erhalten. Es ist zwar anzunehmen, dass sich die beiden beschriebenen Sphären auch angesichts der erwartbar zunehmenden Präzision von Sensoren und Datenspeicher- sowie -weitergabemöglichkeiten noch stärker aufeinander zubewegen werden, als dies jetzt schon der Fall ist. Dies darf aber nicht dazu führen, das vorhandene (höhere) Qualitäts- und Schutzniveau der ersten Sphäre, also im Bereich der gesundheitsbezogenen Forschung und klinischen Praxis, zu nivellieren. Vielmehr gilt es, umgekehrt auch in der zweiten Sphäre darauf hinzuweisen und darauf zu drängen, für den Umgang mit Big Data keine Absenkung sondern eher die Anhebung von Qualitäts- und Schutzstandards anzustreben. In analoger Weise ist dementsprechend mit eventuellen Bemühungen zu verfahren, die existierenden nationalen und EU-rechtlichen Beschränkungen durch Verlagerung in weniger datenschutzsensible Rechtsordnungen zu umgehen. Solche Absichten widersprechen dem hier vertretenen Konzept der Datensouveränität. Um ein darauf aufbauendes Gestaltungs- und Regelungskonzept verantwortlich, effektiv und effizient umzusetzen, spricht der Deutsche Ethikrat nachfolgend unter Berücksichtigung der hier skizzierten Sphärendifferenzierung und unterschiedlichen Verantwortungsmöglichkeiten eine Reihe von Handlungsempfehlungen aus.

6 Empfehlungen

Angesichts der vielfältigen Herausforderungen, die sich aus den rasanten Entwicklungen bei der Erhebung, Verknüpfung und Analyse großer Datenmengen im Gesundheitsbereich ergeben, hält der Deutsche Ethikrat einen grundlegenden Wandel im Verständnis von gesundheitsrelevanten Daten und des Umgangs mit ihnen für erforderlich: Anstelle des bislang geläufigen Modells, das bestimmten Datentypen statische Sensibilitätsgrade zuordnet, sollte angesichts zunehmender Möglichkeiten der Dekontextualisierung und Rekontextualisierung von Daten die kontextabhängig wandelbare Sensibilität von Daten betont werden.

Datensouveränität als Leitkonzept

Der Deutsche Ethikrat empfiehlt ein Gestaltungs- und Regelungskonzept, das sich am zentralen Ziel der Datensouveränität orientiert. Unter Datensouveränität verstehen wir eine den Chancen und Risiken von Big Data angemessene verantwortliche informationelle Freiheitsgestaltung. Um dies zu gewährleisten, ist das traditionelle, primär auf die grundrechtlich geschützte informationelle Selbstbestimmung bezogene Datenschutzrecht weiterzuentwickeln und neu zu gestalten, indem inhaltlich umfassende grundlegende normative Vorgaben einbezogen und instrumentell neue Wege beschritten werden. Um die Chancen, die Big Data im Gesundheitsbereich eröffnet, zu nutzen und zugleich den Risiken neuer Formen asymmetrischer Macht und dadurch bedingten Verlusten an individueller Selbstbestimmung sowie möglicher Benachteiligung und Diskriminierung wirksam entgegenzutreten, bedarf es hinreichender und geeigneter Schutzmechanismen und Gestaltungsstrategien. Wo sich tradierte Instrumente – wie die bislang gängige strikte Orientierung an Datensparsamkeit und enger Zweckbindung – als dysfunktional erweisen, müssen deshalb andere Möglichkeiten, individuelle Freiheit und Privatheit zu wahren und eine gerechte und solidarische Gesellschaft zu gestalten, in den Vordergrund treten. Der Deutsche Ethikrat empfiehlt ein anspruchsvolles, innovationsoffenes Regulierungs- und Gestaltungskonzept, das eine Vielzahl von Akteuren einbindet und dabei sowohl die Unterschiede zwischen als auch die zunehmende Annäherung und Überschneidung von zwei Sphären beachtet: der medizinbezogenen Forschung und klinischen Praxis einerseits und der heterogenen gesundheitsrelevanten Angebote des freien Marktes andererseits. Ein solches Konzept verlangt eine umfassende gesamtgesellschaftliche Anstrengung, die rechtliche wie außerrechtliche Elemente einbezieht, technische Weiterentwicklungen aufnimmt und deren grundrechtswahrende Verfügbarkeit für alle gesellschaftlichen Akteure gewährleistet.

Das vom Deutschen Ethikrat vorgeschlagene Gestaltungs- und Regelungskonzept sollte die folgenden Einzelmaßnahmen umfassen und ihre zeitnahe Verwirklichung und Finanzierung gewährleisten:

A. Potenziale erschließen

Um die Potenziale von Big Data im Gesundheitsbereich zu realisieren, ist eine möglichst reibungsfreie Kooperation zwischen zahlreichen Akteuren aus der klinischen Praxis, medizinbezogenen Grundlagenforschung, in gesundheitsrelevanten Feldern tätigen Unternehmen und individuellen Datengebern nötig. Sie sollte nicht nur auf die prospektive Sammlung und nachhaltige Bereitstellung von Datensätzen abzielen, sondern es auch ermöglichen, bereits vorhandene Datensätze aus Klinik und Forschung mit jeweils neu gewonnenen Daten in ethisch verantwortbarer Weise zu verknüpfen.

A1. Infrastrukturelle Grundvoraussetzungen schaffen

Die Nutzung der Chancen von Big Data im Gesundheitsbereich hängt entscheidend von der Verfügbarkeit einer leistungsfähigen Infrastruktur zur Erfassung, Speicherung, Analyse und Übertragung großer Datenmengen ab. Um problematische Abhängigkeiten von kommerziellen Anbietern infrastruktureller Dienstleistungen, die zudem häufig nicht den deutschen bzw. europäischen Schutzstandards unterliegen, zu vermeiden, sollte die öffentliche Hand gewährleisten, dass eine derartige Infrastruktur – insbesondere für die klinische Praxis und medizinbezogene Grundlagenforschung zeitnah und mit angemessenen Zugangsmöglichkeiten und öffentlicher Kontrolle geschaffen bzw. weiterentwickelt wird.

A2. Datenaustausch und -integration erleichtern

Ebenso sind der verantwortungsvolle Austausch und die Integration von gesundheitsrelevanten Daten zwischen vielfältigen institutionellen Akteuren durch eine Reihe von Maßnahmen und deren ausreichende öffentliche Finanzierung zu gewährleisten:

A2.1. Standardisierte Verfahren der Interoperabilität von Daten entwickeln und bereitstellen

Um eine adäquate Zusammenführung von Daten aus unterschiedlichen Quellen unter Berücksichtigung der Privatheitsansprüche der Datengeber zu ermöglichen, müssen Daten miteinander vergleichbar sein, das heißt einheitlich benannt und angemessen annotiert sein. Eine wesentliche Voraussetzung hierfür ist die Standardisierung von Datenformaten und die Schaffung von Möglichkeiten zur Qualitätskontrolle einschließlich einer transparenten Dokumentation der durchlaufenen Schritte.

A2.2. Kooperatives Forschungsdatenmanagement weiterentwickeln

Die bestehenden Initiativen zur Etablierung effizienter Kommunikations-, Kollaborations- und Koordinationsstrukturen zwischen beteiligten Einrichtungen sollten gebündelt, intensiviert und auf Dauer gestellt werden. Dabei ist auch auf geeignete Schnittstellen zur Telematikinfrastruktur sowie auf eine angemessene Verzahnung mit der im E-Health-Gesetz vorgesehenen Weiterentwicklung der Vernetzung im Gesundheitswesen zu achten.

A3. Daten- und Forschungsqualität fördern und schützen

Eine zentrale Zukunftsaufgabe ist es, die Qualität der Daten zu kontrollieren, um auf diese Weise zu hinreichend verlässlichen Aussagen zu gelangen. Dafür sind folgende Maßnahmen geboten:

A3.1. Epistemische Standards einhalten, insbesondere die der evidenzbasierten Medizin

Bei der Weiterentwicklung von Kontrollmechanismen für die Sicherheit und Wirksamkeit medizinischer Maßnahmen, die bisher nicht auf Big-Data-Anwendungen zugeschnitten waren, dürfen die etablierten Maßstäbe der evidenzbasierten Medizin nicht unterschritten werden. Auch Big-Data-basierte Verfahren müssen sich für medizinische Verwendungszwecke den etablierten klinischen Prüfungen zur Wirksamkeit und Sicherheit unterziehen.

A3.2. Einheitliche Daten- und Dokumentationsstandards einführen

Nicht nur im Sinne der Interoperabilität und Kooperation, sondern auch zur Sicherstellung einer effektiven Qualitätskontrolle ist es sinnvoll, einheitliche Standards einzuführen. Das umfasst beispielsweise Fragen der Formate der Daten selbst, der sie beschreibenden Metadaten, der Rekonstruktion der Verarbeitungsschritte und Versionskontrolle sowie die möglichst einheitliche Abbildung von semantischen Verknüpfungen und Hierarchien von Daten. Zu den die Datenqualität sichernden Standards zählen namentlich Dokumentationspflichten, mit deren Hilfe die Herkunft von Daten nachvollzogen werden kann und ihre weitere Nachverfolgbarkeit zumindest erleichtert wird.

A3.3. Datengütesiegel etablieren

Um die genannten Qualitätsstandards und die damit verbundenen Anforderungen transparent zu machen, sollten entsprechende Konformitätsbescheinigungen („Gütesiegel“) vergeben werden, die die Herkunft und Qualität der Originaldaten und ihrer Verarbeitungsschritte nachweisbar darstellen (zum Beispiel durch Verwendung der Blockchain-Technologie). Weil die Qualitätssicherung auch im Eigeninteresse der jeweiligen Akteure liegt, ist primär auf wissenschafts- und wirtschaftsinterne Kontrollmechanismen zu setzen. Soweit diese sich indes als defizitär erweisen, sind auch übergreifende rechtliche Vorgaben einzuführen.

A4. Rechtliche Rahmenbedingung für die Datennutzung zu Forschungszwecken anpassen

A4.1. Sekundärnutzung von Forschungsdaten weiterentwickeln

Wo es nach geltendem Datenschutzrecht zulässig ist, personenbezogene Daten auf der Grundlage einer sorgfältigen Interessenabwägung auch ohne Einwilligung zu verarbeiten, wenn dies wissenschaftlichen, historischen oder statistischen Zwecken dient und für diese erforderlich ist

(§ 27 BDSG n. F.), sollten im Interesse der Datensouveränität grundsätzlich entsprechende zusätzliche, prozedurale Schutz- und Gestaltungsmaßnahmen wie das Kaskadenmodell (siehe Empfehlung B2) zum Einsatz kommen.

A4.2. Rechtliche Möglichkeit für Individuen schaffen, die umfassende Nutzung ihrer Daten für die medizinische Forschung zu erlauben („Datenspende")

Das traditionelle Einwilligungsmodell setzt für die Erhebung personenbezogener Daten prinzipiell eine enge Zweckbindung voraus. Gerade weil am Einwilligungsmodell grundsätzlich festzuhalten ist, sind hier nicht nur prozedurale Erweiterungen, sondern auch bereichsbezogene Öffnungen sinnvoll. Namentlich sollte es ermöglicht werden, im Sinne einer umfassenden Zustimmung Datennutzung ohne enge Zweckbindung zugunsten der klinischen und medizinbezogenen Grundlagenforschung zu erlauben („Datenspende"). Voraussetzung ist eine umfassende Aufklärung über mögliche Konsequenzen, insbesondere mit Blick auf die Rechte anderer, etwa mitbetroffener Familienmitglieder. Notwendig ist ferner die wissenschaftlich begleitete Entwicklung einer entsprechenden Infrastruktur für die Erfassung, Speicherung, Pflege, Verarbeitung und Weitergabe von gespendeten Daten.

A5. Digitale Entscheidungshilfesysteme in der klinischen Praxis fördern

Der wechselseitige Wissenstransfer zwischen Forschung und klinischer Praxis und die Zulassung digitaler Angebote zur Unterstützung von Entscheidungen für eine verbesserte Versorgung von Patienten sollten beschleunigt werden. Zu diesem Zweck ist für dazu legitimierte Akteure ein – unter Wahrung der Datensouveränität – möglichst umfassender Zugang zu Forschungs- bzw. Versorgungsdaten und geeigneten gesundheitsrelevanten Big-Data-Anwendungen notwendig.

A6. Internationale Anschlussfähigkeit fördern

Mit Blick auf den internationalen Austausch von Daten sollten Standardisierungsbemühungen nicht auf das nationale Territorium beschränkt bleiben. Vielmehr bedarf es weitreichender Anstrengungen auf allen Ebenen (der Politik, der Wissenschaft und Technologieentwicklung) zur Angleichung von Standards.

Um die internationale Wettbewerbsfähigkeit deutscher bzw. europäischer Digitalanwendungen im Gesundheitsbereich einschließlich der damit verbundenen hohen Qualitäts- und Datenschutzstandards zu fördern und um diesbezüglich problematischen Abhängigkeiten entgegenzuwirken, sollten zudem Investitionen im Bereich Medizininformatik deutlich höher ausfallen und schneller umgesetzt werden, als bislang geplant. Sinnvoll erscheint insbesondere eine zielgerichtete Förderung des Datenmanagements in öffentlichen Krankenhäusern.

B. Individuelle Freiheit und Privatheit sichern

Die Bereitschaft, personenbezogene Daten zur Verfügung zu stellen, ist als Teil der informationellen Freiheitsgestaltung der Datengeber zu verstehen. Deshalb müssen sie dazu befähigt werden, souverän mit diesen Daten umzugehen und ihre Privatsphäre zu gestalten. Zudem müssen die Rahmenbedingungen geschaffen werden, um entsprechend angemessene Handlungsspielräume zu garantieren.

B1. Datenhoheit bewahren

Die Bestimmungsmacht des Datengebers über die eigenen personenbezogenen Daten ist angesichts der Zweckoffenheit und Verknüpfungsmöglichkeiten von Big Data so umfassend wie möglich zu wahren.

B1.1. Programmatische Schnittstellen für Datengeber öffnen („Datenagenten")

Insbesondere dort, wo die Datennutzung nicht vorab präzise eingegrenzt werden kann oder wenn eine Datensammlung und -verarbeitung kontinuierlich erfolgt, sollten in Ergänzung zu gängigen Zustimmungsmodellen geeignete Software-Werkzeuge („Datenagenten") zur Verfügung gestellt werden, die die eingespeisten Daten fortdauernd nach den Vorstellungen der Datengeber verwalten und damit größere Kontrolle, Transparenz und Nachvollziehbarkeit ermöglichen. Es sollte eine Standardisierung entsprechender programmatischer Schnittstellen durch Selbstregulation oder gesetzgeberische Maßnahmen erfolgen, die die Entwicklung solcher Datenagenten erleichtert. Die korrekte Funktionsweise der Schnittstellen und Datenagenten sollte durch Auditierungs- bzw. Zertifizierungsmaßnahmen unterstützt werden.

B1.2. Mitbestimmung bei der Datenweitergabe erleichtern

Bei der Weitergabe von Daten sollte grundsätzlich die Reversibilität der Datenerhebung sichergestellt werden: Jedes System, das personenbezogene Daten sammelt und als Input akzeptiert, muss – von wohlbegründeten Ausnahmen abgesehen – in der Lage sein, diese Daten ganz oder teilweise auch wieder zu löschen. Auch hier sollte daher ein Modell von Datenagenten, die als Kontrollinstanz in Datenpipelines integriert werden, zum Einsatz kommen. Durch geeignete Kommunikationskanäle (etwa eine entsprechende App) sollte der Datengeber nachträglich um Zustimmung zur Weitergabe ersucht werden und diese je nach Fall auch relativ einfach einschränken oder widerrufen können.

B1.3. Rechtsprobleme eines vermeintlichen Eigentums an Daten klären

Datensouveränität ist nicht mit einem „Eigentum" an Daten zu verwechseln. Soweit der Eigentumsbegriff seine wesentlichen rechtlichen Elemente impliziert – dauerhaft feste Beziehung und absolute Ausschlussmacht gegenüber Dritten –, ist er für die Zwecke der Gewährleistung von Datensouveränität wenig geeignet. Weil andererseits aber eine gewisse (allerdings flexible)

Datenhoheit des Einzelnen anzuerkennen ist, ist es sinnvoll, sich stattdessen intensiver auf die rechtlichen Rahmenbedingungen der Nutzung von Daten zu konzentrieren. Der Deutsche Ethikrat empfiehlt, zu diesem Themenkomplex eine umfassende, nicht nur mit juristischem Sachverstand, sondern interdisziplinär besetzte Expertenkommission einzurichten.

B2. Kaskadisch strukturierte Einwilligungsmodelle etablieren

Grundsätzlich sollte in der klinischen Praxis und medizinbezogenen Forschung weiterhin ein einwilligungsbasiertes Regelungskonzept Verwendung finden (Opt-in-Modell). Wann immer möglich, sollten Kaskadenmodelle der persönlichen Einwilligung eingesetzt werden, die verschiedene, dynamisierte Möglichkeiten bieten, Einwilligungsentscheidungen einmalig, regelmäßig oder für jeden Entscheidungsfall neu zu treffen oder zu delegieren (etwa an unabhängige Einrichtungen/Treuhänder oder Ähnliches). Unter der Voraussetzung, dass die in der Stellungnahme entwickelten Sicherungs- und Qualitätsstandards und privatsphärenfreundliche Grundeinstellungen gewährleistet sind, sollten bereits praxiserprobte, erfolgreiche Vorbilder, insbesondere aus dem Bereich der Biobanken, auch auf andere Sektoren übertragen bzw. angepasst werden.

B3. Privatsphärenfreundliche Grundeinstellungen gewährleisten

Weil Datengeber aus Zeitmangel, fehlendem Verständnis, subjektiv empfundener Alternativlosigkeit oder aus gutem Glauben häufig die vorgegebenen Einstellungen von Daten sammelnden und Daten verarbeitenden Anwendungen übernehmen, sollten Grundeinstellungen technisch entwickelt und weiter rechtlich abgesichert werden, die von vornherein einen angemessenen Schutz der Privatsphäre bieten (*privacy by design/privacy by default*). Dies gilt insbesondere für den bislang vergleichsweise unregulierten Bereich privater Angebote, zum Beispiel gesundheitsrelevante Apps für Mobilgeräte und entsprechende Messgeräte. Über die Vorgaben der Datenschutz-Grundverordnung zu nutzerfreundlichen Einstellungen hinaus ist durch zusätzliche Aufklärung darauf hinzuwirken, dass Nutzer die Konsequenzen einer Änderung der Grundeinstellungen tatsächlich verstehen.

B4. Einsatz von Algorithmen transparent machen und erläutern

Über die rechtlich ohnehin vorgesehenen Auskunftspflichten hinaus sollten die Zielvorgaben, Funktions- und Wirkweisen der Datenakkumulation und der verwendeten Algorithmen so erläutert werden, dass sie auch für Nichtspezialisten nachvollziehbar sind. Insbesondere sollte dies – unter Berücksichtigung der jeweiligen Erfordernisse des Schutzes von geistigem Eigentum – die folgenden Aspekte umfassen:

- welche Nutzerdaten als Eingabe in welche Analysen, Vorhersagemodelle und Entscheidungs- oder Auswahlprozesse einfließen bzw. welche Attribute, etwa zur Vermeidung von Diskriminierung, ausdrücklich nicht erhoben und einbezogen werden,

- welche Ableitungen, Schlüsse, Vorhersagen, Selektionen oder Entscheidungen auf der Basis dieser Daten mittels Algorithmen getroffen werden,
- ob und inwiefern Profile des Datengebers erstellt werden und welche erwartete Aussagekraft solche abgeleiteten Größen haben,
- in welcher Form personenbezogene Daten in anonymisierter Form in (statistische) Modelle einfließen und wer über deren Nutzungsrechte verfügt.

B5. Täuschung und Manipulation entgegenwirken

Es ist zu unterscheiden zwischen offenen, transparenten Methoden der Einflussnahme auf andere einerseits und problematischeren verdeckten Eingriffen, die sich daher der kognitiven Kontrolle der Adressaten entziehen, andererseits. Eine manipulative Datengewinnung und -nutzung, die die Datengeber etwa über Art und Zweck der Erhebung täuscht und/oder ihre mangelnde Einsichtsfähigkeit ausnutzt, ist rechtlich wie moralisch unzulässig. Insbesondere in sozialen Netzwerken, bei Apps und Online-Spielen sollten nicht nur staatliche Instanzen, sondern auch die Betreiber selbst entsprechenden Tendenzen strikter entgegenwirken.

B6. Digitale Bildung fördern

Datensouveränität setzt Grundkenntnisse über die Bedeutung und den Wert von Big Data und die damit verbundenen Risiken voraus. Da bereits Kinder digitale Anwendungen nutzen und dabei Daten generieren, sollte eine entsprechende Nutzerkompetenz schon in der Schule vermittelt werden. Über die rein technischen Aspekte der gängigen Digitalisierungsstrategien schulischen Unterrichts hinaus sollte dies als Querschnittsaufgabe für alle Fächer des schulischen Curriculums ausgestaltet sein, um der gerade bei Kindern und Jugendlichen virulenten informationellen Selbstgefährdung entgegenzuwirken und schon früh ein Bewusstsein für die rechtlichen, sozialen und ethischen Implikationen zu schaffen. Die Vermittlung solcher Nutzerkompetenz sollte daher zukünftig Teil der Lehreraus- und -fortbildung werden. Einrichtungen der Erwachsenenbildung sollten zudem kontinuierlich niedrigschwellige Angebote für alle Altersgruppen vorhalten. Auch Unternehmen und Institutionen sollten regelmäßig entsprechende interne Schulungen durchführen.

B7. Diskurs und Teilhabe stärken

Die kontinuierliche öffentliche Debatte über Big Data sollte stärker gefördert werden. Dafür sollten staatlicherseits verlässliche Informationen zur Verfügung gestellt und partizipative Verfahren etabliert werden. Diese sollten eine breite Beteiligung der Öffentlichkeit und einen Austausch mit der Fachwelt gewährleisten.

C. Gerechtigkeit und Solidarität sichern

C1. Fairen Zugang zu digitalen Angeboten schaffen

Von den Vorteilen der Digitalisierung sind manche Nutzergruppen regelmäßig ausgeschlossen, etwa aufgrund von Bildungshemmnissen. Um dem entgegenzuwirken, bedarf es nicht nur spezieller Informations- und Bildungsangebote, sondern es ist auch Sorge dafür zu tragen, dass digitale Angebote nicht von vornherein so konzipiert werden – zum Beispiel durch unverständliche, unnötig komplizierte Handhabung oder unnötig technische Sprache – dass sie exklusiv wirken. Software und Nutzeroberflächen sollten möglichst barrierefrei gestaltet werden.

C2. Diskriminierung und Stigmatisierung aufdecken bzw. verhindern

Es ist sicherzustellen, dass eine über Big Data erweiterte Entscheidungsbasis für gesundheitsrelevante Allokationsentscheidungen nicht dazu missbraucht wird, Personen oder Personengruppen zu diskriminieren oder zu stigmatisieren. Bei der Verwendung von Erkenntnissen aus Big-Data-Analysen besteht eine Gefahr darin, dass die zugrunde liegenden Daten, die gewählten Randbedingungen der Analyse und angewandten Algorithmen zu Ergebnissen führen können, die eine systematische und nur schwer erkennbare Diskriminierung von Personen oder Gruppen nach sich ziehen. Deshalb ist nicht nur vorab auf die Unzulässigkeit entsprechender Selektionskriterien ohne klare und angemessene Zweckbestimmung hinzuweisen, sondern es sind auch Verfahren zu entwickeln, mit denen eventuelle Verstöße aufgezeigt und sanktioniert werden können. Auch wenn hierfür sektor- bzw. institutioneninterne, subsidiäre Regelwerke durchaus sinnvoll sind, muss es darüber hinaus aber auch justiziable, sanktionsbewehrte hoheitliche Sicherungsmechanismen geben.

C3. Widerspruch bei automatisierten Entscheidungen ermöglichen

Bei algorithmenbasierten Entscheidungen bedarf es strukturierter Widerspruchsmöglichkeiten. Speziell im Bereich privater Versicherungen muss für abgelehnte Antragsteller der Anspruch auf eine für sie verständliche, individuelle Begründung der Ablehnung garantiert sowie ein kostenfreier und niederschwelliger Zugang zu internen und externen Beschwerde- und Schlichtungsinstanzen sichergestellt werden.

C4. Vulnerable Gruppen und Individuen schützen

Besondere Aufmerksamkeit erfordern Personen und Gruppen, die aufgrund individueller oder sozialer Umstände (gegebenenfalls vorübergehend) besonders anfällig dafür sind, dass ihnen mittelbar oder unmittelbar, strukturell oder intentional die Vorteile einer Digitalisierung des Gesundheitssektors vorenthalten oder die Nachteile im Übermaß aufgebürdet werden. Dies gilt in besonderem Maße für Kinder und Jugendliche sowie Menschen mit Behinderung und ältere Menschen. Sie sind nicht nur mit Blick auf den Erwerb der Fähigkeit zur verantwortungsvollen

Inanspruchnahme digitaler Dienste zu unterstützen, sondern müssen in ihrer spezifischen Vulnerabilität auch im Prozess der Datensammlung und -verwendung besonders geschützt werden. Datensouveränität berücksichtigt insoweit auch die keineswegs fixe, sondern individuell und situationsbedingt variierende Verantwortungsfähigkeit der Betroffenen.

C4.1. Einwilligungserfordernisse bei Kindern und Jugendlichen streng beachten

Die Vorgaben der Datenschutz-Grundverordnung zu Einwilligungen von Minderjährigen in Bezug auf Dienste der Informationsgesellschaft sollten strikt und zügig umgesetzt werden. Über die von der Datenschutz-Grundverordnung zugelassene Möglichkeit, das Mindestalter abzusenken, sollte nicht entschieden werden, ohne die Betroffenen (Kinder und Jugendliche) zu beteiligen.

C4.2. Schutzmechanismen für die Datenerhebung an sonstigen Personen mit eingeschränkter Einwilligungsfähigkeit entwickeln

Für die Datenerhebung an sonstigen Personen mit eingeschränkter Einwilligungsfähigkeit sollten besondere Schutzmechanismen entwickelt werden, ohne damit die Chancen einer Big-Data-basierten Forschung mit diesen Personen und zu deren Gunsten zu unterbinden. Die beteiligten Forschungsinstitutionen sollten sicherstellen, dass entsprechend dem Konzept der Entscheidungsassistenz den betroffenen Menschen selbst, ihrer Einsichtsfähigkeit gemäß, und ihren Betreuungspersonen hinreichende Informationen zur Entscheidungsfindung an die Hand gegeben werden.

C4.3. Einsatz von Chatbots restriktiv regeln

Der Einsatz von Chatbots zur Datenerhebung an Personen mit eingeschränkter Einsichtsfähigkeit bietet ein besonders hohes Manipulationspotenzial und sollte daher besonders restriktiv geregelt werden.

C5. Zuwendungsorientierte Medizin gewährleisten

Die persönliche Zuwendung zum Patienten in der medizinischen Praxis sollte durch den Einsatz von Big-Data-Anwendungen nicht geschwächt, sondern gestärkt werden. Zeitliche und finanzielle Kapazitäten, die etwa durch die Entlastung des versorgenden Personals von Routine-Tätigkeiten oder die schnellere und präzisere Diagnostik durch digitale Algorithmen frei werden, sollten in mehr persönliche Zuwendung für Patienten umgesetzt werden.

C6. Wirksame Haftung von Unternehmen, die im Gesundheitsbereich mit Daten arbeiten, sicherstellen

Angesichts der mit Big Data verbundenen Risiken erscheint es angemessen, speziell hierauf zugeschnittene Haftungsmodelle zu entwickeln. Hier ist zunächst genau zu beobachten, ob und inwieweit die neuen Regelungen des deutschen Datenschutzrechts, die die Möglichkeiten der

europäischen Datenschutz-Grundverordnung (DSGVO) bislang nicht ausschöpfen, ausreichen. Die DSGVO eröffnet die Möglichkeit, für einen effektiven Schutz von Personen vor Schädigung die Gefährdungshaftung einzuführen. Angesichts der Unsicherheiten der Haftung und der Beweisregelung ist eine derartige, auf die spezifischen Risiken von Big Data zugeschnittene Gefährdungshaftung zu erwägen. Diese Haftung sollte unabhängig von der Befugnis der Verwendung nur dann ausgeschlossen sein, wenn der Schaden unvermeidbar ist. Eine eventuelle summenmäßige Begrenzung der Haftung sollte so hoch sein, dass sie auch gegenüber großen Unternehmen spürbare Wirkung entfaltet.

D. Verantwortung und Vertrauen fördern

D1. Schutz- und Qualitätsstandards garantieren

D.1.1. Bestmögliche Schutzstandards gegen unbefugte Identifizierung von Individuen aus anonymisierten, pseudonymisierten oder aggregierten Datensätzen etablieren

Angesichts der unzureichenden Schutzeffekte der traditionellen Anonymisierung und Pseudonymisierung sollten angemessene ergänzende Schutzstandards etabliert werden, um die Hürden für eine Reidentifizierung zu erhöhen:

- Wo Identifikatoren einen relativ unmittelbaren Rückschluss auf die jeweilige Person erlauben (E-Mail, Login, Geräte-ID, Cookie-ID), sind diese durch anonymisierte Schlüssel zu ersetzen, deren Lebensdauer möglichst kurz zu halten ist.
- Wenn immer ein anonymer Nutzer sich unerwartet oder versehentlich direkt oder indirekt identifiziert, hat der Datensammler Sorge zu tragen, dass die Identifizierung durch Datenlöschung rückgängig gemacht wird (versehentliche Preisgabe von Namen, E-Mail, Telefonnummern, Kreditkartennummer, Ausweisnummer usw.).
- Wo immer ein Datensatz durch die Kombination von Attributen und Daten einen Nutzer mit hoher Wahrscheinlichkeit identifizierbar macht, sind auf jenen die gleichen datenschutzrechtlichen Maßnahmen anzuwenden wie bei expliziten Identifikatoren.
- Datensätze, deren Verbindung eine entsprechende Schutznivellierung mit sich bringt, müssen getrennt gehalten werden oder dürfen nur „flüchtig“ (das heißt ohne persistent in Datenbanken gespeichert zu werden) für wohldefinierte Zwecke verknüpft werden.

D.1.2. Anonymisierungsdefizite durch kontrollierten Zugang zu Daten kompensieren

Angesichts des verbleibenden Reidentifizierungsrisikos kommt der Kontrolle des Datenzugriffs besondere Bedeutung zu. Insbesondere in der klinischen Praxis und der medizinbezogenen Grundlagenforschung ist daher der Zugang zu Daten durch Aufbewahrung von gesundheitsrelevanten Daten in sicheren, technisch getrennten und voneinander unabhängigen Repositorien

und die Etablierung kontrollierter Zugangswege, einschließlich robuster Verifikations- und Authentifizierungssysteme, angemessen auf befugte Akteure zu beschränken.

D.1.3. Umsetzung von Schutzvorgaben gewährleisten und nachweisen

Datensouveränität setzt ein Miteinander von technischen und regulatorischen Standards voraus. In Anknüpfung an existierende Vorgaben zu *privacy by design* sollten Datenverarbeiter und Datennutzer noch stärker darauf achten, dass schon in der Planungs- und Entwicklungsphase datenschutzbezogene Erwägungen oberste Priorität besitzen. Es sollte zudem den betroffenen Einrichtungen (in der Forschung, in der medizinischen Praxis, oder im kommerziellen Bereich) obliegen, für ihren Verantwortungsbereich die Übereinstimmung mit den Datensouveränität sichernden Vorgaben nachzuweisen. In Anknüpfung an die diesbezüglich bestehenden Erfahrungen mit internen Datenschutzbeauftragten lässt sich deren Aufgaben- und Befugnisprofil sinnvoll in diese Richtung (*corporate data governance*) weiterentwickeln.

D1.4. Informationspflicht bei Pannen und Fehlverhalten etablieren

Es ist darauf zu achten, dass mögliche Pannen oder Fehlverhalten nicht verborgen bleiben, sondern in ihrer Relevanz für das Gesamtsystem verstanden und produktiv als Lerneffekt genutzt werden. Deshalb bedarf es einer entsprechenden Informationspflicht gegenüber den potenziell geschädigten Nutzern und – sofern diese nicht zu ermitteln sind – der Öffentlichkeit, sowie einer Meldepflicht gegenüber den Aufsichtsbehörden/-gremien.

D2. Kontrollmechanismen verbessern

D2.1. Datenschutzbeauftragte stärken

Zur Sicherstellung von Datensouveränität bedarf es einer Vielzahl interner (privater) und externer (hoheitlicher) Kontrollstellen. Deren Zuständigkeiten sollten genauer abgegrenzt und gegebenenfalls ihre Kapazitäten und Kompetenzen erweitert werden. Insbesondere ist es sinnvoll und geboten, die Tätigkeit der bestehenden Datenschutzbeauftragten – und zwar sowohl im öffentlichen wie im privaten Bereich – in Richtung Datensouveränität neu zu justieren und gegebenenfalls auszuweiten. Sie sollten die Arbeit von lokalen Kontrollinstanzen, wie etwa Forschungsethikkommissionen, ergänzen und auf der Grundlage transparenter Entscheidungskriterien in Konfliktsituationen moderierend und schlichtend wirken. Soweit sich die existierenden Kontrollstrukturen gegenüber den spezifischen Problemen von Big Data als unzulänglich erweisen, beispielsweise bei überregionalen und internationalen Verbundprojekten, ist eine stärkere Zentralisierung zu erwägen.

D2.2. Datenprüfer etablieren

Gerade mit Blick auf die als gesamtgesellschaftlich bedeutsame Datenqualität, insbesondere in der medizinbezogenen Forschung und klinischen Praxis, sollte eine entsprechende Prüfstruktur etabliert werden. Diese muss nicht notwendig rein hoheitlicher Natur sein, sondern ließe sich – etwa analog zum Abschlusswesen und zur Rechnungslegung im Gesellschaftsrecht – auch als private Regulierung konzipieren.

D2.3. Datentreuhandmodelle einführen

Um Vertrauen zu fördern und Missbrauch zu verhindern, sollten Datenverwender die technischen und organisatorischen Voraussetzungen dafür schaffen, dass Datenbestände nicht unmittelbar an sie selbst übergeben werden müssen, sondern Treuhandmodelle (zum Beispiel gemeinnützige Stiftungen) zwischengeschaltet werden können. Das kann nicht nur Machtungleichgewichte verringern, sondern auch Interessenkollisionen entgegenwirken. Zumindest im Bereich der medizinbezogenen Forschung und klinischen Praxis sollte politisch darauf hingewirkt werden, dass solche Modelle insbesondere auch in Bezug auf Datenverwender im internationalen Kontext (zum Beispiel Google, Apple, Facebook, Amazon und Microsoft) wirksam werden.

D3. Kodizes für Forschung, Klinik und Wirtschaft erarbeiten

Nach dem Vorbild bereits existierender Selbstverpflichtungen sollte konsequent weiter darauf hingewirkt werden, in allen datenschutzsensiblen Bereichen umfassende interne Verhaltensstandards zu etablieren. Dabei gilt es nicht nur die jeweiligen regulatorischen Vorgaben aufzunehmen und gegebenenfalls zu intensivieren, sondern auch – zumindest branchenintern oder mit Blick auf spezifische Anwendungsfelder – internationale Abstimmungen und Harmonisierungen anzustreben.

D4. Gütesiegel für Anbieter und Anwendungen unterstützen und ausbauen

Da eine besondere Berücksichtigung der Datensouveränität auch und gerade im Interesse der Datenverwender liegt, sollten entsprechende marktbasierte, teilweise bereits existierende Klassifizierungen („Gütesiegel“) unterstützt und ausgebaut werden. Über Mindeststandards setzende, zwingende gesetzliche Vorgaben hinausgehende Bemühungen können auf diese Weise zum profilbildenden Wettbewerbsfaktor avancieren. Soweit diese selbstregulativen Mechanismen sich als unzureichend erweisen, sind Koregulierungsmaßnahmen – etwa in Form von Zertifizierungen – einzubeziehen und die staatlichen Kontrollstrukturen einschließlich Haftungsregelungen zu verstärken.

D5. Kompetenz im verantwortungsvollen Umgang mit Daten für alle, die professionell mit Big Data zu tun haben, stärken

In Tätigkeitsfeldern, in denen Big Data rapide zunimmt, muss das Bewusstsein für die ethischen Herausforderungen und für die neuen Verantwortlichkeiten, die sich aus der Nutzung gesundheitsrelevanter Daten ergeben, befördert werden. Für einen solchen Kulturwandel ist bei allen Beteiligten ein besseres Verständnis von Forschungs- und Datenethik sowie wissenschaftstheoretische Reflexionskompetenz erforderlich. Die Förderung solcher Kompetenzen sollte daher verpflichtendes Element in der Aus-, Fort- und Weiterbildung in allen relevanten Fächern und Bereichen werden. Um der Komplexität und Bedeutung des Themas gerecht zu werden, könnten beispielsweise verstärkt betriebs- und institutionenintern Data-Science-Fachabteilungen eingerichtet werden.

Sondervotum

Analog zur medizinischen Ethik, die den Nutzen für das Individuum in den Mittelpunkt stellt und nach dem Grundsatz *nihil nocere* die Schadensabwehr in jedem einzelnen Fall zur obersten Maxime macht, gilt es auch im Umgang mit den Chancen und Risiken großer Datenmengen, die unveräußerlichen Rechte des Individuums und seine Selbstbestimmung als Maßstab für gesellschaftlichen Fortschritt zu nehmen. Diese Rechte stehen nicht im Widerspruch zum Gemeinwohl, sie sind vielmehr für einen freiheitlichen und sozialen Rechtsstaat konstitutiv. Die Bedürfnisse der (Gesundheits-)Wirtschaft nach immer umfassenderem Einblick in die Lebensäußerungen der Menschen sind dies nicht. Auf der anderen Seite birgt Big Data ein großes Potenzial. So können zum Beispiel Zusammenhänge von Gesundheit und ihren sozialen gesellschaftlichen Determinanten erkannt und neue Ansätze zur gesundheitsförderlichen Gestaltung verschiedenster Lebensbereiche erarbeitet werden.

Aber Big Data erweist sich erst dann als nutzbringend für die Gesundheitsvorsorge und die Medizin, wenn der oder die Einzelne als EigentümerIn seiner/ihrer personenbezogenen Daten zu jedem Zeitpunkt entscheiden kann, wem er oder sie diese in welchem Umfang auch im Falle der Sekundärnutzung offenlegen will.

Der Datenschutz bedarf daher einer präzisen gesetzlichen Regelung und das Bundesdatenschutzgesetz einer Präzisierung mit geeigneten Schutzmechanismen und Gestaltungsstrategien, also einer Bestätigung und Ausweitung, die Datensparsamkeit und Zweckbindung beinhalten. Diese funktionalen Instrumente gewährleisten einen Ausbau des Persönlichkeitsschutzes und des Datenschutzes und somit die Implementierung einer bestmöglichen Datensouveränität. Diese muss einen höheren Stellenwert auch gegenüber Forschungsinteressen behalten. Daher ist die Vorab-Analyse möglicher Folgen neuer Verfahren auf den Datenschutz und die informationelle Selbstbestimmung geboten (Datenschutz-Folgenabschätzung). Der Datenschutz und die damit verbundene Datensouveränität, die im Gegensatz zur Empfehlung B1.3 auch das Eigentum an personenbezogenen Daten und somit eine absolute Ausschlussmacht gegenüber Dritten bedeutet, ist ein sehr hohes Gut und deshalb auch regulatorisch und strafrechtlich abzusichern. Nur so kann die der informierten Einwilligung zugrunde liegende Selbstbestimmung gewährleistet werden.

Die technische Realisierung der Auswertung von Datenmassen muss rechtlich eingeschränkt werden, sodass Anwendungen möglich sind, jedoch personenbezogener Missbrauch verhindert wird. Analog dem Gendiagnostikgesetz muss es dezidierte Verbote von diskriminierenden Verwendungen personenbezogener Daten geben. Die Speicherung und Analyse personenbezogener Daten sollte daher nur im eng definierten Rahmen erlaubt sein. Missbräuchliche Datenzugriffe auch bei Sekundärnutzung müssen strafrechtlich sanktioniert werden. Entscheidend ist

eine finanziell effektive Ahndung von Verstößen, die wirksam abschreckt. Festzustellen ist weiterhin, dass es in diesem Bereich weniger ein Regeldefizit als ein massives Vollzugsdefizit gibt.

Ob Big Data in diesem Sinne im Gesundheitsbereich eher Chancen oder vermehrt Risiken bietet, entscheidet sich an folgenden Bedingungen:

- **Keine zentrale Speicherung von PatientInnendaten**

  PatientInnendaten sind prinzipiell nicht auf zentralen Servern, sondern auf dezentralen Speichermedien in der Hand der PatientInnen zu speichern. Ergänzend können die Daten verschlüsselt beim Hausarzt oder der Hausärztin gespeichert werden, die Entscheidung zur (auch begrenzten) Freigabe sollte zur Wahrung der PatientInneninteressen – wenn möglich – mit dem Hausarzt oder der Hausärztin zusammen getroffen werden.

  Im Gegensatz zur Empfehlung A4.1 dürfen jegliche personenbezogenen Daten, einschließlich derer, die mit der elektronischen Gesundheitskarte erhoben werden, daher nicht zentral und nicht ohne strikte individuelle Zustimmung im Falle der Primär- und Sekundärnutzung gespeichert werden. Das Einwilligungsmodell, das für die Erhebung personenbezogener Daten prinzipiell eine enge Zweckbindung voraussetzt, muss somit erhalten bleiben. Die Einführung eines Kaskadenmodells, das gemäß Empfehlung B2 eine breite Nutzungseinwilligung einschließen würde, ist daher abzulehnen.

  Zur Gewährleistung der maximalen Datensouveränität und Datensicherheit ist es notwendig, dass kleine und mittlere lokale Unternehmen sowie die kassenärztlichen Vereinigungen dies umsetzen.

- **Zustimmung der Versicherten hat Priorität vor anderen, auch vor Forschungsinteressen**

  In diesem Sinne muss eine Einsicht Dritter in die dezentralen Datenspeicher ohne Zustimmung der Versicherten auch für die Sekundärnutzung verboten bleiben. Daran sind auch der in A2 empfohlene Austausch und die Integration von gesundheitsrelevanten Daten zwischen vielfältigen institutionellen Akteuren zu messen. Technologische Sicherheitsvorkehrungen müssen getroffen werden, wie eine über die Empfehlung D.1.1 hinausgehende effektive Anonymisierung und Pseudoanonymisierung, die eine Reidentifizierung unmöglich macht.

- **Gesetzliche Regelungen statt freiwilliger Selbstkontrollen**

  Eine Absicherung darf nur durch gesetzliche Verpflichtung und nicht durch freiwillige Selbstkontrollen und Koregulierungsmaßnahmen, wie in Empfehlung D3 vorgeschlagen, geschehen. Selbstregulierung funktioniert nicht, denn durch Selbstregulierung entwickeln Industrieverbände ihre eigenen Kodizes und schaffen eigene (Pseudo-)Verfahren, um auf

Beschwerden zu reagieren. Selbstregulierungsgremien wird die Verantwortung dafür übertragen, dass ihre jeweiligen Mitglieder sich an die Regeln halten und dass bei Bedarf Abhilfemaßnahmen und Sanktionen angewendet werden.

Sich auf Koregulierungsmaßnahmen wie auf wirtschaftsinterne Kontrollmechanismen bei der Vergabe eines „Datengütesiegels" zu verlassen ist als naiv zu bezeichnen. Zum Beispiel zeigen die letzten Datenskandale im Sommer 2017 bei dem Datendienstleister Equifax mit 143 Millionen betroffenen Kunden das Versagen von internen Kontrollmechanismen.[365] PatientInnendaten sind zudem nach Warnungen des FBI auf dem Schwarzmarkt zehnmal teurer als Kreditkartennummern.[366] Bereits 2015 wurden beim zweitgrößten amerikanischen Gesundheitsdienstleister Anthem Inc. Gesundheitsdaten von 80 Millionen AmerikanerInnen von Hackern erbeutet.[367] Gesetzlich und regulatorisch müssen klare Rahmenbedingungen für persönliche Daten geschaffen werden, damit offensichtlich ist, was erlaubt und was Missbrauch ist.

Auf der technischen Ebene ist es ein Erfordernis des Datenschutzes, bereits in der Entwicklungsphase von Hardware, Software und Algorithmen Sicherheitslevel zu definieren und auf unterschiedliche Schutzgüter anzuwenden. Den Betroffenen muss die Einwilligung und ein „Recht auf Vergessen" eingeräumt werden, um die personenbezogenen Daten nachzuverfolgen und gegebenenfalls in jedem Fall und ausnahmslos löschen zu lassen. Ausnahmen, wie in Empfehlung B1.2 vorgeschlagen, sind nicht hinzunehmen. Entsprechend hierzu sind die Hersteller zu verpflichten, schon bei der Planung und Herstellung neuer Produkte darauf zu achten, dass den Grundprinzipien der Datensouveränität/des Datenschutzes ausnahmslos entsprochen wird (*privacy by design*) und somit ein kompletter Schutz der Privatsphäre gewährleistet ist. Die Erprobung von innovativen Konzepten in Kliniken ist vorrangig durch öffentliche Forschung zu gewährleisten.

- **Bedingungen für eine Datenspende zur Vermehrung des Gemeinwohls**

  Die Schaffung der rechtlichen Möglichkeit für Individuen, ihre Daten für medizinische Forschung zu spenden, sowie das Zur-Verfügung-Stellen von personenbezogenen Daten darf – im Gegensatz zur Empfehlung A4.2 – nur unter Wahrung eines strengen und effektiven Datenschutzes mit konkreter Zweckbindung geschehen. Dies ist zur Vermehrung des Gemeinwohls notwendig und ergänzt die Forderung, PatientInnendaten primär auf dezentralen Speichermedien in der Hand der PatientInnen zu speichern. Das Hasso-Plattner-Insitut

[365] Siehe https://investor.equifax.com/news-and-events/news/2017/09-07-2017-213000628 [09.11.2017].
[366] Siehe https://www.reuters.com/article/us-cybersecurity-hospitals/your-medical-record-is-worth-more-to-hackers-than-your-credit-card-idUSKCN0HJ21I20140924 [09.11.2017].
[367] Siehe https://www.forbes.com/sites/gregorymcneal/2015/02/04/massive-data-breach-at-health-insurer-anthem-reveals-social-security-numbers-and-more [09.11.2017].

strebt mit seiner Gesundheitscloud eine solche effektive Kontrolle über die eigenen Gesundheitsdaten an und beinhaltet die Möglichkeit zur altruistischen Spende von effektiv anonymisierten Daten für Forschungszwecke.[368]

- **Eigentum an Daten**

  Das Auskunftsrecht des Einzelnen über die zu seiner Person gespeicherten Daten reicht nicht mehr aus. Nur wenn transparent ist, welche Daten in die jeweiligen Auswertungen und Bewertungsprozesse einfließen, nach welchen Kriterien die Klassifikation erfolgt und wie sie Entscheidungen beeinflussen, lassen sich Aussagen zu deren Rechtmäßigkeit und ethischen Vertretbarkeit gewinnen. Anders als in Empfehlung B1.3 gefordert, ist ein „Eigentum" an Daten sicherzustellen, das eine Ausschlussmacht gegenüber Dritten beinhaltet.

Fazit: Sollte ein umfassender Datenschutz, die Umsetzung effektiver Anonymisierungs- und Pseudoanonymisierungsstandards und das Recht auf Vergessen nicht gewährleistet werden können, wäre ein Verzicht auf die Nutzung von Big Data zu Forschungszwecken oder anderen Anwendungen die notwendige Folge.

Christiane Fischer

[368] Siehe https://hpi.de/open-campus/hpi-initiativen/gesundheitscloud.html [09.11.2017].

Literaturverzeichnis

Acar, G. et al. (2014): The web never forgets. Persistent tracking mechanisms in the wild. In: Association for Computing Machinery (Hg.): Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. Scottsdale, 674-689.

Acquisiti, A.; Brandimarte, L.; Loewenstein, G. (2015): Privacy and human behavior in the age of information. In: Science, 347 (6221), 509-514.

Albrecht, U.-V. (Hg.) (2016): Chancen und Risiken von Gesundheits-Apps (CHARISMHA). http://www.digibib.tu-bs.de/?docid=00060000 [17.10.2017].

Amma, N. G. B. (2016): Big data mining. In: M. K. Singh; K. G. Dileep (Hg.): Effective Big Data Management and Opportunities for Implementation. Hershey, 53-59.

Andelfinger, V. P. (2016): Gesundheitsportale und Private Krankenversicherung – was in anderen Ländern passt, das passt auch in Deutschland – oder? In: Andelfinger, A.P.; Hänisch,T. (Hg.): eHealth. Wie Smartphones, Apps und Wearables die Gesundheitsversorgung verändern werden. Wiesbaden, 105-107.

Augsberg, I. (2015): Artikel 8 GRC. Schutz personenbezogener Daten. In: H. v. d. Groeben; J. Schwarze; A. Hatje (Hg.): Europäisches Unionsrecht (7. Aufl.). Baden-Baden, 142-158.

Augsberg, S. (2016): Big Data im Recht der Transplantationsmedizin - Vom „Ende der Theorie“ zum „Ende der Aporie“? In: Medizinrecht, 34 (9), 699-705.

Aumann, I.; Frank, M.; Pramann, O. (2016): Kapitel 12 - Gesundheits-Apps in der Gesetzlichen und Privaten Krankenversicherung. In: U.-V. Albrecht (Hg.): Chancen und Risiken von Gesundheits-Apps (CHARISMHA). http://www.digibib.tu-bs.de/?docid=00060000 [17.10.2017], 244-280.

Axer, P. (2011): Normengeflecht und Wissensrezeption in der gesetzlichen Krankenversicherung. In: Gesundheit und Pflege, 1 (6), 201-209.

Bachrach, P.; Baratz, M. S. (1962): Two faces of power. In: American Political Science Review, 56 (4), 947-952.

Barocas, S.; Selbst, A. D. (2016): Big data’s disparate impact. In: California Law Review, 104 (3), 671-732.

Baumgartner, U.; Gausling, T. (2017): Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen. Was Unternehmen jetzt nach der DS-GVO beachten müssen. In: Zeitschrift für Datenschutz, 7 (7), 308-313.

Bayertz, K. (1996): Staat und Solidarität. In: K. Bayertz (Hg.): Politik und Ethik. Stuttgart, 305-330.

Beauchamp, T. L.; Childress, J. F. (2001): Principles of Biomedical Ethics (5. Aufl.). New York.

Becker, M. (2017): Ein Recht auf datenerhebungsfreie Produkte. In: Juristenzeitung, 72 (4), 170-181.

Bedi, G. (2015): Automated analysis of free speech predicts psychosis onset in high-risk youths. In: npj Schizophrenia, 1, Art.-Nr.: 15030. DOI: 10.1038/npjschz.2015.30.

Bedi, G. et al. (2014): A window into the intoxicated mind? Speech as an index of psychoactive drug effects. In: Neuropsychopharmacology, 39 (10), 2340-2348.

Bennett, J. (2017): Your Face Belongs to Us. http://www.thedailybeast.com/how-facebook-fights-to-stop-laws-on-facial-recognition [17.10.2017].

Berdin, J. (2017): Biobank-Governance. Unter besonderer Berücksichtigung von Trust-Modellen, Baden-Baden.

Berlin, I. (1969): Four Essays on Liberty. Oxford.

Berndt, C. (2013): Resilienz. Das Geheimnis der psychischen Widerstandskraft. München.

Bertelsmann Stiftung (Hg.) (2017): Rethinking Privacy Self-Management and Data Sovereignty in the Age of Big Data. https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/rethinking-privacy-self-management-and-data-sovereignty-in-the-age-of-big-data [20.10.2017].

Binder, J.-H. (2012): Regulierungsinstrumente und Regulierungsstrategien im Kapitalgesellschaftsrecht. Tübingen.

Bleicher, A. (2017): Demystifying the black box that is AI. https://www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai [17.11.2017].

Bocanegra, C. L. S. et al. (2017): HealthRecSys: a semantic content-based recommender system to complement health videos. In: BMC Medical Informatics & Decision Making, 17 (1), 63.

Bodin, J. (1976): Über den Staat. München.

Böhme, G. (2008): Ethik leiblicher Existenz. Frankfurt am Main.

Bormann, F. J. (2006): Soziale Gerechtigkeit zwischen Fairness und Partizipation. John Rawls und die katholische Soziallehre. Freiburg im Breisgau.

Brandt, R. B. (1992): Morality, utilitarianism and rights. Cambridge.

Braun, M.; Dabrock, P. (2016): Ethische Herausforderungen einer sogenannten Big-Data basierten Medizin. In: Zeitschrift für medizinische Ethik, 62 (4), 313-329.

Britz, G. (2010): Informationelle Selbstbestimmung zwischen rechtswissenschaftlicher Grundsatzkritik und Beharren des Bundesverfassungsgerichts. In: W. Hoffmann-Riem (Hg.): Offene Rechtswissenschaft. Tübingen, 561-596.

Bröckling, U. (2017): Gute Hirten führen sanft. Über Menschenregierungskünste. Berlin.

Bublitz, J. C.; Merkel, R. (2014): Crimes against minds: on mental manipulations, harms and a human right to mental self-determination. In: Criminal Law and Philosophy, 8 (1), 51-77.

Buchner, B. (2017): Art. 22. Automatisierte Entscheidungen im Einzelfall einschließlich Profiling. In: J. Kühling; B. Buchner (Hg.): Datenschutz-Grundverordnung. München, 471-481.

Buchner, B. (2006): Informationelle Selbstbestimmung im Privatrecht. Tübingen.

Buchner, B.; Kühling, J. (2017): Art. 7. Bedingungen für die Einwilligung. In: J. Kühling; B. Buchner (Hg.): Datenschutz-Grundverordnung. München, 260-280.

Budin-Ljøsne, I. et al. (2017): Dynamic Consent: a potential solution to some of the challenges of modern biomedical research. In: BMC Medical Ethics, 18 (1), 4.

Bundesverband Digitale Wirtschaft (Hg.) (2017): Digitale Gesundheit. Die fünf großen Blockaden auf dem Weg zur digitalen Gesundheit – und wie wir sie überwinden können. Düsseldorf.

Buyx, A. (2010): Können, sollen, müssen? Public Health-Politik und libertärer Paternalismus. In: Ethik in der Medizin, 22 (3), 221-234.

Cartledge, C. (2017): Big Data. Data Wrangling Boot Camp. Big Data Vs. http://www.cs.odu.edu/~ccartled/Teaching/2017-Spring/DataWrangling/Presentations/030-bigDataVs.pdf [17.10.2017].

Caulfield, T. (2007): Biobanks and blanket consent: the proper place of the public good and public perception rationales. In: Kings Law Journals, 18 (2), 209-226.

Chadwick, R.; Berg, K. (2001): Solidarity and equity: new ethical frameworks for genetic databases. In: Nature Reviews Genetics, 2 (4), 318-321.

Chatziastros, A.; Drepper, J.; Semler, S. C. (2014): Big Data and Healthcare Analytics – datenschutzrechtliche Herausforderungen. In: mdi, 2/2014, 53-57.

Cisco (Hg.) (2017): The Zettabyte Era: Trends and Analysis. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.pdf [17.10.2017].

Condliffe, J. (2016): China turns big data into big brother. https://www.technologyreview.com/s/602987/china-turns-big-data-into-big-brother [17.11.2017].

Conley, A. et al. (2012) Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry. In: Maryland Law Review, 71 (3), 773-847.

Dabrock, P. (2012): Befähigungsgerechtigkeit. Ein Grundkonzept konkreter Ethik in fundamentaltheologischer Perspektive. Gütersloh.

Daniels, N. (1990): Equality of what: welfare, resources, or capabilities. In: Philosophy and Phenomenological Research, 50 (Suppl.), 273-296.

Datta, A.; Tschantz, M. C. Datta, A.; (2015): Automated experiments on ad privacy settings. A tale of opacity, choice, and discrimination. In: Proceedings on Privacy Enhancing Technologies, 1/2015, 92-112.

Davis-Turak, J. et al. (2017): Genomics pipelines and data integration: challenges and opportunities in the research setting. In: Expert Review of Molecular Diagnostics, 17 (3), 225-237.

Determann, L.; Weigl, M. (2016): EU-US-Datenschutzschild und Alternativen für internationale Datentransfers. In: Europäische Zeitschrift für Wirtschaftsrecht, 27 (21), 811-816.

Deutscher Bundestag (Hg.) (2017): Gesetzentwurf der Bundesregierung. Entwurf eines Gesetzes zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und -Umsetzungsgesetz EU – DSAnpUG-EU). BT-Drs. 18/11325. http://dipbt.bundestag.de/doc/btd/18/113/1811325.pdf [17.10.2017].

Deutscher Bundestag (Hg.) (2016): Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Maria Klein-Schmeink, Renate Künast, Dr. Konstantin von Notz, weiterer Abgeordneter und der Fraktion BÜNDNIS 90/DIE GRÜNEN. Drucksache 18/9058. Verhaltensbasierte Versicherungstarife – Apps und Wearables in der gesetzlichen Krankenversicherung. BT-Drs. 18/9243. http://dipbt.bundestag.de/doc/btd/18/092/1809243.pdf [17.10.2017].

Deutscher Bundestag (Hg.) (2014): Kleine Anfrage der Abgeordneten Harald Weinberg, Kathrin Vogler, Jan Korte, weiterer Abgeordneter und der Fraktion DIE LINKE. Datensammlungen über Versicherte in der privaten Krankenversicherung. BT-Drs. 18/3633. http://dip21.bundestag.de/dip21/btd/18/036/1803633.pdf [24.10.2017].

Deutscher Bundestag (Hg.) (2005): Gesetzentwurf der Fraktionen der CDU/CSU und SPD. Entwurf eines Gesetzes zur Verbesserung der Wirtschaftlichkeit in der Arzneimittelversorgung. BT-Drs. 16/194. http://dipbt.bundestag.de/doc/btd/16/001/1600194.pdf [17.10.2017].

Deutscher Ethikrat (Hg.) (2016): Patientenwohl als ethischer Maßstab für das Krankenhaus. Berlin.

Deutscher Ethikrat (Hg.) (2010): Humanbiobanken für die Forschung. Berlin.

Deutsches Institut für Vertrauen und Sicherheit im Internet (Hg.) (2017): Digitalisierung – Deutsche fordern mehr Sicherheit. Hamburg.

Dijck, J. v. (2016): Big data, grand challenges: on digitization and humanities research. In: KWALON, 21 (1), 8-18.

Do, C. B. et al. (2011): Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. In: PLoS Genetics, 7 (6), Art.-Nr.: e1002141. DOI: 10.1371/journal.pgen.1002141.

Dorloff, A. (2017): Strafen und Belohnen per Big Data. http://www.deutschlandfunk.de/ueberwachung-in-china-strafen-und-belohnen-per-big-data.697.de.html?dram:article_id=387147 [17.11.2017].

Dorniok, D. (2015): Die Funktionalität eines Rechts auf Nichtwissen. Wiesbaden.

Drexl, J. (2017): Neue Regeln für die Europäische Datenwirtschaft? Ein Plädoyer für einen wettbewerbspolitischen Ansatz. In: Neue Zeitschrift für Kartellrecht, 5 (8), 339-344.

Duttge, G. (2010): Das Recht auf Nichtwissen in der Medizin. In: Datenschutz und Datensicherheit, 34 (1), 34.

Engeler, M.; Felber, W. (2017): Entwurf der ePrivacy-VO aus Perspektive der aufsichtsbehördlichen Praxis. Reguliert der Entwurf an der technischen Realität vorbei? In: Zeitschrift für Datenschutz, 7 (6), 251-257.

Ensthaler, J. (2016): Industrie 4.0 und die Berechtigung an Daten. In: Neure Juristische Wochenschrift, (69) 48, 3473-3478.

Ernst, S. (2017): Die Einwilligung nach der Datenschutzgrundverordnung. In: Zeitschrift für Datenschutz, 7 (3), 110-114.

Esteva, A. et al. (2017): Dermatologist-level classification of skin cancer with deep neural networks. In: Nature, 542 (7639), 115-118.

Europäische Kommission (Hg.) (2017): Synopsis report of the public consultation on the evaluation and review of the eprivacy directive. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=40777 [17.10.2017].

Europäische Kommission (Hg.) (2016): Draft Code of Conduct on privacy for mobile health applications. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=16125 [17.11.2017]

Europäische Kommission (Hg.) (2015): Manual on borderline and classification in the Community Regulatory framework for medical devices. Version 1.17 (09-2015). https://www.emergogroup.com/sites/default/files/eu-manual-borderline-and-classification.pdf [17.10.2017].

Europäische Kommission (Hg.) (2014): Grünbuch über Mobile-Health-Dienste („mHealth").http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=5186 [17.11.2017].

Europäische Kommission (Hg.) (2012a): High-Performance Computing: Europe's place in a Global Race. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0045:FIN:EN:PDF [14.11.2017].

Europäische Kommission (Hg.) (2012b): Guidelines on the Qualification and Classification of Stand Alone Software Used in Healthcare within the Regulatory Framework of Medical Devices. https://ec.europa.eu/docsroom/documents/17921/attachments/1/translations/en/renditions/native [17.10.2017].

Europäische Kommission (Hg.) (2012c): Biobanks for Europe. A challenge for governance. http://www.irdirc.org/wp-content/uploads/2013/07/biobanks_for_Europe.pdf [17.10.2017].

European Data Protection Supervisor (Hg.) (2015): Mobile-Health-Dienste: Wie lassen sich technologische Innovation und Datenschutz miteinander vereinbaren? https://edps.europa.eu/sites/edp/files/publication/15-05-21_mhealth_de.pdf [17.10.2017].

Fachforum „Digitalisierung und Gesundheit" im Hightech-Forum (Hg.) (2017): Die Zukunft der Medizin ist digital. Szenario zur Digitalisierung in der Gesundheitsversorgung. München.

Fähnrich, C. et al. (2015): Surveillance and outbreak response management system (SORMAS) to support the control of the ebola virus disease outbreak in West Africa. In: Eurosurveillance, 20 (12), Art.-Nr.: pii=21071. DOI: 10.2807/1560-7917.ES2015.20.12.21071.

Feinberg, J. (1989): The Moral Limits of the Criminal Law Volume 3: Harm to Self. Oxford; NewYork; Toronto.

Fetzer, T. (2015): Plädoyer für ein neues Datenrecht. In: MultiMedia und Recht, 16 (12), 777-778.

Fezer, K.-H. (2017a): Dateneigentum der Bürger. Ein originäres Immaterialgüterrecht sui generis an verhaltensgenerierung Informationsdaten der Bürger. In: Zeitschrift für Datenschutz, 7 (3), 99-104.

Fezer, K.-H. (2017b): Dateneigentum. Theorie des immaterialgüterrechtlichen Eigentums an verhaltensgenerierten Personendaten der Nutzer als Datenproduzenten. In: MultiMedia und Recht, 18 (1), 3-5.

Floridi, L. (2014): The Fourth Revolution – How the infosphere is reshaping human reality. Oxford.

Forst, R. (2007): Das Recht auf Rechtfertigung: Elemente einer konstruktivistischen Theorie der Gerechtigkeit. Berlin.

Forst, R. (1996): Kontexte der Gerechtigkeit. Politische Philosophie jenseits von Liberalismus und Kommunitarismus. Frankfurt am Main.

Foucault, M. (1976): Überwachen und Strafen. Die Geburt des Gefängnisses. Frankfurt am Main.

Frackowiak, R.; Ailamaki, A.; Kherif, F. (2016): Federating and Integrating What We Know About the Brain at All Scales: Computer Science Meets the Clinical Neurosciences. Cham.

Francke, R.; Hart, D. (2008): Bewertungskriterien und -methoden nach dem SGB V. In: MedizinRecht, 26 (1), 2-24.

Friedrichsen, M.; Bisa, P.-J. (Hg.) (2016): Digitale Souveränität. Vertrauen in der Netzwerkgesellschaft. Wiesbaden.

Gärtner, A. (2010): MDD 2007/47/EG: Software als Medizinprodukt. http://www.e-health-com.eu/fileadmin/user_upload/dateien/Downloads/Gaertner-Medizinprodukte-Gesetz.pdf [17.10.2017].

Gassner, U. (2016): Software als Medizinprodukt – zwischen Regulierung und Selbstregulierung. In: Medizin Produkte Recht, 16 (4), 109-115.

Gassner, U. (2015): MedTech meets M-Health. In: Medizin Produkte Recht, 15 (3), 73-82.

Gellman, R.; Dixon, P. (2016): Failures of privacy self-regulation in the United States. In: D. Wright; P. De Hert (Hg.): Enforcing Privacy. Regulatory, Legal and Technological Approaches. Cham, 53-77.

Gerhardt, V. (1999): Selbstbestimmung. Das Prinzip der Individualität. Stuttgart.

Gesang, B. (2013): Rechte und Autonomie im Utilitarismus. In: H. Greif; M. G. Weiss (Hg.): Ethics, Society, Politics: Proceedings of the 35th International Wittgenstein Symposium, Kirchberg am Wechsel, Austria 2012 (Publications of the Austrian Ludwig Wittgenstein Society – New Series, Band 20). Berlin, 227-240.

Gesang, B. (2003): Eine Verteidigung des Utilitarismus. Stuttgart.

Gesang, B. (2000): Der Nutzenbegriff des Utilitarismus. In: Erkenntnis, 52 (3), 373-401.

Gethmann, C. F. (1996): Heilen: Können und Wissen. Zu den philosophischen Grundlagen der wissenschaftlichen Medizin. In: J. P. Beckmann (Hg.): Fragen und Probleme einer medizinischen Ethik. Berlin, 68-93.

GKV-Spitzenverband (Hg.) (2014): Stellungnahme des GKV-Spitzenverbandes vom 27.06.2014 zur öffentlichen Konsultation der Europäischen Kommission zum Grünbuch über Mobile-Health-Dienste (SWD (2014) 135 final). http://dsv-europa.de/lib/02_Positionspapiere/2014_GKV-SV_mHealth_de.pdf [17.10.2017].

Gossen, H.; Schramm, M. (2017): Das Verarbeitungsverzeichnis der DS-GVO. Ein effektives Instrument zur Umsetzung der neuen unionsrechtlichen Vorgaben. In: Zeitschrift für Datenschutz, 7 (1), 7-13.

Gräßer, F. et al. (2017): Therapy decision support based on recommender system methods. In: Journal of Healthcare Engineering, Art.-Nr.: 8659460. DOI: 10.1155/2017/8659460.

Greenstein, S.; Nagle, F. (2014): Digital dark matter and the economic contribution of Apache. In: Research Policy, 43 (4), 623-631.

Grimm, O.; Maiß, S. (2015): Bewerbersuche, Datenschutz und Headhunting. Social-Media-Recruiting. In: Arbeit und Arbeitsrecht, 70 (5), 270-272.

Habl, C. et al. (2016): Study on Big Data in Public Health, Telemedicine and Healthcare. Final Report. https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata_report_en.pdf. [17.10.2017].

Halpern, D. (2015): Inside the Nudge Unit. How small changes can make a big difference. London.

Hanfeld, M. (2017): Profitieren am Ende wieder nur Google und Co.? http://www.faz.net/aktuell/wirtschaft/die-reform-der-eprivacy-nutzt-am-ende-nur-google-und-co-15266269.html [21.11.2017].

Hart, D. (2000): Evidenz-basierte Medizin und Gesundheitsrecht – Überlegungen zu rechtlichen Konsequenzen der Verwissenschaftlichung der Medizin. In: MedizinRecht, 18 (1), 1-5.

Hauner, W. (2016): Big Data Analysis @ Munich Re. https://www.munichre.com/site/daktylos/get/documents_E912162360/mr/assets.dakylos/Documents/presentation/life-forum/Big-Data-Analytics-Munich-Re.pdf [17.10.2017].

Heidbrink, L.; Langbehn, C.; Loh, J. (2017): Handbuch Verantwortung. Wiesbaden.

Heimhalt, D.; Rehmann, W. (2014): Gesundheits- und Patienteninformationen via Apps. In: Medizin Produkte Recht, 8 (6), 197-205.

Helfrich, M. (2017): Art. 22. Automatisierte Entscheidungen im Einzelfall einschließlich Profiling. In: G. Sydow (Hg.): Europäische Datenschutzgrundverordnung. Baden-Baden, 559-574.

Heller, C. (2011): Post-Privacy. Prima leben ohne Privatsphäre. München.

Hellgardt, A. (2016): Regulierung und Privatrecht. Tübingen.

Henn, W. (2016): Prädiktive und individualisierte Genmedizin: Ethische und unternehmenspolitische Herausforderungen. Zeitschrift für Versicherungswesen, 8 (18), 563-566.

Hermstrüwer, Y. (2016): Informationelle Selbstgefährdung. Tübingen.

Heuvel, M. v. d. (2016): Gefährlicher Schatz. https://www.deutsche-apotheker-zeitung.de/news/artikel/2016/03/30/gefahrlicher-schatz [03.11.2017].

Hey, T.; Tansley, S.; Tolle, K. (Hg.) (2009): The Fourth Paradigm. Data Intensive Scientific Discovery. Redmond.

Hobbes, T. (1999): Leviathan - oder Stoff, Form und Gewalt eines kirchlichen und bürgerlichen Staates. Herausgegeben von I. Fetscher. Frankfurt am Main.

Hoffmann-Riem, W. (2000): Verwaltungsrecht in der Informationsgesellschaft – Einleitende Problemskizze. In: W. Hoffmann-Riem; E. Schmidt-Aßmann (Hg.): Verwaltungsrecht in der Informationsgesellschaft. Baden-Baden, 9-58.

Hofmann, J. M.; Johannes, P. C. (2017): DS-GVO: Anleitung zur autonomen Auslegung des Personenbezugs. In: Zeitschrift für Datenschutz, 7 (5), 221-226.

Horak, P. et al. (2017): Precision oncology based on omics data: the NCT Heidelberg experience. In: International Journal of Cancer, 141 (5), 877-886.

Hyman, J.; Steward, H. (Hg.) (2004): Agency and Action. Cambridge.

Institute for Government (Hg.) (2010): Mindspace. Influencing behavior through public policy. https://www.instituteforgovernment.org.uk/sites/default/files/publications/MINDSPACE.pdf [17.10.2017].

International Bioethics Committee (2017): Report of the IBC on Big Data and Health. http://unesdoc.unesco.org/images/0024/002487/248724E.pdf [17.11.2017].

Jones, D. T. W. et al. (2013): Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. In: Nature Genetics, 45 (8), 927-932.

Kamnitsas, K. et al. (2017): Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. https://arxiv.org/pdf/1612.08894.pdf [17.10.2017].

Kampert, D. (2017): Art. 8. Bedingungen für die Einwilligung eines Kindes in Bezug auf Dienste der Informationsgesellschaft. In: G. Sydow (Hg.): Europäische Datenschutzgrundverordnung. Baden-Baden, 381-388.

Kane, R. (Hg.) (2011): The Oxford Handbook of Free Will (2. Aufl.). Oxford; New York.

Karaalp, R. N. (2016): Der Schutz von Patientendaten für die medizinische Forschung in Krankenhäusern. Wiesbaden.

Katko, P.; Babaei-Beigi, A. (2014): Accountability statt Einwilligung? - Führt Big Data zum Paradigmenwechsel im Datenschutz? In: MultiMedia und Recht, 15 (6), 360-364.

Kaye, J. et al. (2015): Dynamic consent: a patient interface for twenty-first century research networks. In: European Journal of Human Genetics, 23 (2), 141-146.

Keppeler, L.; Berning, W. (2017): Technische und rechtliche Probleme bei der Umsetzung der DS-GVO-Löschpflichten. Anforderungen an Löschkonzepte und Datenbankstrukturen. In: Zeitschrift für Datenschutz, 7 (7), 314-318.

Keßler, O. (2017): Intelligente Roboter – neue Technologien im Einsatz. In: MultiMedia und Recht, 18 (9), 589-594.

Kettemann, M. C. (2015): Völkerrecht in den Zeiten des Netzes. Perspektiven auf den Schutz von Grund- und Menschenrechten in der Informationsgesellschaft zwischen Völkerrecht, Europarecht und Staatsrecht. Bonn.

Kim, Y.; Huang, J.; Emery, S. (2016): Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. In: Journal of Medical Internet Research, 18 (2), Art.-Nr.: 41. DOI: 10.2196/jmir.4738.

Kitchin, R. (2014): Big data, new epistemologies and paradigm shifts. In: Big Data & Society, 1 (1), 1-12.

Kitsiou, S. et al. (2017): Effectiveness of mHealth interventions for patients with diabetes: An overview of systematic reviews. In: PloS ONE, 12 (3), Art.-Nr.: e0173160. DOI: 10.1371/journal.pone.0173160.

Klein, R. A. (2016): Depotenzierung der Souveränität. Tübingen.

Klement, J. H. (2017): Öffentliches Interesse an Privatheit. Das europäische Datenschutzrecht zwischen Binnenmarkt, Freiheit und Gemeinwohl. In: JuristenZeitung, 72 (4), 161-170.

Kneer, G. (2012): Die Analytik der Macht bei Michel Foucault. In: P. Imbusch (Hg.): Macht und Herrschaft. Wiesbaden, 265-283.

Kollek, R. (2012): Systembiologie als Nexus zwischen Genen und Gesundheit? TA-Implikationen konzeptioneller Innovation in den Lebenswissenschaften. In: Technikfolgenabschätzung – Theorie und Praxis, 21 (2), 21-28.

Koppernock, M. (1997): Das Grundrecht auf bioethische Selbstbestimmung. Zur Rekonstruktion des allgemeinen Persönlichkeitsrechts. Baden-Baden.

Krajewski, J. et al. (2014): A phonetic approach for detecting sleepiness from speech in simulated air-traffic controller-communication. In: D. d. Waard et al. (Hg.): Human Factors of Systems and Technology. Maastricht, 147-155.

Kramer, A. D. I.; Guillory, J. E.; Hancock, J. T. (2014): Experimental evidence of massive-scale emotional contagion through social networks. In: Proceedings of the National Academy of Sciences of the United States of America, 111 (24), 8788-8790.

Kreße, B. (2017): Artikel 82. Haftung und Recht auf Schadensersatz. In: G. Sydow (Hg.): Europäische Datenschutzgrundverordnung. Baden-Baden, 1260-1268.

Krohm, N. (2016): Abschied vom Schriftformgebot der Einwilligung – Lösungsvorschläge und künftige Anforderungen. In: Zeitschrift für Datenschutz, 6 (8), 368-373.

Krügel, T. (2017): Das personenbezogene Datum nach der DS-GVO. Mehr Klarheit und Rechtssicherheit? In: Zeitschrift für Datenschutz, 7 (10), 455-459.

Krüger, P.-L. (2016): Datensouveränität und Digitalisierung. In: Zeitschrift für Rechtspolitik, 49 (7), 190-192.

Kucklick, C. (2016): Die granulare Gesellschaft. Wie das Digitale unsere Wirklichkeit auflöst. Berlin.

Kühling, J.; Seidel, C. (2015): Grundlagen – Allgemeiner Teil. In: T. Kingreen; J. Kühling (Hg.): Gesundheitsdatenschutzrecht. Baden-Baden, 29-185.

Ladeur, K.-H. (2016): „Big Data" im Gesundheitsrecht – Ende der „Datensparsamkeit"? In: Datenschutz und Datensicherheit, 40 (6), 360-364.

Laney, D. (2001): 3-D Data Management: Controlling Data Volume, Velocity and Variety. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf [17.10.2017].

Langanke, M. et al. (2013): Gesundheitliche Eigenverantwortung im Kontext Individualisierter Medizin. In: Ethik in der Medizin, 25 (3), 243-250.

Langkafel, P. (2015): Auf dem Weg zum Dr. Algorithmus? Potenziale von Big Data in der Medizin. In. Aus Politik und Zeitgeschichte, 65 (11-12), 27-32.

Langkafel, P. (Hg.) (2014): Big Data in Medizin und Gesundheitswirtschaft. Heidelberg.

Lauss, G. et al. (2011): Embracing complexity and uncertainty: An analysis of three orders of ELSA research on biobanks. In: Genomics, Society and Policy, 7 (1), 47-64.

LeCun, Y.; Bengio, Y.; Hinton, G. (2015): Deep learning. In: Nature, 521 (7553), 436-444.

Lek, M. et al. (2016): Analysis of protein-coding genetic variation in 60,706 humans. In: Nature, 536 (7616), 285-291.

Leong, L. et al. (2017): Magic Quadrant for Cloud Infrastructure as a Service, Worldwide. Herausgegeben von Gartner. https://www.gartner.com/doc/reprints?id=1-2G2O5FC&ct=150519&st=sb [17.10.2017].

Lewinski, K. v. (2016): Privacy Shield – Notdeich nach dem Pearl Harbor für die transatlantischen Datentransfers. Europarecht, 13 (4), 405-421.

Lewinski, K. v.; Herrmann, C. (2016): Cloud vs. Cloud – Datenschutz im Binnenmarkt. Verantwortlichkeit und Zuständigkeit bei grenzüberschreitender Datenverarbeitung. Zeitschrift für Datenschutz, 6 (10), 467-474.

Litjens, G. (2017): A survey on deep learning in medical image analysis. In: Medical Image Analysis, 42, 60-88.

Livingston, K. M. et al. (2013): Representing annotation compositionality and provenance for the semantic web. In: Journal of Biomedical Semantics, 4, Art.-Nr.: 38. DOI: 10.1186/2041-1480-4-38.

Lob-Hüdepohl, A. (2007): Schwierige Willensbekundung. Garantieren Patientenverfügungen würdevolles Sterben? In: Herder Korrespondenz, 61 (2), 83-87.

Lovett, F. (2007): Power. In: R. E. Goodin; P. Pettit; T. W. Pogge (Hg.): A Companion to Contemporary Political Philosophy (2. Aufl.). Malden (MA), 709-718.

Lucht, M.; Boeker, M.; Kramer, U. (2017): Gesundheits- und Versorgungsapps. Herausgegeben vom Universitätsklinikum Freiburg. http://www.tk.de/centaurus/servlet/contentblob/724464/Datei/143235/Studie-Gesundheits-und-Versorgungs-Apps.pdf [17.10.2017].

Lukes, S. (2005): Power. A Radical View (2. Aufl.). Basingstoke; New York.Martini, M. (2016): Do it yourself im Datenschutzrecht. In: Neue Zeitschrift für Verwaltungsrecht, 35 (6), 353-354.

Martini, M. (2014): Big Data als Herausforderung für den Persönlichkeitsschutz und das Datenschutzrecht. Deutsches Verwaltungsblatt, 129 (23), 1481-1489.

Martini, M.; Nink, F. (2017): Wenn Maschinen entscheiden … - vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz. In: Neue Zeitschrift für Verwaltung, 36 (10), 1-14.

Mayer-Schönberger, V.; Cukier, K. (2013): Big Data. A Revolution That Will transform How We Live, Work, and Think. Boston; London.

McGoogan, C. (2017): NHS illegally handed Google firm 1.6m patient records, UK data watchdog finds. http://www.telegraph.co.uk/technology/2017/07/03/googles-deepmind-nhs-misused-patient-data-trial-watchdog-says [17.11.2017].

Merkel, R. et al. (2007): Intervening in the Brain. Changing Psyche and Society. Berlin; Heidelberg.

Mill, J. S. (1859): On Liberty. London.

Molnár-Gábor, F.; Kaffenberger, L. (2017): EU-US-Privacy-Shield – ein Schutzschild mit Löchern? In: Zeitschrift für Datenschutz, 7 (1), 18-24.

Montjoye, Y.-A. d. et al. (2013): Unique in the crowd. The privacy bounds of human mobility. In: Scientific Reports, 3, Art.-Nr.: 1376. DOI: 10.1038/srep01376.

Mooy, M. D. (2017): Rethinking Privacy Self-Management and Data Sovereignty in the Age of Big Data. Herausgegeben von der Bertelsmann Stiftung. Gütersloh.

Mostert, M. et al. (2016): Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach. In: European Journal of Human Genetics, 24 (7), 956-960.

Mota, N. B. et al. (2012): Speech graphs provide a quantitative measure of thought disorder in psychosis. In: PLoS One, 7 (4), Art.-Nr.: e34928. DOI: 10.1371/journal.pone.0034928.

Müller, H.; Hanbury, A. (2016): Forschungsanwendungen in der digitalen Radiologie. In: Der Radiologe, 56 (2), 176-180.

Nakamoto, S. (2009): Bitcoin. A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf [17.10.2017].

Nationaler Ethikrat (Hg.) (2007): Prädiktive Gesundheitsinformationen beim Abschluss von Versicherungen. Berlin.

Nationaler Ethikrat (Hg.) (2005): Prädiktive Gesundheitsinformationen bei Einstellungsuntersuchungen. Berlin.

Nationaler Ethikrat (Hg.) (2004): Biobanken für die Forschung. Berlin.

Nationales Aktionsbündnis für Menschen mit Seltenen Erkrankungen (Hg.) (2016): Zwischenbericht zur Umsetzung des Nationalen Aktionsplans für Menschen mit Seltenen Erkrankungen. http://www.achse-online.de/de/was_tut_ACHSE/namse/pdf/nationaler_Aktionsplan/namse_monitoringbericht_2016.pdf [17.10.2017].

Nicholson Price II, W. (2015): Black-box medicine. In: Havard Journal of Law & Technology, 28 (2), 419-468.

Nissenbaum, H. (2011): A contextual approach to privacy online. In: Daedalus, 140 (4), 32-48.

Nissenbaum, H. (2009): Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford.

Nussbaum, M. (2015): Fähigkeiten schaffen: neue Wege zur Verbesserung menschlicher Lebensqualität. Freiburg.

Nussbaum, M. (2014): Die Grenzen der Gerechtigkeit. Behinderung, Nationalität und Spezienszugehörigkeit. Berlin.

Nussbaum, M. (1998): Gerechtigkeit oder Das gute Leben. Frankfurt am Main.

Obermeyer, Z.; Emanuel, E. J. (2016): Predicting the future - big data, machine learning, and clinical medicine. In: New England Journal of Medicine, 375 (13), 1216-1219.

OECD (Hg.) (2013): The OECD Privacy Framework. http://oecd.org/sti/ieconomy/oecd_privacy_framework.pdf [17.10.2017].

Oen, R. (2009): Software als Medizinprodukt. In: Medizin Produkte Recht, 3 (2), 55-57.

Ohm, P. (2015): Sensitive information. In: Southern California Law Review, 88 (5), 1125-1196.

Ohm, P. (2010): Broken promises of privacy: responding to the surprising failure of anonymization. In: UCLA Law Review, 57 (6), 1701-1777.

Ortner, R.; Daubenbüchel, F. (2016): Medizinprodukte 4.0 – Haftung, Datenschutz, IT-Sicherheit. In: Neue Juristische Wochenschrift, 69 (40), 2918-2924.

Paal, B.; Hennemann, M. (2017): Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen. In: Neue Juristische Wochenschrift, 70 (24), 1697-1701.

Pannenbecker, A. (2013): § 14 Grundzüge des Arzneimittel- und Medizinprodukterechts. In: M. Terbille; T. Clausen; J. Schroeder-Printzen (Hg.): Münchener Anwaltshandbuch Medizinrecht (2. Aufl.). München, 1165-1283.

Pearl, J. (2010): An introduction to causal inference. In: The International Journal of Biostatistics, 6 (2), 1-59.

Pelz, W. (2017): SWOT-Analyse. Definition, Beispiele und Vorlagen zum Erstellen einer SWOT-Analyse. http://www.wpelz.de/swot-analyse/SWOT-Analyse.pdf [17.10.2017].

Pillay, N. (2014): The right to privacy in the digital age. Report of the Office of the United Nations High Commissioner for Human Rights. UN-Dokument A/HRC/27/37. http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session27/Documents/A-HRC-27-37_en.doc [17.10.2017].

Ploug, T.; Holm, S. (2016): Meta consent – a flexible solution to the problem of secondary use of health data. In: Bioethics, 30 (9), 721-732.

Poelzig, D. (2012): Normdurchsetzung durch Privatrecht. Tübingen.

Polonetsky, J.; Tene, O. (2013): Privacy and big data. Making ends meet. In: Stanford Law Review, 66, Art.-Nr.: 25. SSRN: 2628412.

Prainsack, B.; Buyx, A. (2017): Solidarity in Biomedicine and Beyond. Cambridge.

Prainsack, B.; Buyx, A. (2016): Das Solidaritätsprinzip. Ein Plädoyer für eine Renaissance in Medizin und Bioethik. Frankfurt am Main.

Radic, D. et al. (2016): Big Data im Krankenversicherungsmarkt. Relevanz, Anwendungen, Chancen und Hindernisse. http://s.fhg.de/krankenversicherung [17.11.2017].

Rat für Informationsinfrastrukturen (Hg.) (2017): Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI). http://www.rfii.de/?wpdmdl=2269 [17.10.2017].

Rawls, J. (1972): A Theory of Justice. Oxford.

Regalado, A. (2017): EmTech: Illumina Says 228,000 Human Genomes Will Be Sequenced This Year. https://www.technologyreview.com/s/531091/emtech-illumina-says-228000-human-genomes-will-be-sequenced-this-year [17.10.2017].

Reich, N. (2014): Fehlerhaftigkeit von Medizinprodukten. Europäische Zeitschrift für Wirtschaftsrecht, 25 (23), 898-900.

Reinsel, D.; Gantz, J.; Rydning, J. (2017): Data Age 2025. The Evolution of Data to Life-Critical. Herausgegeben von IDC. http://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf [17.10.2017].

Rhees R. (Hg.) (1969): Schriften von Ludwig Wittgenstein. Band I: Tractatus logico-philosophicus; Tagebücher; Philosophische Untersuchungen. Frankfurt am Main.

Richter, G.; Buyx, A. (2016): Breite Einwilligung (broad consent) zur Biobank-Forschung – die ethische Debatte. In: Ethik in der Medizin, 28 (4), 311-325.

Riechert, A. (2016): Stellungnahme zu rechtlichen Aspekten eines Einwilligungsassistenten. https://stiftungdatenschutz.org/fileadmin/Redaktion/Bilder/Abschluss_Studie_30032017/stiftungdatenschutz_Stellungnahme_Rechtliche_Aspekte_eines_Einwilligungsassistenten_Anhang_1_final.pdf [17.10.2017].

Roberto, C. A.; Kawachi, I. (2015): Behavioral Economics and Public Health. Oxford.

Rorty, R. (1989): Contingency, Irony and Solidarity. Cambridge.

Roessler, J.;Eilan, N. (Hg.) (2003): Agency and Self-Awareness. Issues in Philosophie and Psychology. Oxford.

Rössler, B. (2001): Der Wert des Privaten. Berlin.

Roßnagel, A.; Nebel, M. (2015): (Verlorene) Selbstbestimmung im Datenmeer – Privatheit im Zeitalter von Big Data. Datenschutz und Datensicherheit, 39 (7), 455-459.

Rübsamen, K. (2015): Rechtliche Rahmenbedingungen für mobileHealth. In: MedizinRecht, 33 (7), 485-491.

Rudolf, W. (2011): § 90 - Recht auf informationelle Selbstbestimmung. In: D. Merten; H.-J. Papier (Hg.): Handbuch der Grundrechte. Band 4. Heidelberg, 233-290.

Samuelson, P. (1975): Foundations of Economic Analysis (10. Aufl.). Cambridge.

Schaar, K. (2017): Anpassung von Einwilligungserklärungen für wissenschaftliche Forschungsprojekte. Die informierte Einwilligung nach der DS-GVO und den Ethikrichtlinien. In: Zeitschrift für Datenschutz, 7 (5), 213-220.

Schaar, P. (2014): Überwachung total. Wie wir in Zukunft unsere Daten schützen. Berlin.

Schantz, P. (2016): Die Datenschutz-Grundverordnung – Beginn einer neuen Zeitrechnung im Datenschutzrecht. In: Neue Juristische Wochenschrift, 69 (26), 1841-1847.

Scherer, S. et al. (2013): Investigating voice quality as a speaker-independent indicator of depression and PTSD. https://pdfs.semanticscholar.org/d7dd/8fc189f93d9a14714582904dd9542f952f32.pdf [17.10.2017].

Schmitz, B.; Dall'Armi, J. v. (2017): Datenschutz-Folgenabschätzung – verstehen und anwenden. Wichtiges Instrument zur Umsetzung von Privacy by Design. In: Zeitschrift für Datenschutz, 7 (2), 57-63.

Schneider, J. (2017): Schließt Art. 9 DS-GVO die Zulässigkeit der Verarbeitung bei Big Data aus? Überlegungen, wie weit die Untersagung bei besonderen Datenkategorien reicht. In: Zeitschrift für Datenschutz, 7 (7), 303-307.

Schröder, P. (2011): Politische Strategien (2. Aufl.). Herausgegeben von der Friedrich-Naumann-Stiftung für die Freiheit. Potsdam.

Schroeder, A. (2015): Das Recht auf Nichtwissen im Kontext prädiktiver Gendiagnostik. Eine Studie zum ethisch verantworteten Umgang mit den Grenzen des Wissens. Wiesbaden.

Schwartz, P. M.; Solove, D. J. (2011): The PII problem: privacy and a new concept of personally identifiable information. In: New York University Law Review, 86 (6), 1814-1894.

Seemann, M. (2011): Vom Kontrollverlust zur Filtersouveränität. In: Heinrich-Böll-Stiftung (Hg.): #public_life. Digitale Intimität, die Privatsphäre und das Netz. Berlin, 74-80.

Seife, C. (27. November 2013): 23andMe is terrifying, but not for the reasons the FDA thinks. https://www.scientificamerican.com/article/23andme-is-terrifying-but-not-for-the-reasons-the-fda-thinks [17.10.2017].

Sen, A. (2009): The idea of justice. Cambridge.

Sen, A. (1979): Equality of what? In: S. McMurrin (Hg.): The Tanner Lectures on Human Values. Cambridge, 195-220.

Simitis, S. (2014a): § 4a. Einwilligung. In: S. Simitis (Hg.): Bundesdatenschutzgesetz (8. Aufl.). Baden-Baden, 470-503.

Simitis, S. (2014b): § 28. Datenerhebung und -speicherung für eigene Geschäftszwecke. In: S. Simitis (Hg.): Bundesdatenschutzgesetz (8. Aufl.). Baden-Baden, 1182-1271.

Sivridis, A.; Seidel, C.; Kühling, J. (2015): Datenschutzrecht (3. Aufl.). Heidelberg.

Smart, J. J. C. (1961): An outline of a system of utilitarian ethics. Melbourne.

Smart, J. J. C.; Williams, B. (1973): Utilitarianism For and Against. Cambridge.

Solove, D. J. (2011): Nothing to hide. The false tradeoff between privacy and security. New Haven.

Solove, D. J. (2008): Understanding Privacy. Cambridge.

Solove, D. J. (2007): "I've got nothing to hide" and other misunderstandings of privacy. In: San Diego Law Review, 44 (4), 745-772.

Solove, D. J.; Hartzog, W. (2014): The FTC and the New Common Law of Privacy. In: Columbia Law Review, 114 (3), 583-676.

Specht, L. (2017): Daten als Gegenleistung – Verlangt die Digitalisierung nach einem neuen Vertragstypus? In: JuristenZeitung, 72 (15-16), 763-770.

Spielkamp, M. (2017): Sind Algorithmen die besseren Richter? In: Technology Review, Heft 8/2017, 36-37.

Stallberg, C. (2010): Evidenzbasierte Medizin als Rechtsbegriff – Funktion, Inhalte, Grenzen. In: Pharma Recht, 32 (1), 5-12.

Stegmüller, W. (1983): Probleme und Resultate der Analytischen Philosophie. Band I (2. Aufl.). Berlin.

Steinsbekk, K. S.; Kåre Myskja, B.; Solberg, B. (2013): Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? In: European Journal of Human Genetics, 21 (9), 897-902.

Stephens, Z. D. et al. (2015): Big Data. Astronomical or genomical? In: PLoS Biology, 13 (7), Art.-Nr.: e1002195. DOI: 10.1371/journal.pbio.1002195.

Stone, M. L. (2014): Big Data for Media. Herausgegeben vom Reuters Institute for the Study of Journalism und der University of Oxford. http://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-04/Big%20Data%20For%20Media_0.pdf [17.10.2017].

Strategy&; PricewaterhouseCoopers (Hg.) (2016): Weiterentwicklung der eHealth-Strategie. Sudie im Auftrag des Bundesministeriums für Gesundheit. http://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/E/eHealth/BMG-Weiterentwicklung_der_eHealth-Strategie-Abschlussfassung.pdf [17.10.2017].

Struppek, D. (2010): Patientensouveränität im Pflegeheim - Möglichkeiten und Grenzen aus der Sicht von hochaltrigen, mehrfach erkrankten Pflegeheimbewohnern, ihren Ärzten, Pflegekräften und privaten Bezugspersonen. Berlin.

Taigman, Y. (2014): Web-scale training for face identification. http://www.cs.tau.ac.il/~wolf/deeplearningmeeting/pdfs/deepface_masterclass.pdf [17.10.2017].

Taigman, Y. et al. (2014): Deepface: Closing the gap to human-level performance in face verification. https://www.cs.toronto.edu/~ranzato/.../taigman_cvpr14.pdf [17.10.2017].

Tamm, M. (2011): Verbraucherschutzrecht. Europäisierung und Materialisierung des deutschen Zivilrechts und die Herausbildung eines Verbraucherschutzprinzips. Tübingen.

Tene, O.; Polonetsky, J. (2012): Privacy in the age of big data: a time for big decisions. In: Stanford Law Review Online, 64 (25), 63-69.

Teramoto, A. et al. (2017): Automated classification of lung cancer types from cytological images using deep convolutional neural networks. In: BioMed Research International, Art.-Nr.: 4067832. DOI: 10.1155/2017/4067832.

Thaler, R. H.; Sunstein, C. R. (2008): Nudge. Wie man kluge Entscheidungen anstößt. Berlin.

Thorbom, N. (2015): Datenschutz in der medizinischen Forschung. In: T. Kingreen; J. Kühling (Hg.): Gesundheitsdatenschutzrecht. Baden-Baden, 344-371.

Thüsing, G.; Schmidt, M.; Forst, G. (2017): Das Schriftformerfordernis der Einwilligung nach § 4a BDSG im Pendelblick zu Art. 7 DS-GVO. In: Recht der Datenverarbeitung, 33 (3), 116-222.

Thouvenin, F. (2017): Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumsbegriffs. In: Schweizerische Juristen-Zeitung, 113 (2), 21-32.

Towfigh, E. V.; Ulrich, J. (2017): Artikel 44. Allgemeine Grundsätze der Datenübermittlung. In: G. Sydow (Hg.): Europäische Datenschutzgrundverordnung. Baden-Baden. 862-872.

Tsanas, A. et al. (2012): Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. In: IEEE Transactions on Biomedical Engineering, 59 (5), 1264-1271.

Ulbricht, M.-R.; Weber, K. (2017): Adieu Einwilligung? Neue Herausforderungen für die informationelle Selbstbestimmung im Angesicht von Big Data-Technologien. In: M. Friedewald; J. Lamla; A. Roßnagel (Hg.): Informationelle Selbstbestimmung im digitalen Wandel. Wiesbaden, 265-286.

Verbraucherzentrale Nordrhein-Westfalen (Hg.) (2015): Bonusprogramme der gesetzlichen Krankenkassen. https://www.verbraucherzentrale.nrw/sites/default/files/migration_files/media236794A.pdf [24.10.2017].

Vodafone Institute for Society and Communications (Hg.) (2016): Big Data. A European Survey on the Opportunities and Risks of Data Analytics. http://www.vodafone-institut.de/wp-content/uploads/2016/01/VodafoneInstitute-Survey-BigData-en.pdf [17.10.2017].

Vossenkuhl, C. (2013): Der Schutz genetischer Daten. Unter besonderer Berücksichtigung des Gendiagnostikgesetzes. Berlin.

Walzer, M. (1983): Spheres of justice: a defense of pluralism and equality. New York.

Wartenberg, T. E. (1990): The Forms of Power: From Domination to Transformation. Philadelphia.

Warren, S. D.; Brandeis, L. D. (1890): The right to privacy. In: Harvard Law Review, 4 (5), 193-220.

Weber, M. (1980): Wirtschaft und Gesellschaft. Tübingen.

Wee, R. (2013): Dynamic consent in the digital age of biology. In: Journal of Primary Health Care, 5 (3), 259-261.

Weichert, T. (2014a): Medizinisches Big Data und Datenschutz. In: P. Langkafel (Hg.): Big Data in Medizin und Gesundheitswirtschaft. Heidelberg, 161-174.

Weichert, T. (2014b): Big Data, Gesundheit und der Datenschutz. In: Datenschutz und Datensicherheit, 38 (12), 831-838.

Weichert, T. (2013): Big Data und Datenschutz. Chancen und Risiken einer neuen Form der Datenanalyse. In: Zeitschrift für Datenschutz, 3 (6), 251-259.

Weilert, A. K. (2015): Gesundheitsverantwortung zwischen Markt und Staat. Baden-Baden.

Weitzel, T. et al. (2016): Active Sourcing und Social Recruiting. https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/wiai_lehrstuehle/isdl/Recruiting_Trends_2016_-_Active_Sourcing_und_Social_Recruiting_v_WEB.PDF [17.10.2017].

Werkmeister, C.; Brandt, E. (2016): Datenschutzrechtliche Herausforderungen für Big Data. In: Computer und Recht, 32 (4), 233-238.

Wiebe, A. (2017): Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken. In: Gewerblicher Rechtsschutz und Urheberrecht, 119 (4), 338-345.

Wiebe, A.; Schur, N. (2017): Ein Recht an industriellen Daten im verfassungsrechtlichen Spannungsverhältnis zwischen Eigentumsschutz, Wettbewerbs- und Informationsfreiheit. In: Zeitschrift für Urheber- und Medienrecht, 61 (6), 461-476.

Wigge, P. (2000): Evidenz-basierte Richtlinien und Leitlinien – Qualitäts- oder Steuerungsinstrumente in der GKV? In: MedizinRecht, 18 (12), 574-585.

Wilbanks, J. T.; Topol, E. J. (2016): Stop the privatization of big data. In: Nature, 353 (7612), 345-348.

Williams, H. et al. (2015): Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. In: JMIR Medical Informatics, 3 (1), Art.-Nr.: e3. DOI: 10.2196/medinform.3525.

Wirth, J. (2015): Goldgrube Big Scientific Data. In: GIT Spezial BIOforum, 38 (2), 24-26.

Wright, D.; De Hert, P. (Hg.) (2016): Enforcing Privacy. Regulatory, Legal and Technological Approaches. Cham.

Zhelyazkova, A. et al. (2017): Munich Digital Healthcare Footprint. Food, Wellbeing & Trade. Herausgegeben vom Munich Digital Institute. https://www.munich-digital.com/download/article/?token=c461296644ad0420c35ceee5fd8229840d85ef9c [12.9.2017].

Zirfas, J. (2017): Smart Health rechtsverträglich gestaltet: Ubiquitous Computing in der Gesundheitspflege und -vorsorge. Wiesbaden.

Zook, M. et al. (2017): Ten simple rules for responsible big data research. In: PLoS Computational Biology, 13 (3), Art.-Nr.: e1005399. DOI: 10.1371/journal.pcbi.1005399.

Entscheidungsverzeichnis

BVerfG, 15.12.1983 – 1 BvR 209/83, 1 BvR 269/83, 1 BvR 362/83, 1 BvR 420/83, 1 BvR 440/83, 1 BvR 484/83 (BVerfGE 65, 1)

BVerfG, 11.03.2008 – 1 BvR 2074/05, 1 BvR 1254/07 (BVerfGE 120, 378)

BVerfG, 26.02.2008 – 1 BvR 1602/07, 1 BvR 1606/07, 1 BvR 1626/07 (BVerfGE 120, 180)

BVerfG, 11.06.1991 – 1 BvR 239/90 (BVerfGE 84, 192)

BVerfG, 17.07.2013 – 1 BvR 3167/08 (NJW 2013, 3086)

EuGH, 19.10.2016 – C-582/14 (NJW 2016, 3579)

BGH, 28.01.2014 – VI ZR 156/13 (NJW 2014, 1235)

BGH, 18.04.2013 – I ZR 53/09 (NJW-RR 2014, 46)

VG Darmstadt, 24.06.2004 – 1 E 470/04 (3) (NVwZ-RR 2006, 566)

Abkürzungsverzeichnis

a. F.	alte Fassung
AA	Akademieausgabe
ABl. EG	Amtsblatt der Europäischen Gemeinschaften
ABl. EU	Amtsblatt der Europäischen Union
Abs.	Absatz
AEUV	Vertrag über die Arbeitsweise der Europäischen Union
AGB	Allgemeine Geschäftsbedingungen
Art.	Artikel
Aufl.	Auflage
BDSG	Bundesdatenschutzgesetz
BGB	Bürgerliches Gesetzbuch
BGBl.	Bundesgesetzblatt
BGH	Bundesgerichtshof
BMBF	Bundesministerium für Bildung und Forschung
BT-Drs.	Bundestagsdrucksache
BVerfG	Bundesverfassungsgericht
BVerfGE	Entscheidungen des Bundesverfassungsgerichts
bzw.	beziehungsweise
ca.	circa
CDISC	Clinical Data Interchange Standards Consortium
CT	Computertomografie
DSGVO	Datenschutz-Grundverordnung
ebd.	ebenda
EDSB	Europäische Datenschutzbeauftragter
EDV	Elektronischer Datenverarbeitung
EGA	European Genome-phenome Archive
EMRK	Europäische Menschenrechtskonvention
engl.	englisch
EU	Europäische Union
EuGH	Europäischer Gerichtshof

Fn.	Fußnote
GenDG	Gendiagnostikgesetz
GG	Grundgesetz
GKV	gesetzliche Krankenversicherung
GRC	Charta der Grundrechte der Europäischen Union
HTTP	Hypertext Transfer Protocol
ID	Identifikator
IT	Informationstechnik
lit.	littera (Buchstabe)
m. w. N.	mit weiteren Nachweisen
MPG	Medizinproduktegesetz
n. F.	neue Fassung
NCT	Nationales Centrum für Tumorerkrankungen
NJW	Neue Juristische Wochenschrift
NJW-RR	Neue Juristische Wochenschrift-Rechtsprechungs-Report
Nr.	Nummer
NVwZ-RR	Neue Zeitschrift für Verwaltungsrecht-Rechtsprechungs-Report
OECD	Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
PKV	private Krankenversicherung
RFID	radio-frequency identification
Rn.	Randnummer
S.	Satz
SGB	Sozialgesetzbuch
SSO	single sign-ons
StGB	Strafgesetzbuch
UK	United Kingdom (Vereinigtes Königreich)
UrhG	Urheberrechtsgesetz
US	United States (Vereinigte Staaten)
usw.	und so weiter
VG	Verwaltungsgericht
vgl.	vergleiche
WPS	Wi-Fi Positioning System

Mitglieder des Deutschen Ethikrates

Prof. Dr. theol. Peter Dabrock (Vorsitzender)
Prof. Dr. med. Katrin Amunts (Stellvertretende Vorsitzende)
Prof. Dr. phil. Dr. h. c. Dipl.-Psych. Andreas Kruse (Stellvertretender Vorsitzender)
Prof. Dr. med. Claudia Wiesemann (Stellvertretende Vorsitzende)

Constanze Angerer
Prof. Dr. iur. Steffen Augsberg
Prof. Dr. theol. Franz-Josef Bormann
Prof. Dr. med. Alena M. Buyx
Prof. em. Dr. iur. Dr. h. c. Dagmar Coester-Waltjen
Dr. med. Christiane Fischer
Prof. em. Dr. phil. habil. Dr. phil. h. c. lic. phil. Carl Friedrich Gethmann
Prof. Dr. rer. nat. Dr. phil. Sigrid Graumann
Bischof Prof. Dr. theol. Martin Hein
Prof. Dr. med. Wolfram Henn
Prof. Dr. iur. Wolfram Höfling
Prof. Dr. (TR) Dr. phil. et med. habil. Ilhan Ilkilic
Prof. Dr. rer. nat. Ursula Klingmüller
Stephan Kruip
Prof. Dr. phil. Adelheid Kuhlmey
Prof. Dr. med. Leo Latasch
Prof. Dr. iur. Dr. h. c. Volker Lipp
Prof. Dr. theol. Andreas Lob-Hüdepohl
Prof. em. Dr. iur. Reinhard Merkel
Prof. Dr. phil. Gabriele Meyer
Prof. Dr. med. Elisabeth Steinhagen-Thiessen
Dr. phil. Petra Thorn