



Indra Spiecker genannt Döhmann

6.1 Künstliche Intelligenz und BigData im Spannungsfeld von Wissenschaftsfreiheit und Forschungsevaluation

Forschung an und mit Daten und auf der Basis von Informationen zur Wissensgewinnung ist allgegenwärtig, nicht erst seit Künstliche Intelligenz und Big Data (siehe dazu Kap. 1 und 2 in diesem Band)¹ mediale Aufmerksamkeit und Forschungsetats treiben. Wird mit Daten geforscht, stehen rechtlich vor allem Belange der

Dieses Kapitel entstand durch intensive Diskussion und Zusammenarbeit mit der interdisziplinären Projektgruppe „Digitale Arbeitswelten“, der Carl Friedrich Gethmann, Peter Buxmann, Julia Distelrath, Bernhard G. Humm, Stephan Lingner, Verena Nitsch, Jan C. Schmidt und Indra Spiecker genannt Döhmann angehörten. In besonderer Weise hat der Text Schärfungen erfahren durch die Überlegungen von Carl Friedrich Gethmann und C. Jan Schmidt. Danke gebührt außerdem meinem Mitarbeiter Ass. iur. Jeremy Stevens für Input zur allgemeinen Regulierung von Künstlicher Intelligenz und stud. iur. Sophie Danner für Unterstützung bei Formalia und Recherche.

¹Die Diskussion um die Abgrenzung der Begrifflichkeiten von Daten, Information und Wissen ist ubiquitär und weiterhin nicht abgeschlossen. In diesem Text werden im Wesentlichen Daten und Information als gleich verstanden. Am Einführungssatz wird aber deutlich, dass grundsätzlich ein Unterschied besteht: Daten sind die Basis von Informationen, die den Kontext miteinbeziehen; Wissen ist die gezielte Verwendung von Informationen, z. B. durch Rekombination. Vgl. dazu Spiecker genannt Döhmann 2010, S. 250 ff.

Rechtlich wird allenfalls zwischen Daten und Information begrifflich unterschieden, wobei z. B. die Europäische Datenschutzgrundverordnung als wesentliches Datenschutzrecht oder die Informationsfreiheitsgesetze eine Vermengung der Begriffe von Datum und Information praktizieren und beides erfassen, vgl. Hamer 2020, S. 40; Karg 2019, Art. 4 Nr. 1 Rn. 25.

I. Spiecker genannt Döhmann (✉)
Goethe Universität Frankfurt a.M., Frankfurt, Deutschland
E-Mail: spiecker@jur.uni-frankfurt.de

Privatheit und des Datenschutzes im weiteren Sinne im Raum, rechtlich abgesichert etwa durch Datenschutzrecht, Urheberrecht, Betriebs- und Geschäftsgeheimnisrecht oder in jüngerer Zeit auch Wettbewerbsrecht. Zudem treten – rechtlich bisher kaum erfasst – weitere Problembereiche wie die Qualitätssicherung, Archivierung und Zugänglichkeit von Datenbeständen, die Zugehörigkeit von Daten zu Forscher/innen bzw. deren Forschungsinstitutionen oder die Teilbarkeit und Übertragbarkeit von Rohdaten in das Sichtfeld regulatorischer Überlegungen. Eingebettet ist die Suche nach Informationen zudem fast immer in die Problematik der Herstellung, Darstellung und Kontrolle von privaten und staatlichen Entscheidungen unter Unsicherheit (siehe zu letzterem etwa Spiecker genannt Döhmman 2021). Denn Informationsgewinnung, wie sie im Mittelpunkt des Einsatzes von Künstlicher Intelligenz in der Forschung steht, verweist auf den damit in Aussicht gestellten erwarteten Erkenntnisgewinn und die (vermeintlich) besser informierte Entscheidungslage. Das aber ruft Fragen auf, wie mit Informationen verschiedener Quellen und verschiedenartiger Bearbeitungsdichte und erst recht verschiedenartiger Gewissheit umzugehen ist, gerade dann, wenn staatliche Entscheidungen auf dieser Basis getroffen werden. Regulatorische Eingriffe in die Gewinnung, Nutzung und Verbreitung von Daten und des Einsatzes von Künstlicher Intelligenz können also aus vielerlei Perspektiven geboten sein.

Blickt man, wie es dieser Beitrag tut, in besonderem Maße auf die Wissenschaftler/innen und ihre Nutzung von Daten und dies im besonderen Anwendungsfeld der Forschung mit und an Künstlicher Intelligenz, ist zusätzlich die Forschungsfreiheit als ein zentraler rechtlicher Belang zu beachten. Dieser wird bisher nur wenig berücksichtigt in der Auseinandersetzung mit Künstlicher Intelligenz; ein Forschungsfreiheitsgesetz oder wenigstens allgemeine Richtlinien, die verfassungsrechtliche Vorgaben konkretisierten und Kollisionen mit anderen Rechtspositionen klärten, fehlen sowohl auf nationaler als auch auf internationaler/europäischer Ebene und selbst Leitlinien sind derzeit Mangelware. Auch der Entwurf einer KI-Verordnung, wie sie die EU jüngst vorgelegt hat (Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence – Artificial Intelligence Act), COM/2021/206 final), konzentriert sich auf die Anwendungsfelder und bezieht die eigentliche Forschungsperspektive kaum dezidiert ein.

Methoden des Machine Learnings und der Künstlichen Intelligenz werden in einer Vielzahl von Forschungsdisziplinen genutzt. Sie ermöglichen die Verarbeitung großer, bisher häufig als unbewältigter eingeschätzter Datenmengen, um hierdurch ggf. neue und unerwartete Muster und Zusammenhänge zu finden. Hierzu zählt beispielsweise die Suche nach sowie die Prognose der Eigenschaften neuer Materialien/Werkstoffe oder Arzneimittel, aber ebenso die Analyse von archäologischen Fundstücken bis hin zur Vorhersage menschlichen Verhaltens, z. B. bei Wahlen oder der Partner/innensuche. Die spezifische Art und Weise der Nutzung dieser Technologien unterscheidet sich je nach Anwendungsfall. Zum Teil erleichtert die Automation bestimmte Arbeitsschritte, vor allem die Untersuchung einer deutlich größeren Zahl an Datensätzen mit der Entwicklung erster Hypothesen, die dann von den Forschern ausgewertet, weiter bearbeitet und verfeinert werden können. Zum Teil basieren die Forschungsergebnisse jedoch derart auf den verwendeten Algorithmen, dass eine wesentliche Forschungsleistung vom algorithmischen System übernommen wird. Hier zeigt sich beim Einsatz von Künstlicher Intelligenz häufig das Problem, dass die Reproduzierbarkeit der Forschung erschwert wird, weil ein exakt gleich aufgebautes algorithmisches System mit ähnlichen Trainings- und Entwicklungsdaten nicht gegeben ist.

Die Verbreitung des Einsatzes von Methoden der Künstlichen Intelligenz und von Big Data in der Wissenschaft führt dazu, dass sich auch die Einschätzung von Wissenschaft verändert. Darauf basierende Methodik kommt zu anderen, bereits existenten und anerkannten Methoden der Beurteilung, Begutachtung und Bewertung hinzu oder ersetzt sogar bereits bestehende Vorgehensweisen. Zudem verändern sich die Möglichkeiten und das Spektrum der Beurteiler: Öffentlichkeit, Gesellschaft, Wirtschaft, Verlage, Publikationsorgane, Politik, die eigenen Community oder Forschungsinstitutionen benötigen möglicherweise andere Herangehensweisen, um die Qualität und Bedeutung von Wissenschaft unter Bedingungen der Künstlichen Intelligenz überhaupt einschätzen zu können, erhalten aber möglicherweise gleichzeitig auch andere Grundlagen und Ausgangspunkte zur Bewertung.

Dieser erste allgemeine Befund lässt erkennen, dass sich Wissenschaftler/innen nicht nur mit der Methodik von Künstlicher Intelligenz selbst befassen müssen, um sie überhaupt korrekt anwenden und für ihre Zwecke nutzen zu können, sondern dass sie ihrerseits auch das Verhältnis des Einsatzes von Künstlicher Intelligenz/Big Data im Verhältnis zu anderer Forschungsmethodik begründen und möglicherweise auch rechtfertigen müssen. Dies gilt ebenso auch für den Fall, dass sie dies *nicht* tun. Die Methodenfrage wird also erweitert.

Diese Methodenfrage stellt sich zum einen im Internum der Forschung, also bei der Ausgestaltung des eigentlichen Forschungsansatzes. Gewinnt die Verfolgung einer Forschungsfrage in besonderer Weise durch den Einsatz von Künstlicher Intelligenz? Oder sind die Kosten-Nutzen-Relationen zu ungünstig, die Ergebnisse dadurch besonderen Belastungen ausgesetzt, die zu erwartende Unsicherheit und die Schwierigkeiten der angemessenen Darstellung und Rezeption der Ergebnisse womöglich noch größer?

Diese Methodenfrage stellt sich aber auch im Externum der Forschung, also bei der Beurteilung von Forschungsleistungen und Forschungsaussagen. Diese erfolgt typischerweise wissenschaftsintern, also beispielsweise bei der Beurteilung von Publikationen durch Schriftleiter/Redakteure oder im begutachtenden Peer Review Verfahren. Zudem aber ist Forschung auch zunehmend ein Beurteilungsfaktor über den Forschenden selbst, also z. B. bei der Beantragung und Bewilligung von Drittmitteln, bei der Bewertung für wissenschaftsrelevante Tätigkeiten oder aber bei der Beurteilung seiner selbst im Rahmen der modernen Bezügeausgestaltung unter Leistungsanreizen der ökonomisierten Hochschule. Überall hier gilt es den Forschenden, seine Forschungsleistung und seinen Forschungsansatz – unter Einsatz von Künstlicher Intelligenz oder gerade nicht – adäquat zu beurteilen. Die Entscheidung über die Nutzung oder Nicht-Nutzung von Künstlicher Intelligenz und deren konkrete Ausgestaltung wird zu einem neuen Beurteilungsfaktor von Wissenschaft.

6.2 Regulierungsansätze von Künstlicher Intelligenz

So allgegenwärtig wie Künstliche Intelligenz derzeit in Wissenschaft, Wirtschaft und Öffentlichkeit zu sein scheint, sind konkrete Regulierungsansätze für den Bereich von Forschung und Wissenschaft dennoch noch wenige zu konstatieren. Dieser Befund erstreckt sich auch auf die Lehre und Bildung: Die Vermittlung von Kenntnissen zu deren

Einsatz verläuft weitgehend ohne Vorgaben und fehlt in den meisten Ausbildungsprofilen und Studiengängen. Dagegen gibt es bereits eine Reihe von Empfehlungen, Leitlinien, Vorschlägen und Einschätzungen verschiedenster Fachgesellschaften, Kommissionen, Verbänden und Gremien zu Einsatz und Nutzung von Künstlicher Intelligenz im Allgemeinen bis hin zum weltweit ersten Entwurf einer EU-Verordnung für Künstliche Intelligenz (Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence – Artificial Intelligence Act), COM/2021/206 final). Es ist, auch wenn das bisher noch kaum gesehen wird, zu erwarten, dass mit dem verbreiteten Einsatz von Künstlicher Intelligenz in vielen Lebensbereichen auch der Forschungsbereich zunehmend stärkerer Beobachtung und in der Folge auch stärkerer Setzung von Leitplanken ausgesetzt sein wird. Daher wird hier der Rechtsrahmen solcher potenziellen Regulierungsanstrengungen speziell für die Wissenschaft untersucht.

Der Regulierung, also der Beschränkung von Wissenschaft und Forschung, sind im nationalen Recht durch Art. 5 Abs. 3 S. 1 Grundgesetz (GG) bzw. auf europarechtlicher Ebene Art. 13 Grundrechte-Charta (GRCh) enge Grenzen gesetzt (siehe dazu Abschn. 6.4.2). Die Einschränkung der Forschungsfreiheit bedarf daher einer überzeugenden Begründung unter Einbeziehung entgegenstehender, überwiegender verfassungsrechtlich abgesicherter Rechtspositionen. Typische Fälle anerkannter forschungsrelevanter Regulierung sind die allgemeinen Anforderungen an eine ethische Forschung, Vorgaben hinsichtlich der Nutzung bestimmter als gefährlich eingestufte Materialien und Stoffe sowie die Regulierung aus ethischen Gesichtspunkten, insbesondere im Hinblick auf Forschung am Menschen und genetischen Materialien.

Forschungsvorhaben können in bestimmten Fällen unmittelbar Risiken für das Leben bzw. andere Rechtsgüter Einzelner oder größeren Gruppen von Personen bedeuten. So können bei Experimenten mit Gefahrstoffen, wie bestimmten Chemikalien oder radioaktiven Stoffen, negative Auswirkungen für die Beteiligten, aber auch für Dritte, entstehen. Verschiedene Vorschriften regeln daher den Umgang mit diesen Stoffen, die auch auf Forschungsvorhaben Anwendung finden (z. B. die deutsche Verordnung zum Schutz vor Gefahrstoffen (Gefahrstoffverordnung – GefStoffV), die europäische REACH-Verordnung für Chemikalien (Verordnung (EG) Nr. 1907/2006 zur Registrierung, Bewertung, Zulassung und Beschränkung chemischer Stoffe (REACH)) oder das deutsche Chemikaliengesetz (Gesetz zum Schutz vor gefährlichen Stoffen (Chemikaliengesetz – ChemG)).

Im Hinblick auf die Forschung an Menschen und Tieren spielen ethische Gesichtspunkte eine besondere Rolle. Diese Vorstellungen sind häufig auch rechtlich im Schutz der Würde des Menschen, Art. 1 Abs. 1 GG, seiner Persönlichkeitsrechte, Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG, des Schutzes von Leib und Leben, Art. 2 Abs. 2 GG, oder dem Tierschutz, Art. 20a GG, verankert. Der Schutz der Grundrechte sowie der Tierschutz erfordern insofern verfahrensbezogene sowie inhaltliche, materiell-rechtlich wirkende Vorgaben zum Ausgleich der verschiedenen Rechtsgüter. Dies wird z. T. durch Einschränkung bestimmter Forschung, etwa von Tierversuchen nur zu bestimmten Zwecken und nach Einhaltung strikter Vorgaben, oder durch Vorgaben zur Begleitung und zum Monitoring dieser Forschung, etwa im Bereich der Arzneimittelzulassung für klinische Versuche, erreicht. Vielfach sind zudem weitere Sicherungen vorgesehen, z. B. Ethikkommissionen, die Forschungsvor-

haben proaktiv beurteilen und kontinuierlich begleiten sollen. Daneben erfolgt die Regulierung von Forschung teilweise auch im Hinblick auf die möglichen Folgen und Implikationen der Forschung. Dies betrifft beispielsweise die Forschung zur Schaffung neuer Waffen (z. B. Biowaffenkonvention (Gemeinsame Aktion 2006/184/GASP des Rates vom 27. Februar 2006 zur Unterstützung des Übereinkommens über das Verbot von biologischen Waffen und Toxinwaffen (BWÜ) im Rahmen der Strategie der EU gegen die Verbreitung von Massenvernichtungswaffen), die Forschung im Bereich der Präimplantationsdiagnostik (Gesetz zur Regelung der Präimplantationsdiagnostik (Präimplantationsdiagnostikgesetz – PräimpG)) oder die Forschung an bestimmten Forschungsgegenständen, etwa Stammzellen (Gesetz zur Sicherstellung des Embryonenschutzes im Zusammenhang mit Einfuhr und Verwendung menschlicher embryonaler Stammzellen (Stammzellgesetz – StZG)).

Die Notwendigkeit einer Regulierung von Forschung mit Einsatz von Künstlicher Intelligenz wird gegenwärtig überwiegend dahingehend diskutiert, dass der Einsatz solcher algorithmischer Systeme häufig zu Nachweis- und Reproduktionsproblemen führt: Nur ein exakt identischer Datensatz und damit die Kenntnis der verwendeten Daten ließe überhaupt eine genaue Reproduzierbarkeit und damit eine Kontrolle durch andere Forscher/innen erwarten; häufig sind die Ergebnisse aber aufgrund des Machine Learnings nicht nachvollziehbar, nicht transparent und nicht begründbar (Roßnagel 2019a, b, Art. 5 Rn. 148; Gausling 2019, S. 335 f.). Es besteht das Risiko, dass – insbesondere bei empirischer Forschung – die ursprüngliche Studie eventuell bloße Zufallsartefakte gemessen hat und die Ergebnisse daher keine Aussagekraft haben (siehe Huston 2018; Stewart 2018). Daher ist zu beobachten, dass – statt einer Restriktion solcher Forschung – zunächst eine Förderung von Forschung gefordert wird (und auch erfolgt) mit dem Ziel, erklärbar bzw. kontrollierbare Künstliche Intelligenz zu schaffen (High Level Expert Group 2019, S. 3). Dies soll jedenfalls zum Teil mit Mitteln der Künstlichen Intelligenz selbst erfolgen.

Ein weiterer Fokus der Debatte um den Einsatz Künstlicher Intelligenz liegt auf dem Einsatz in der Beurteilung von Menschen und den daraus resultierenden Problemen durch Manipulation, Dark Patterns, fehlende Begründbarkeit der Entscheidung und mögliche Diskriminierung (siehe z. B. Datenethikkommission 2019, S. 97 ff.; Spiecker genannt Döhmman 2020; Ruschemeier 2020). Weiterreichender sind Besorgnisse, dass die Entscheidungen auf der Basis von Künstlicher Intelligenz aufgrund unklarer Entscheidungsregeln und unklarer Dynamik normativ weitere nicht wünschenswerte Auswertungen und Ergebnisse produzieren, z. B. durch die Wiederholung und Verfestigung von Vorurteilen und Diskriminierungen (Sachverständigenkommission für den Dritten Gleichstellungsbericht der Bundesregierung 2021, S. 137; Fröhlich und Spiecker genannt Döhmman 2018), aber auch durch inhärente Vorgaben der Normativität (Sommerer 2020, S. 101 ff.). Gelegentlich rückt auch die Datenqualität und ihre -sicherung in den Mittelpunkt von Überlegungen zur Qualität Künstlicher Intelligenz (Stevens 2020, S. 74 ff.; Hoeren 2016, S. 9 ff.). Diese Aspekte betreffen allerdings zumeist weniger die Forschung mit und an Künstlicher Intelligenz als den Einsatz dieser Verfahren in Realbereichen.

Ein kohärentes Framework zum Umgang mit dem Einsatz von Künstlicher Intelligenz in der Forschung ist in der Konsequenz daher nicht bekannt. Nur hinsichtlich einiger Aspekte, wie eben der Reproduzierbarkeit oder Transparenz, beginnend mit

Diskriminierungen, existiert eine breitere und kenntnisreiche Diskussion, an die angeknüpft werden kann.² In weiten Bereichen adressiert diese allerdings weniger gezielt die Forschung, als dass sie generell Probleme im Umgang mit Künstlicher Intelligenz aufruft. Lösungsansätze sind überwiegend dadurch gekennzeichnet, dass durch Investitionen in Forschung und Entwicklung technische Lösungen für identifizierte spezielle Probleme des Einsatzes von Künstlicher Intelligenz gefunden werden sollen. Überlegungen zur Nutzung von Machine Learning zu Forschungszwecken und den Auswirkungen auf die Forschung selbst werden in diesem Szenario meistens nicht oder lediglich cursorisch thematisiert, selbst wenn Künstliche Intelligenz im Arbeitskontext analysiert wird.³

Insgesamt zeigen sich daher zwar erste Ansätze, warum der Einsatz von Künstlicher Intelligenz reguliert werden sollte und wie das geschehen kann, diese sind allerdings zumeist nicht forschungsspezifisch. In diesem Sinne bildet das Positionspapier des Wissenschaftsrats vom Herbst 2020 „Zum Wandel in den Wissenschaften durch datenintensive Forschung“ eine Ausnahme und einen ersten großen Schritt. Allerdings werden weniger Regulierungsansätze entwickelt, sondern zunächst einmal das Problemfeld sondiert und – häufig gerade auch zugunsten einer rechtssicheren Forschung mit Künstlicher Intelligenz – eher in bestimmten Bereichen das Fehlen von Regulierungsvorgaben und klaren rechtlichen Regeln moniert (WissRat 2020, S. 30 ff.).

6.3 Der Einsatz Künstlicher Intelligenz im allgemeinen Spannungsfeld von Wissenschaftsfreiheit und Datenschutz: Zwei Konstellationen

Künstliche Intelligenz – wie jedes algorithmische System – verarbeitet Daten zur Vorbereitung bzw. Durchführung von Entscheidungen. Damit können Einzelentscheidungen erfasst sein oder auch Teilschritte einer Gesamtentscheidung. Obliegt in letzterem Fall die finale Entscheidung bei einem Menschen oder einer menschlichen Institution, z. B. dem Staat oder einem Unternehmen, ist der Einsatz Künstlicher Intelligenz ein Mittel zur Assistenz menschlichen Entscheidens. In dieser stellen sich ihrerseits wiederum eigene Probleme nach dem Verhältnis von maschinellem und individuellem Entscheiden und der Verantwortung zwischen beiden

²Nicht näher eingegangen wird in diesem Kontext auf die Forschung an ethisch problematischen bzw. riskanten Anwendungsfällen, insbesondere autonome Waffensysteme (Lethal Autonomous Weapon Systems (LAWS)), siehe hierzu beispielsweise die 2012 gegründete Campaign to Stop Killer Robots sowie die Entschließung des Europäischen Parlaments 2018/2752(RSP); rechtswissenschaftliche Behandlung beispielsweise bei Dederer 2018, S. 380 ff.; Rauch 2020, S. 5 ff.

³So finden sich bei der *High Level Expert Group* nur Verweise auf die Relevanz von Forschung, aber keine konkreten Überlegungen zu Differenzierungen. Anders aber *Wissenschaftsrat*, Zum Wandel in den Wissenschaften durch datenintensive Forschung. Positionspapier, 2020, der eine spezifischere Perspektive einnimmt (WissRat 2020, S. 26 ff.). Ähnlich wenig aussagekräftig auch der Dritter Gleichstellungsbericht 2021, S. 99.

Systemen. Zudem wird hier virulent, dass maschinelles Entscheiden oftmals den Eindruck von (höherer) Kompetenz, Objektivität und Unvoreingenommenheit weckt und somit der menschliche Entscheider oftmals in Rechtfertigungslasten gerät, warum er von der – vorbereitenden – maschinellen Entscheidung abweicht. Diese Fragestellungen sind indes nicht spezifisch für den Einsatz von algorithmischen Systemen in der Forschung und für die Forschung.

Es sind zwei Konstellationen vorstellbar, in denen die Forschung an und mit Daten bei Künstlicher Intelligenz einer rechtlichen Regulierung unterworfen wird. Die erste und offenkundige Variante ist die Forschung an Daten durch Wissenschaftler. Der Zweck des Datenumgangs ist dann die Forschung selbst, die Daten sind Mittel zum forschersischen Zweck. Regulierung bemisst sich dann anhand der Forschung und ihrer Ziele; die Forschungsfreiheit widersetzt sich dem Bemühen um Einhegung und Eingrenzung. Der folgende Beitrag befasst sich mit diesem Problemfeld.

Der andere Fall entsteht dann, wenn Künstliche Intelligenz eingesetzt wird, um den Wissenschaftler zum Gegenstand von deren Anwendung zu machen, also z. B. zu Zwecken der Evaluation oder sonstiger Beurteilung und Steuerung von Forschungsinhalten, Publikationserfolgen oder zur Budgetierung. Der Zweck des Datenumgangs ist dann oftmals ein institutioneller Zweck – die Hochschule, welcher der Wissenschaftler/in angehört, möchte seine Tätigkeit vergleichbar machen, eine Governance effektiv gestalten, Schwerpunkte identifizieren und ausweisen, Forschungsleistungen und Rezeption beurteilen, die Nachfrage nach Forschungsinhalten und Forschenden näher bestimmen und vielleicht sogar die Arbeitszeit und -verteilung, etwa auf Lehre und Forschung, kontrollieren.

Auch hier bestimmt sich der anwendbare Rechtsrahmen vor allem aus diesen Zwecken der Forschungsinstitution. Der Rechtsrahmen entstammt daher vor allem dem Hochschul- und Forschungsrecht. Dieses als allgemeines Recht setzt (oft genug allerdings keine) Grenzen, weil es – vieles in die Regelungshoheit der Hochschulen selbst belassend – spezifische Regelungen für diesen Bereich zumeist nicht kennt. Daher muss zur Identifikation von Bindungen der Forschungsinstitutionen auf noch allgemeinere Regelungen zurückgegriffen werden. Diese bilden aber oftmals das hochschul- bzw. forschungsinstitutionell-spezifische Umfeld nicht ab. Da Künstliche Intelligenz auf Daten über Wissenschaftler/innen basiert und die daraus gewonnenen Erkenntnisse auf diese übertragen werden, ist in aller Regel das allgemeine Datenschutzrecht anwendbar.

Damit ergibt sich eine interessante multipolare Gemengelage, die der anderen Situation des Einsatzes von Künstlicher Intelligenz nicht entspricht. Nunmehr streitet die Forschungsfreiheit des Wissenschaftlers oftmals auf Seiten des Datenschutzrechts, während die gleichzeitig bestehende Forschungsfreiheit der Institution entgegengesetzte Interessen unterstützt. Für den Bereich der Publikation kommt dann noch möglicherweise die Pressefreiheit der Publikationsorgane als weiteres zu berücksichtigendes rechtliches Interesse hinzu.

Da in diesem Bereich Künstliche Intelligenz und Big Data bisher wenig eingesetzt werden, nicht zuletzt auch wegen der geringen Standardisierungsmöglichkeiten, soll diesem Themenfeld aber hier keine größere Bedeutung zugewiesen werden. Es

ist allerdings nicht ausgeschlossen, dass Anwendungsfelder für Problembereiche warten, sobald Künstliche Intelligenz auch hier Einzug hält: Man denke nur an die Beurteilung der Reichweite von Veröffentlichungen oder auch schon vorverlagert, die Abfassung und Kontrolle von Publikationen mit Hilfe von Künstlicher Intelligenz.

6.4 Der Einsatz Künstlicher Intelligenz als Mittel der Forschung

6.4.1 Der grundsätzliche Konflikt zwischen Datenschutzrecht und Forschungsfreiheit

Der anwendbare Rechtsrahmen für die Forschung mit Daten im Rahmen des Einsatzes Künstlicher Intelligenz bestimmt sich also zum einen aus der Verwendung und Zielrichtung dieser Daten heraus. Der Zweck bestimmt den Rechtsrahmen, etwa wenn die Daten einem Forschungszweck zugewendet werden sollen, der untersagt oder beschränkt ist. Zum anderen aber gibt es auch rechtliche Regelungen, die bereits an der Verfügbarkeit der Daten selbst ansetzen und daher nicht erst ihre Verwendung als regulatorischen Ansatzpunkt wählen. Sämtliche Rechtsregime, die Daten schützen, verfolgen diesen Ansatz, bei weitem nicht nur das Datenschutzrecht, das sich auf die personenbezogenen Daten bezieht. So werden möglicherweise auch Daten verwendet, die einem Betriebs- und Geschäftsgeheimnis unterliegen oder es werden Daten herangezogen, die aus anderen Gründen nur beschränkt verfügbar sein sollen.

Wissenschaft erfährt beim Einsatz von Künstlicher Intelligenz allerdings nicht nur Beschränkungen. Vielmehr stehen den Forschenden auch Rechte zu, allen voran die Wissenschaftsfreiheit. Darauf kann sich unter bestimmten Umständen auch die Institution, für die er oder sie forscht, berufen. Der Einsatz von Daten, wie er bei Künstlicher Intelligenz unentbehrlich ist, kann also gerade auch von Freiheitsrechten besonders geschützt sein und einen besonderen Freiraum auch rechtlich beanspruchen.

Im Folgenden wird der daraus entstehende Konflikt näher betrachtet, nämlich der – mögliche, nicht zwingende – Gegensatz zwischen Wissenschaftsfreiheit einerseits und Datenschutzrecht andererseits. Dabei werden die Konsequenzen für diejenigen betrachtet, deren Daten für wissenschaftliche Zwecke verwendet werden. Die spätere Übertragung der Erkenntnisse auf einzelne Personen, also die Anwendung der Forschungsergebnisse unter Einsatz von Künstlicher Intelligenz, soll nicht näher untersucht werden; auch dort stellen sich aber ähnliche Probleme und kann sich der Konflikt in ähnlicher Weise auftun.

Dieses Themenfeld, der Konflikt zwischen Datenschutzrecht und Wissenschaftsfreiheit, ist gegenwärtig eines der großen ungelösten Probleme beim Einsatz von Künstlicher Intelligenz und Big Data. Dies liegt nicht nur, aber ganz besonders

daran, dass die Zweckbestimmung wissenschaftlicher Arbeit mit Daten nicht immer klar zu bestimmen ist und dies teilweise auch explizit nicht gewünscht wird.

6.4.2 Wissenschaftsfreiheit

Die Wissenschaftsfreiheit ist in Europa umfänglich bekannt. Sie wird auf verfassungsrechtlicher Ebene geschützt. So formuliert etwa Art. 13 Abs. 1 EU-GR-Charta: „[...] und Forschung sind frei.“ Als Parallelnorm dazu sieht Art. 5 Abs. 3 GG für das nationale Recht vor: „[...] und Wissenschaft, Forschung und Lehre sind frei.“ Eine Besonderheit der Forschungsfreiheit liegt darin, dass sie ohne sog. Schranken gewährleistet ist und insoweit als sog. vorbehaltloses Grundrecht einen herausgehobenen Stellenwert einnimmt (Caspar 2019, Art. 98 Rn. 3).

Grundsätzlich verdrängt das EU-Recht nationales Recht, so dass vorrangig das Verständnis von Forschung und Wissenschaft des EU-Rechts auf die Beurteilung anzuwenden wäre. Allerdings sieht Art. 4 Abs. 3 AEUV Besonderheiten für den Gegenstand der „Forschung“ vor; hier existiert der Sonderfall, dass sich die Zuständigkeit der EU auf Unterstützungs-, Koordinierungs- und Ergänzungsmaßnahmen erstreckt (Calliess 2016, Art. 4 AEUV Rn. 21 unter Verweis auf Nettesheim 2004, S. 530). Daher ist das nationale Verständnis von Forschungsfreiheit in vielen Bereichen gleichwohl von Bedeutung.

Die Forschungsfreiheit erstreckt sich auf alle Elemente des forschersischen Tuns, also auch auf die angewandten Methoden. Dies schließt Künstliche Intelligenz und den Einsatz von Big Data grundsätzlich ein. Wie weit dieser Schutz aber konkret reicht, wird in der Abwägung mit dem Datenschutzrecht noch zu konkretisieren sein (siehe dazu Abschn. 6.4.4).

6.4.3 Datenschutzrecht

Hintergrund und Zielrichtung des Schutzes

Das Datenschutzrecht hält derzeit – neben dem Urheberrecht – wohl das umfassendste Informationsregulierungsrecht vor. Gleichzeitig ist es von der Schutzrichtung her wesentlich komplexer. Es schützt – aus nationalem Verständnis – die „informationelle Selbstbestimmung“ (Hornung und Spiecker genannt Döhmann 2019, Einleitung, Rn. 237 ff.), im internationalen Kontext undeutlicher „Privatheit“. Dahinter verbergen sich allerdings vielschichtige Konzepte, die den spezifischen Gefahren automatisierter Datenverarbeitung begegnen und somit die Verbindung von Daten und darauf basierenden Entscheidungen in den Blick nehmen. Daher beinhaltet Datenschutz auch Diskriminierungsschutz und Schutz vor Intransparenz (Hornung und Spiecker genannt Döhmann 2019, Einleitung, Rn. 10; Kühling und Buchner 2018, Art. 1 Rn. 14).

Während die Wissenschaftsfreiheit ein eher zurückhaltendes Dasein in der rechtlichen Beachtung fristet, hat das Datenschutzrecht in den letzten Jahren erheblich an Gewicht zugelegt. Das hängt zum einen damit zusammen, dass zunehmend erkannt worden ist, dass Digitalisierung durchaus Nebenwirkungen und unerwünschte Konsequenzen zeitigen kann. Zum anderen aber wirkt sich aus, dass gerade auf europäischer Ebene das Datenschutzrecht von einem „zahnlosen Tiger“ zu einem „scharfen Schwert“ gewandelt worden ist. Dies ist vor allem dem EuGH und der Datenschutzgrundverordnung (DSGVO) geschuldet. Beide beziehen sich dabei auf wesentliche Gewährleistungen im Europäischen Primärrecht, insbesondere Art. 7 EU-GR-Charta, wonach „Jede Person [...] ein Recht auf Achtung ihres Privat [...] lebens, ihrer Wohnung sowie ihrer Kommunikation“ hat, sowie Art. 8 Abs. 1 EU-GR-Charta, der „Jede Person hat ein Recht auf Schutz der sie betreffenden personenbezogenen Daten“ regelt.

Hinter dem Datenschutzrecht, das seit dem Aufkommen automatisierter Verarbeitungen in den 1960er-Jahren entwickelt wurde und u. a. in Hessen 1970 seine erste gesetzliche Fassung erhalten hat, stehen diverse Anliegen. Das Rechtsgebiet will verhindern, dass es aufgrund der Digitalisierungstechnik zu informationellen Machtasymmetrien kommt. Diese Gefahr wird u. a. dadurch gesehen, dass bei einer intensiven Datenverarbeitung Entscheidungen des Einzelnen dadurch vorherbestimmt werden können und somit der Einzelne Macht über die Umstände, Voraussetzungen und Folgen seiner persönlichen Entscheidungen verliert. In der Folge kommt es dann zu einem Freiheitsverlust und einer Selbstbeschränkung, weil der Einzelne von seinen Freiheiten keinen Gebrauch mehr macht, da er Sorge um die Konsequenzen aus der Beobachtung hat (Simitis et al. 2019, Einleitung, Rn. 12 ff.). Diese Grundzüge hat für das nationale Recht 1983 das Bundesverfassungsgericht im Volkszählungsurteil formuliert (BVerfGE 65, 1 (43)) und damit eine allgemeingültige Umschreibung gefunden. Auch der EuGH hat sich in verschiedenen Entscheidungen in ähnlicher Weise zu den Grundanliegen des Datenschutzes geäußert (EuGH, Urt. v. 21.12.2016 – C-203/15, C-698/1). Auf eine Kurzformel gebracht, bildet das Datenschutzrecht als ein Persönlichkeitsschutzrecht das Rückgrat von Demokratie, Individualität, Würde und Autonomie und damit den Eckpfeilern der freiheitlich-demokratischen Grundordnung (Spiecker genannt Döhmman 2018).

Die DSGVO hat in der Umsetzung dieser Anliegen eine Vielzahl an spezifischen Anforderungen für Datenverarbeiter aufgebracht; im Kern ist sie eine Fortführung des bisher geltenden Rechts und seines Regulierungsansatzes (Hornung und Spiecker genannt Döhmman 2019, Einleitung, Rn. 212; Paal und Pauly 2021, Einleitung, Rn. 7). Dieser sieht einige Kernanforderungen an Datenumgang vor, zu denen u. a. das Erfordernis einer Rechtsgrundlage für jegliche Datenverarbeitung (Gesetz oder Einwilligung) und ein insgesamt präventives Grundkonzept gehört.

Auf den Einsatz Künstlicher Intelligenz ist dies grundsätzlich unmittelbar anwendbar, weil die EU-Datenschutzgrundverordnung (Verordnung (EU) 2016/679 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung – (DSGVO)) jegliche Verarbeitung von personenbezogenen Daten nach Art. 4 Nr. 1 DSGVO ihrer Anwendung unterwirft. Wenn also personen-

bezogene Daten verarbeitet werden, sind erhebliche rechtliche Vorgaben zu beachten, wenn mit Künstlicher Intelligenz Daten verarbeitet werden. In der Regel knüpfen diese Vorgaben aber nicht an die Verwendung Künstlicher Intelligenz, sondern an die Verwendung von personenbezogenen Daten an (Karg 2019, Art. 4 Nr. 1 Rn. 3, 19; Ernst 2021, Art. 4 Rn. 20).

Personenbezogene Daten

Entsprechend umstritten ist der Begriff trotz der Legaldefinition des Art. 3 Nr. 1 DSGVO. Danach sind personenbezogene Daten „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden „betroffene Person“) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann.“

Von wesentlicher Bedeutung ist also die Personenbeziehbarkeit, also ob die Verbindung zwischen einem scheinbar sachlichen Datum und einer Person hergestellt werden kann. Der EuGH hat den schon unter der Vorgängernorm des Art. 2 lit. a Datenschutz-Richtlinie (DSRL) (Richtlinie 95/46/EG zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr) schwelenden Streit darüber, auf wessen Zusatzwissen es für die Herstellung der Personenbeziehbarkeit bzw. Identifizierbarkeit ankommt (Karg 2019, Art. 4 Nr. 1 Rn. 6), inzwischen entschieden. Danach ist ein objektivierender Begriff zugrunde zu legen, d. h. grundsätzlich kommt es darauf an, dass überhaupt eine Verbindung hergestellt werden kann. Diese Grundüberlegung erfährt allerdings Einschränkungen. So ist u. a. einzubeziehen, wenn ein Dritter eine solche Verbindung herstellen kann (und nicht allein der Datenverarbeiter – dann handelt es sich um den subjektiven Begriff). Dies soll allerdings nur gelten, wenn diese Verbindung legal hergestellt werden kann (EuGH, Urt. v. 19.10.2016 – C-582/14).

Der Leitplankenansatz der DSGVO – insbesondere die Zweckbindung der Erhebung „sensibler“ Forschungsdaten

Ein zentrales Prinzip des Datenschutzrechts mit besonderer Bedeutung für den Einsatz künstlicher Intelligenz in der Forschung ist der Grundsatz der Zweckbindung, Art. 5 Abs. 1 lit. b) DSGVO. Dieser Grundsatz verlangt, dass jegliche Datenverarbeitung personenbezogener Daten für „festgelegte, eindeutige und legitime Zwecke“ erfolgen muss. Dieses Gebot erstreckt sich nicht nur auf die Datenerhebung selbst, sondern auch auf jegliche Weiterverarbeitung. Aufgegriffen wird damit eine Vorgabe aus dem primär-, d. h. verfassungsrechtlichen Gebot des Art. 8 Abs. 2 S. 1 EU-GR-Charta, wonach Daten nur „für festgelegte Zwecke“ verarbeitet werden dürfen (Roßnagel 2019a, b, Art. 5 Rn. 64).

Der Zweckbindungsgrundsatz ist also eine zentrale Einschränkung der Verarbeitung von Daten und beschränkt damit auch den Zugriff und die Weiterver-

arbeitung. Gleichzeitig wird damit auch die Verwendungsdauer reguliert, denn mit Zweckerreichung erlischt auch der Zweck und damit die Befugnis zur Verarbeitung (vgl. Roßnagel 2019a, b, Art. 5 Rn. 95). Das im nicht-europäischen Rechtskreis gelegentlich vertretene Prinzip, wonach man alle Daten nutzen kann, die man einmal auf rechtmäßigem Wege erhalten hat, (vgl. hierzu EuGH, Urt. v. 16.07.2020 – C-311/18) gilt also im europäischen Rechtskreis unter Geltung der DSGVO und der GR-Charta gerade nicht. Denn der Verarbeiter ist gefordert, für die konkrete Verarbeitung eine Zweckbestimmung zu formulieren. Da als Verarbeitung auch die Vorhaltung und Speicherung (Art. 4 Nr. 2 DSGVO) für noch nicht näher konkretisierte Zwecke gilt, ist eine Vorratshaltung von Daten grundsätzlich nicht legitim. Das hat auch der EuGH bereits mehrfach, u. a. für die sog. Vorratsdatenspeicherung, so entschieden (EuGH, Urt. v. 08.04.2014 – C-293/12, C-594/12).

Was allerdings als Beschränkung formuliert ist, ist gleichzeitig auch eine wesentliche Grundlage für die Sicherheit der Verarbeitung für den Datenverarbeiter. Denn der Verarbeiter gewinnt damit auch das Vertrauen desjenigen, dessen Daten verarbeitet werden. Der Zweckbindungsgrundsatz sorgt dafür, dass für den Betroffenen, das sog. Datensubjekt, vorhersehbar ist, was mit den Daten passiert; dies wiederum steigert die Bereitschaft, Daten preiszugeben und bestimmten Verarbeitungen zugänglich zu machen. Diese Funktionalität des Zweckbindungsgrundsatzes wird häufig verkannt.

Dreh- und Angelpunkt legitimer Datenverarbeitung ist damit der jeweilige und konkrete Zweck, der von der DSGVO und dem Datenschutzrecht allgemein als Beschränkung der Datenverarbeitung konstruiert ist. Ein Zweck, der so weit gefasst ist, dass er faktisch keine Begrenzungen vorsieht, kann also dem Zweckbindungsgrundsatz von vorneherein nicht genügen (vgl. Roßnagel 2019a, b, Art. 5 Rn. 70). Je nachdem, wie weit der Zweck gefasst ist, können mehr oder weniger Datenverarbeitungen vereinbar sein. Je präziser ein Zweck gefasst ist, umso mehr Datenverarbeitungen schließt er aus. Gleichzeitig steigt aber die Wahrscheinlichkeit, dass mit einer präzisen Zweckbestimmung die Zulässigkeit der Datenverarbeitung gegeben ist. Denn in diesem Fall ist der Maßstab der Beurteilung eher zu fassen, während Zweifel zu Lasten des Datenverarbeiters gehen. Grundsätzlich gilt, dass der Zweck umso konkreter gefasst sein muss, je schwerer der mit der Datenverarbeitung erfolgende Eingriff in die Rechte des Datensubjekts wiegt (Roßnagel 2019a, b, Art. 5 Rn. 71; Schantz 2020, Art. 5 Rn. 15). Dies entspricht dem Grundsatz der Verhältnismäßigkeit.

Daher ist als erste Schlussfolgerung für die Verwendung von Daten unter Einsatz künstlicher Intelligenz in der Forschung festzuhalten, dass die Verwendung von personenbezogenen Daten zu allgemeinen Zwecken der „Forschung“ den Anforderungen aus Art. 5 Abs. 1 lit. b) DSGVO nicht genügend kann. Dazu bedürfte es einer näheren Spezifikation, um welche Art von Forschung, durch wen, unter Beteiligung welcher weiteren Personen/Institutionen und welchen Weiterverwertungsstrategien es sich handelt. Umgekehrt lässt sich aber auch nicht abstrakt feststellen, wie hoch die jeweiligen Anforderungen an die Zweckbeschreibung sein müssen. Denn Forschungsanliegen unterscheiden sich: Eine Krebsforschung unter Einbeziehung genetischer Faktoren stellt einen schwerwiegenderen Eingriff in die

Rechte des Datensubjekts dar, als eine Forschung, die sich mit der Häufigkeit des Zahnarztbesuchs befasst und unterliegt entsprechend anderen Anforderungen.

Für die Verwendung von Daten im Umfeld von Künstlicher Intelligenz bedeutet der Zweckbindungsgrundsatz jedenfalls dann eine erhebliche Einschränkung für die Forschenden, wenn Daten zweckfrei verwendet und vorgehalten werden sollen. Eine Auswertung ohne Hypothesen oder vorherige Zielrichtung ist mit der DSGVO nicht vereinbar.

Typisch allerdings dürfte es im Bereich der Forschung sein, eine Vielzahl von Daten speichern zu wollen, ohne den späteren Verwendungszweck bereits zum Zeitpunkt und während des Verlaufs der Speicherung festlegen zu können. Dahinter steht der Gedanke, einmal verfügbare und einmal ermittelte Daten nicht wieder preisgeben bzw. löschen zu wollen, u. a. aus Effizienzerwägungen oder auch, weil diese Daten in der Zukunft nicht ohne weiteres replizierbar sein könnten. Ein solches Vorgehen ist mit der DSGVO nicht vereinbar, denn als bestimmter Zweck genügt es nicht, eine mögliche spätere Verwendung in einem noch unbekanntem Kontext zu noch unbestimmten Zwecken anzugeben (Roßnagel 2019a, b, Art. 5 Rn. 72; Schantz 2020, Art. 5 Rn. 15 f.). Dies wäre eine zweckfreie Bevorratung von Daten, die den verfassungsrechtlichen Grundsätzen widerspricht.

Diese Einschränkung betrifft allerdings nur einen engen Teilbereich von Künstlicher Intelligenz, der zudem weniger den eigentlichen Vorgang der Auswertung erfasst als Vorbereitungshandlungen. Zudem wird davon eine wissenschaftlich umstrittene Vorgehensweise mittels Künstlicher Intelligenz, nämlich ohne Hypothese Zufallsergebnisse produzieren zu lassen, erfasst. Jedenfalls die letztere Einsatzmöglichkeit dieser speziellen Technologie kann im Hinblick auf ihre Schutzwürdigkeit durchaus kritisch betrachtet werden, weil von einem planmäßigen Vorgehen zu meist wohl kaum die Rede sein kann.

Es ist aber zu betonen, dass die DSGVO einer Verwendung von Daten im Rahmen von selbstlernenden oder optimierenden Prozessen nicht prinzipiell entgegensteht – im Gegenteil. Das Datenschutzrecht ist – anders als gelegentlich behauptet wird – kein Verhinderungsrecht. Es ist vielmehr ein vielseitiges Regulierungsrecht, das, wie andere Regulierungsrechte auch, verschiedene Interessen einschließlich wirtschaftlicher Interessen in einen Ausgleich zu bringen versucht (Hornung und Spiecker genannt Döhmman 2019, Einleitung, Rn. 249).

Informationstechniken wie Künstliche Intelligenz müssen bei Verwendung personenbezogener Daten die vielfältigen Datenspuren der Nutzer/innen zusammenführen und Auswertungen von Verhalten erstellen können, um den Anforderungen an eine funktionsfähige (und möglicherweise auch kontrollierbare) Künstliche Intelligenz überhaupt genügen zu können.

Es gilt dann aber sehr genaue Strukturierungen und prozessuale Sicherungen in Ergänzung zu weiteren inhaltlichen Unterscheidungskriterien vorzusehen, um sicherzustellen, dass die Vorteile von neuen Technologien wie Künstlicher Intelligenz genutzt werden können, ohne aber die Befürchtungen des Datenschutzes vor Kontrollverlust und Aufgabe der Individualität wahr werden zu lassen (vgl. in diese Richtung auch Roßnagel 2019a, b, Art. 5 Rn. 112 f.). Dazu gehört gerade auch, Daten im Rahmen von Künstlicher Intelligenz nicht für jegliche Zwecke zur Ver-

wendung freizugeben. Vielmehr muss der Einsatz der Technologie als ein Mittel zur Erreichung eines Zwecks, für den weitere Datenerhebungen und -verwendungen erforderlich sind, verstanden und ausgestaltet werden. Dies kann dann dazu führen, dass zur Einhaltung des Zweckbindungsgrundsatzes die Zwecke der Datenverarbeitung als Sub-Zwecke einer funktionierenden Künstlichen Intelligenz in einem bestimmten Einsatzgebiet zu bestimmen sind.

Künstliche Intelligenz als Datenverarbeiter und Verantwortlicher i. S. v. Art. 26 DSGVO?

Die DSGVO ist, wie schon ihre Vorgängerin, die DSRL, davon geprägt, beteiligten Akteuren klare Verantwortlichkeiten zuzuschreiben. Dazu gehört auch, dass derjenige, der die Datenverarbeitung vornimmt, als Verantwortlicher einer Vielzahl von Verpflichtungen ausgesetzt ist. Darunter fallen sowohl primäre Pflichten, etwa die Einhaltung der Vorgaben des Art. 5 Abs. 1 DSGVO oder des Art. 6 Abs. 1 DSGVO zur Legitimität von Datenverarbeitungen. Zu beachten sind aber auch sekundäre Verpflichtungen, wie diejenige zu Schadensersatz nach Art. 82 DSGVO.

Angesichts der Möglichkeiten von Künstlicher Intelligenz lässt sich in Frage stellen, ob sich diese Zuschreibung klarer Verantwortlichkeiten aufrechterhalten lässt: Kann möglicherweise die Eigenständigkeit der Forschung, die mittels Künstlicher Intelligenz vorangetrieben wird, soweit reichen, dass sich die Verantwortlichkeit des Wissenschaftlers als Datenverarbeiter dahinter zurückzieht? Dies würde den Forschenden erheblich entlasten: Nicht er wäre verpflichtet, das Datenschutzrecht zu beachten, sondern die Künstliche Intelligenz, und bei Verstößen haftete nicht er, sondern die Künstliche Intelligenz.

Die Antwort auf diese Frage muss sich zunächst damit auseinandersetzen, ob man einem algorithmischen System eine Rechtspersönlichkeit zubilligen könnte. Denn nur in diesem Falle ließe sich eine Künstliche Intelligenz mit Rechten und – hier von besonderem Interesse – mit Pflichten denken. Obwohl diese Ansicht gelegentlich vertreten wird (z. B. Specht und Herold 2018, S. 43 f.; Schimer 2016, S. 664), ist nicht erkennbar, dass die Gründe, die gegen eine solche Annahme sprechen (siehe z. B. Datenethikkommission 2019, S. 219; Teubner 2018, S. 160 ff.), für den Einsatz Künstlicher Intelligenz in der Forschung anders zu bewerten wären. Zum anderen ist technisch jedenfalls gegenwärtig sehr fraglich, ob man einer Künstlichen Intelligenz so weitreichende Eigenständigkeit zubilligen kann, dass sie tatsächlich zum beherrschenden Akteur wird und ihre Funktion als ein Mittel vollständig aufgibt. Dies kann zumindest derzeit mit guten Gründen bezweifelt werden (siehe Kap. 2 in diesem Band).

Selbst wenn man davon ausgeht, dass algorithmischen Systemen eine Rechtspersönlichkeit zugeschrieben werden könnte, ginge damit noch keine vollständige Entlastung des Wissenschaftlers ein bzw. umgekehrt eine vollständige Verantwortlichkeit der Künstlichen Intelligenz. Denn die DSGVO kennt in Art. 26 DSGVO den sog. „gemeinsamen Verantwortlichen“ als Adressaten ihrer Vorgaben. Diese Vorschrift reagiert auf die zunehmend vernetzte Verarbeitung von Daten, die häufig arbeitsteilig erfolgt (Petri 2019, Art. 26 Rn. 1). Sie weist aber nicht etwa unter mehreren Datenverarbeitern einem die Alleinverantwortlichkeit zu und entlastet damit

alle anderen. Vielmehr obliegt es den Verantwortlichen, die konkrete Zuteilung zu regeln; nach außen bleiben alle Verantwortlicher (Petri 2019, Art. 26 Rn. 4). Damit wird vermieden, dass die arbeitsteiligen und vernetzten Strukturen zum Nachteil der Betroffenen wirken können, weil diese womöglich nicht sicher wissen, an wen sie ihre Ansprüche richten müssen (Petri 2019, Art. 26 Rn. 2). Aus der Rechtsprechung des EuGH (EuGH, Urt. v. 05.06.2018 – C-210/16) wird deutlich erkennbar, dass für die Annahme einer gemeinsamen Verantwortlichkeit keine erhebliche Einwirkungsmöglichkeit aller Beteiligten gegeben sein muss, sondern die Nutzung eines Angebots (etwa eines Plattformbetreibers) und damit die Nutzung einer Datenverarbeitung zu eigenen Zwecken ausreichen kann. Insofern ist schwer vorstellbar, dass die Wissenschaftler/innen, die eine Künstliche Intelligenz auf den Weg gebracht haben, nicht weiterhin wenigstens gemeinsam Verantwortlicher blieben. Zudem bliebe auch im Innenverhältnis fraglich, wer dann die Hauptlast trüge: Denn da der Forschende die Künstliche Intelligenz eingesetzt hat, bleibt eine erhebliche Einwirkungsmöglichkeit auf seiner Seite.

6.4.4 *Das Austarieren von Datenschutz und Wissenschaftsfreiheit nach dem Verständnis der DSGVO sowie des deutschen Rechts*

Forschungsfreiheit und Datenschutzrecht sind verfassungsrechtlich gewährleistete Rechte. Weder kann sich also das eine noch das andere vollständig durchsetzen; vielmehr müssen beide im Konfliktfall in einen Ausgleich gebracht werden. Dies ist zuvörderst Aufgabe des Gesetzgebers. In der DSGVO sind für diesen Fall Regelungen vorgesehen. Diese sind allerdings nicht ganz einfach zu verstehen und umzusetzen, wie im Folgenden gezeigt wird.

Einführung

Ist eine präzise und rechtlich tragfähige Bestimmung des Zwecks der Forschung mit Künstlicher Intelligenz schon nicht immer ganz einfach, erschwert die DSGVO mit einigen unklaren Formulierungen die Datenverarbeitung im Bereich der Künstlichen Intelligenz zusätzlich. Das liegt daran, dass die DSGVO selbst dem Anliegen der Wissenschaftsfreiheit aktiv Rechnung trägt, also bereits in ihren Vorschriften einen Ausgleich der grundsätzlich ähnlich wertigen Rechte des Datenschutzes und der Forschungsfreiheit herstellt. Dies geschieht allerdings unter Berücksichtigung der besonderen Kompetenzen der EU. Daher ist an dieser Stelle eine komplizierte rechtliche Gemengelage entstanden. In dieser greifen EU-Regelungen und nationale Regelungen ineinander, zum Teil können sogar entgegen der herkömmlichen Lehre die nationalen Regelungen den europäischen Vorgaben vorgehen.

Hintergrund dieses schwierigen rechtlichen Bestands sind mehrere Sonderregelungen innerhalb der DSGVO zum Thema „Forschung“ (vgl. zu einer Übersicht der verstreuten Vorschriften Caspar 2019, Art. 89 Rn. 31 ff.), allen voran

Art. 89 DSGVO und Art. 5 Abs. 1 lit. b) 2. HS DSGVO. Diese Vorschriften privilegieren auf unterschiedliche Weise die „Forschung“ gegenüber anderen Verarbeitungszwecken und Verarbeitungsschritten. U.a. kann nach Art. 5 Abs. 1 lit. b) 2. HS DSGVO eine Weiterverwendung von Daten, die zu einem anderen als einem Forschungszweck erhoben worden sind, mit derselben Rechtsgrundlage für Forschungszwecke erfolgen. Bei oberflächlicher Betrachtung könnte man damit die eben beschriebene Zweckbeschränkung jedenfalls für Forschungszwecke als aufgehoben betrachten. Und Art. 89 Abs. 2 DSGVO überlässt den Mitgliedstaaten in einer weiten Öffnungsklausel die Möglichkeit, für die Zwecke der Forschung Rechte der Betroffenen weitgehend zurückzuschneiden, solange Ausgleichsmaßnahmen anderer Art dafür vorgesehen sind.

Der Begriff der Forschung in der DSGVO

Angesichts solcher erheblichen Privilegierungen ist die Frage aufgerufen, was unter dem Begriff der „Forschung“ i. S. d. DSGVO überhaupt zu verstehen ist, wer und was also überhaupt von dieser erheblichen Privilegierung profitieren kann. Fest steht zunächst nur, dass die Berufung auf „Forschung“ nicht dazu führen kann, dass jegliche Bindung an das Datenschutzrecht aufgelöst wäre. Denn das führte dazu, dass die verfassungsrechtlich abgesicherte Forschungsfreiheit das ebenso verfassungsrechtlich abgesicherte Datenschutzrecht vollständig verdrängte und sich diesem gegenüber umfassend durchsetzte. Angesichts der weitreichenden Gleichrangigkeit beider Verfassungsrechte ist dies aber nicht zulässig.

Eine Legaldefinition des Forschungsbegriffs in der DSGVO fehlt, wie man ja auch von einem gesicherten Verständnis auf der verfassungs-/primärrechtlichen Ebene nicht ausgehen kann (Roßnagel 2019a, b, S. 158). Art. 13 EU-GR-Charta kennt den Forschungsbegriff und versteht ihn eng, nämlich unter Ausschluss der Lehre (Jarass 2021, Art. 13 Rn. 9, Bernsdorff 2019, Art. 13 Rn. 14). Die eigentliche Definition kann durchaus vergleichbar einer vom deutschen Bundesverfassungsgericht entwickelten Formel formuliert werden (Roßnagel 2019a, b, S. 158), zumal das deutsche Verfassungsrecht als Inspirationsquelle für die europäische Fassung diene (Ruffert 2016, Art. 13 GRCh Rn. 1). Danach ist Forschung alles, was „nach Inhalt und Form als ernsthafter planmäßiger Versuch zur Ermittlung von Wahrheit anzusehen ist“ (Zit. als: BVerfGE 47, 327 (367)). Geschützt ist also jeder als Person und alles als Gegenstand, was zum Zweck des Erkenntnisgewinns nach wissenschaftlichen Methoden unternommen wird (Ruffert 2016, Art. 13 GRCh Rn. 8). Die Anwendung ist davon nicht erfasst, d. h. Regulierung kann ohne Berücksichtigung der Forschungsfreiheit – dann aber möglicherweise unter Berücksichtigung anderer Grundrechte wie Berufs- und Eigentumsfreiheit – erwogen werden.

Dem Erwägungsgrund 159 der DSGVO lässt sich zudem für das Verständnis der DSGVO entnehmen, dass nicht nur die Grundlagen, sondern auch die anwendungsorientierte und auch die privatfinanzierte Forschung erfasst sein sollen. Ferner führt Erwägungsgrund 159 aus, dass der Begriff für die DSGVO „weit ausgelegt“ werden soll.

Diese Formulierungen der DSGVO verbergen, dass mit Art. 89 DSGVO und den weiteren Vorschriften zur Forschung ein verfassungsrechtlich verlangter Ausgleich

zweier Grundrechte vorgenommen wird. Denn auch wenn der Forschungsfreiheit des Art. 13 EU-GR-Charta und ebenso des Art. 5 Abs. 3 GG keine unmittelbare Beschränkung zu entnehmen ist, gilt trotzdem, dass im nationalen wie im europäischen Recht die Rechte anderer und die anerkannten Gemeinwohlbelange von Verfassungsrang bzw. aus Art. 52 Abs. 1 EU-GR-Charta entgegenstehen können (Ruffert 2016, Art. 13 GRCh Rn. 11). Die DSGVO mit ihren Privilegierungen für Forschungsanliegen stellt also einen Ausgleich zweier potenziell konfligierender Grundrechte dar und sorgt dafür, dass beide in einem angemessenen Verhältnis zueinander zur Geltung kommen können. Weder kann die Forschungsfreiheit noch das Datenschutzrecht das jeweils andere Grundrecht vollständig verdrängen.

Daher bedeutet die in Erwägungsgrund 159 DSGVO angesprochene „weite“ Lesart des Forschungsbegriffs keine vollständige Übernahme des verfassungsrechtlichen Begriffs und sie führt auch nicht dazu, dass angesichts des Vorliegens jeglicher Art von Forschungsanliegen der Datenschutz zurücktreten müsste. Eine weite und quasi unbeschränkte Lesart von Forschung im Rahmen der DSGVO würde dazu führen, dass nahezu alles, das irgendeinem Erkenntnisgewinn dient und egal von wem die Forschung und zu welchen Zwecken sie betrieben wird, in den Genuss der weitreichenden Privilegierungen der DSGVO käme. Damit würde die grundsätzliche Beschränkung, welche die DSGVO mit Prinzipien wie Zweckbindung und Datenminimierung in Art. 5 Abs. 1 DSGVO vorsieht, unterlaufen. Es träte ein Verdrängen der Anliegen der Datenschutzgrundrechte durch die Forschungsfreiheit ein. Sie entspräche auch nicht der Überschrift von Kap. 9 der DSGVO, wonach es um „besondere Verarbeitungssituationen“ gehen soll, also um Ausnahme- und nicht um Regeltatbestände (ebenso Caspar 2019, Art. 89 Rn. 10). Daher ist insgesamt aus dem Kontext der DSGVO, den allgemeinen Zielen, Zwecken und der Systematik eine restringierende Lesart zu entwickeln, welche dem Kernanliegen einen klugen Ausgleich zwischen Datenschutzrecht und Forschungsfreiheit herzustellen, gerecht wird, ohne aber damit gleichzeitig das Datenschutzrecht auszuhöhlen und leer laufen zu lassen. Das lässt sich auch aus Art. 89 Abs. 1 DSGVO auslesen, der nämlich bei Reduktion der Anforderungen der DSGVO aus Gründen der Forschungsfreiheit begleitende Garantien für die Rechte und Freiheiten fordert, also deutlich macht, dass ein Weniger an Datenschutz durch Forschung insgesamt nicht hingekommen werden soll.

Auch aus einer Metaperspektive auf die Bedeutung des Datenschutzrechts kann ein weiter Forschungsbegriff nicht zu einer vollständigen Verdrängung des Datenschutzrechts führen. Das Datenschutzrecht ist das ureigene, zentrale Recht der Informationsgesellschaft und der Regulierung der Informationstechnologie. Es befasst sich mit automatisierter Datenverarbeitung wie kein anderes Rechtsgebiet; es ist seit seinen Ursprüngen auf die Bewältigung der automatisierten Datenverarbeitung angelegt (Simitis et al. 2019, Einleitung, Rn. 1). Sein Regulierungsanliegen ist es also gerade, die technische Innovation und die damit verbundene massenhafte Auswertung von Daten zu beschränken, allerdings nicht grundsätzlich zu verbieten. Deshalb darf auch der Forschungsbegriff nicht zur Beliebigkeit verkümmern (Caspar 2019, Art. 89 Rn. 15).

Daher ist die Umschreibung der Privilegierung für Forschung, z. B. in Art. 5 Abs. 1 lit. b) DSGVO, durch den europäischen Gesetzgeber von „wissenschaftlicher Forschung“ (in Abgrenzung zu „historischer Forschung“) nur scheinbar tautologisch (Caspar 2019, Art. 89 Rn. 15). Daraus lässt sich nämlich folgern, dass gerade die Entwicklung eines Forschungsgegenstands und der hierzu verwendeten Methodik ihrerseits wiederum wissenschaftlichen Anforderungen und Standards genügen muss. Daher sind z. B. Korrelationen und Zufälligkeiten, die methodisch herangezogen werden, nicht ausreichend wissenschaftlich im Vergleich zu echter statistischer Auswertung und Belastbarkeit sowie Nachvollziehbarkeit und Kontrollierbarkeit der Verfahren und darauf basierenden Ergebnisse (Spiecker genannt Döhmann 2017, S. 60). Eine Künstliche Intelligenz, die sich diese Ansätze ohne wissenschaftliche Basierung und ohne wissenschaftliches Konzept zunutze machen wollte, hätte Schwierigkeiten vom Forschungsprivileg der DSGVO zu profitieren, das aus der Forschungsfreiheit erwächst.

Andererseits muss einer zu engen Lesart die Stirn geboten werden, die vorschnell unter Heranziehung des unklaren Wortlauts nur eine öffentlich finanzierte Forschung genügen lassen würde. Zum einen spricht Erwägungsgrund 159 ausdrücklich davon, dass auch privatfinanzierte Forschung privilegiert sein kann, wenngleich dies nicht als Zwangsläufigkeit („muss“) formuliert ist. Daher kann es sehr wohl Fälle geben, in denen eine privat finanzierte Forschung vom Forschungsprivileg ausgeschlossen wird – wie es eben auch Fälle geben kann, in denen eine öffentlich finanzierte Forschung ausgeschlossen wird, wie soeben erläutert. Vor allem aber entspricht die deutsche Übersetzung nicht dem englischen Text: Während die deutsche Fassung „im öffentlichen Interesse“ voranstellt, so dass man es auf alle drei genannten privilegierten Tatbestände (also Archivzwecke, wissenschaftliche und historische Forschungszwecke) beziehen kann, stellt die englische Fassung „in the public interest“ nach und zwar allein und ausschließlich den Archivzwecken.

Gleichwohl ist nicht jede privatfinanzierte Forschung eine privilegierte Forschung, wie sich aus dem Einschränkungsanliegen des 9. Kapitels der DSGVO folgern lässt. Forschungsfreiheit ist nicht um ihrer selbst geschützt – dafür hätte die Berufsfreiheit genügt. Sondern Forschungsfreiheit ist deshalb geschützt, weil die Wissensvermehrung grundsätzlich auch der Allgemeinheit zur Verfügung steht und ihr nützt. Es bedarf daher zur Anerkennung im Rahmen der DSGVO einer sozialen, allgemeinwohlorientierten Zwecksetzung, die nicht allein wirtschaftlichen Interessen dient. Rein kommerzielle Forschungen, die nicht auch der Allgemeinheit zu dienen bestimmt sind, sondern gerade darauf abzielen, durch Geheimhaltung und Verbergen der Erkenntnisse Vorsprung vor Konkurrent/innen zu gewinnen oder Defizite bei Kunden und Nutzer/innen auszunutzen, fallen nicht darunter (Caspar 2019, Art. 89 Rn. 16 ff.). Sie können die vielseitigen Allgemeinwohlbelange, die das Datenschutzrecht verfolgt, nicht aushebeln.

Das ergibt sich auch daraus, dass eine Abgrenzung der Berufsfreiheit und der Forschungsfreiheit zu treffen ist. Im Umkehrschluss kann man sagen, dass jedenfalls sog. drittmittelfinanzierte Forschung an öffentlichen Forschungseinrichtungen unter den privilegiierenden Forschungsbegriff fallen dürfte (so wohl auch Caspar 2019, Art. 89 Rn. 17). So darf zwar das Thema durch Dritte vorgegeben werden,

nicht aber die Methodik. Und der Ablauf der Forschung darf nicht von Unternehmensentscheidungen oder den Entscheidungen Dritter abhängen, wie es aber für eine Auftragsforschung typisch ist. Die Grenzlinie ist also durchaus schwierig zu ziehen; sie ist graduell je nach Grad der Unabhängigkeit zu bestimmen.

Ferner ist zu belegen, dass die Forschung dem Allgemeinwohl zu dienen bestimmt ist, weil sie auf Veröffentlichung und damit auf Teilung der Ergebnisse und ihre Diskussion angelegt ist. Eine Forschung, die ihre Ergebnisse abgrenzen und allein für sich bzw. die Unternehmenszwecke nutzen will, genügt diesen Anforderungen nicht. Sie ist schon nicht wissenschaftlich, weil sie den Aspekt der Selbstkontrolle des Wissenschaftsapparats durch die fachliche Diskussion und Prüfung missachtet und damit einen wesentlichen Definitionsaspekt nicht entfaltet.

Zudem lässt sich auch eine anwendungsorientierte Forschung durchaus unter Art. 89 DSGVO und die weiteren Vorschriften der DSGVO zur Forschung subsumieren. Die Anwendung selbst ist allerdings nicht mehr erfasst, denn diese unterfällt nicht der Forschung selbst. Ansonsten unterfiele jede Verbesserung eines strukturell entwickelten algorithmenbasierten Informationssystems dem Forschungsprivileg der DSGVO (vgl. auch Caspar 2019, Art. 89 Rn. 17).

Um Missverständnisse vorzubeugen, sei darauf verwiesen, dass die Allgemeinwohlorientierung eine notwendige, aber nicht etwa hinreichende Bedingung ist: Denn der Forschungsbegriff umfasst weitere Kriterien, die zudem eingehalten sein müssen. So kann sich auch allgemeinwohlorientierte Forschung Methoden bedienen, die nicht mehr als wissenschaftlich einzustufen sind und damit aus dem Forschungsprivileg der DSGVO herausfallen.

Für eine Künstliche-Intelligenz-gestützte oder sogar zentral auf Künstliche Intelligenz ausgerichtete Forschung bedeutet dies Konsequenzen: Zunächst einmal ist zu identifizieren, wer die Forschung mittels Künstlicher Intelligenz betreibt und wie dies geschieht. Wenn es sich um eine Forschung handelt, die auf Öffentlichkeit ausgerichtet ist, also darauf, das Allgemeinwohl durch einen Erkenntnisgewinn zu steigern, der geteilt wird, greifen die Privilegierungen der DSGVO. Sie erreichen allerdings nur den Bereich, der tatsächlich der Entwicklung, also der eigentlichen Beforschung, dient, nicht aber der Anwendung und deren Verbesserung ohne fundamentale Neuentwicklung. Zudem ist darauf zu achten, dass der Einsatz Künstlicher Intelligenz nach allgemein anerkannten wissenschaftlichen Prinzipien erfolgt.

6.4.5 Die Lockerung der Zweckbindung in Art. 5 Abs. 1 lit. b 2. HS DSGVO

Unterfällt die Entwicklung und Verwendung einer Künstlichen Intelligenz dem Forschungsbegriff, ist eine entscheidende Privilegierung in Art. 5 Abs. 1 lit. b) 2. HS DSGVO enthalten: Danach können nämlich Daten, die der Entwickler der Künstlichen Intelligenz zur Verfügung hat, für die Zwecke dieser Forschung verwendet werden, auch wenn diese Daten ursprünglich zu einem ganz anderen Zweck erhoben und verarbeitet, z. B. gespeichert, worden sind. Das Grundproblem der

Zweckbindung ist damit für diesen besonderen Bereich weitgehend aufgehoben. Nur in Einzelfällen, bei besonders kritischen Daten oder besonders problematischen Anwendungen der Künstlichen Intelligenz, würde die Privilegierung des Art. 5 Abs. 1 lit. b) 2. HS DSGVO nicht greifen.

Einzige Einschränkung der erweiterten Nutzbarkeit von Daten über die Entbindung vom Zweckbindungsgrundsatz ist Art. 6 Abs. 4 DSGVO, wonach der Datenverarbeiter zu prüfen hat, ob der Zweck tatsächlich noch ein vereinbarter Zweck ist. Wegen der Sonderformel des Art. 5 Abs. 1 lit. b) 2. HS DSGVO ist diese Vorschrift allerdings in den meisten Fällen der Forschung gegeben. Art. 5 Abs. 1 lit. b) 2. HS DSGVO stellt dafür eine Vermutung auf, die allerdings widerlegt werden kann. Denn auch bei der Anwendung von Art. 6 Abs. 4 DSGVO im Zusammenspiel mit Art. 5 Abs. 1 lit. b) 2. HS DSGVO ist zu beachten, dass es sich um eine Regelung handelt, die einen Ausgleich zwischen Forschungsfreiheit und Datenschutzgrundrechten herstellt, also weder das eine noch das andere Recht unangemessen im Ergebnis privilegieren darf. Art. 7 Abs. 2 S. 2 EU-GR-Charta verlangt ausdrücklich die Einhaltung einer Zweckbindung und auch wenn diese zugunsten der Forschung zurückgenommen werden kann, darf sie doch nicht obsolet werden. Daher kann Art. 5 Abs. 1 lit. b) 2. HS DSGVO nicht dahingehend verstanden werden, dass jegliche Forschung jegliche Zweckbindung aufgeben kann, sondern dass die Beschränkungen des Art. 6 Abs. 4 DSGVO weiterhin gelten und jeweils zu prüfen sind (so auch Roßnagel 2019a, b, S. 162). Immer dann, wenn besonders eingriffsintensive Maßnahmen vorgenommen werden, darf der Verarbeiter nicht ohne weiteres von der Vereinbarkeit ausgehen, auch wenn der Wortlaut des Art. 5 Abs. 1 lit. b) 2. HS DSGVO dies vermuten lässt.

Für den Einsatz Künstlicher Intelligenz bedeutet dies auch in der Forschung eine hohe Aufmerksamkeit für die Einhaltung des Datenschutzrechts. Denn sie ist grundsätzlich mit ihrem großen und spezifischen Datenbedarf und den intensiven Verknüpfungsmöglichkeiten (auf diese Kriterien stellt auch Roßnagel 2019a, b, S. 162 *ab*) als besonders eingriffsintensiv einzuschätzen. Auch wenn sie als Forschung privilegiert ist, unterliegt sie dennoch besonderen Anforderungen. Das formuliert Art. 89 Abs. 1 DSGVO mit der Forderung nach zusätzlichen Garantien in besonderer Weise und dürfte angesichts der Schwierigkeiten an Nachvollziehbarkeit, Begründbarkeit und Wiederholbarkeit schwer umzusetzen sein. Art. 6 Abs. 4 DSGVO fordert hier mehr als nur eine standardisierte Betrachtung. Daher ist beispielsweise bzgl. der Herkunft der verwendeten Daten genau zu prüfen, ob es eine enge Verbindung mit der Forschung gibt. Zudem sind Beschränkungen der Weiterverwendung für andere Zwecke hilfreich, um eine Legalität der Verwendung Künstlicher Intelligenz zu sichern.

Dass im Übrigen die sonstigen Voraussetzungen der DSGVO eingehalten werden müssen, so z. B. eine Datenschutz-Folgeabschätzung nach Art. 35 DSGVO, soll hier nur als Hinweis aufgenommen werden. Dies bedeutet, dass der Einsatz von Künstlicher Intelligenz als Forschungsinstrument nicht leichtfertig eingesetzt werden sollte. Es erfordert einen aktiven planerischen Zugang auf die datenschutzrechtlichen Gewährleistungen einschließlich umfangreicher Sicherungsmaßnahmen, um Missbrauch und Manipulation, Diskriminierung und andere unerwünschte Effekte zu verhindern.

6.5 Datenschutzrecht und sonstige Anforderungen bei Kooperationen

Nicht unterschlagen werden soll eine weitere offene Flanke von und für Forschung mit Künstlicher Intelligenz, nämlich die unsichere Rechtslage für Kooperationen zwischen Forschungseinrichtungen.

Der erhebliche Ressourcenbedarf der Entwicklung und Begleitung Künstlicher Intelligenz setzt große Anreize, aber auch große Notwendigkeiten für eine Zusammenarbeit zwischen verschiedenen Wissenschaftlern und Forschungseinrichtungen. Dies ist bereits jetzt zu beobachten, wird sich aber noch verstärken, möglicherweise auch von Forschungsförderinstitutionen vorangetrieben (man denke nur an die DFG-Finanzierung des Höchstleistungsrechnens, dazu [WissRat 2012](#)). Ein wichtiger, auch wettbewerbsrelevanter Faktor ist dabei die Verfügbarkeit von Datensätzen: Ein Partner mit großen Sammlungen von Daten, auf die für den Einsatz der Künstlichen Intelligenz zurückgegriffen werden kann, kann für die Qualität der Ergebnisse von entscheidender Bedeutung sein.

Ein Kooperationsrecht für öffentliche Einrichtungen, gar für Hochschulen und Forschungseinrichtungen, gibt es allerdings nicht. Anders als im privatrechtlichen Bereich, wo sich seit Jahrhunderten mit dem Gesellschaftsrecht und dem Eherecht ein umfassendes Kooperationsrecht herausgebildet hat, einschließlich eines teilweise bestehenden Formzangs und der Möglichkeit der Gründung juristischer Personen für eigenständige Rechtspersönlichkeiten auf Dauer, kennt das öffentliche Recht ein solches Kooperationsrecht nur in Ansätzen. In der Folge gibt es kein Rechtsregime, auf das sich Kooperationspartner als Grundsatzregelung stützen können. Vielmehr bedarf es zu einer rechtssicheren Gestaltung umfassender, individuell auszuhandelnder Kooperationsverträge, die zum Teil (z. B. bei länderüberschreitender Kooperation und je nach Wesentlichkeit des Kooperationsgegenstands) sogar parlamentarisch beschlossen werden müssen. Das Spektrum der zu regelnden Bereiche ist groß. Es reicht von urheberrechtlichen Fragen bis hin zu Kontrollrechten und Budgethoheit, erfasst aber auch grundlegende Entscheidungen wie Abstimmungsrechte, Kündigungsregelungen und Weiterverwertung nach Beendigung oder Teilausstieg.⁴

Diese Probleme fehlender rechtlicher Regelungen steigern sich noch, wenn im Rahmen internationaler Forschung nicht nur europäische Partner zusammenkommen, sondern auch Partner anderer Staaten kooperieren, deren Rechtsvorstellungen erheblich divergieren können. Deutlich wird das gerade in den Informations- und Datenrechtsvorstellungen, ganz besonders im Urheber- und im Datenschutzrecht. Denn nur in wenigen Staaten weltweit kann von einem Datenschutzniveau ausgegangen werden, das demjenigen der DSGVO entspricht und das auch als solches anerkannt worden ist (zuletzt Japan Anfang 2019: Durchführungsbeschluss (EU) 2019/419 der Kommission vom 23. Januar 2019 über die Angemessenheit des Datenschutzniveaus in Japan; siehe die EuGH-Entscheidung

⁴Siehe zum Themenkomplex insgesamt Richter und Spiecker genannt Döhmman [2019](#), S. 181 ff.

Schrems II zum unzureichenden Datenschutz auf der Basis des Privacy Shield, EuGH, Urt. v. 16.07.2020 – C-311/18), so dass in diesen Kooperationen von einem rechtlich unproblematischen Transfer der Daten ausgegangen werden kann. Und auch die Zugriffsrechte auf urheberrechtlich geschützte Werke variieren erheblich, wie u. a. die Auseinandersetzung um Google Books – möglicherweise sogar auch für Forschungszwecke – zeigt (siehe dazu Gesmann-Nuissl und Wünsche 2012, S. 228 ff.).

Eine weitere ungelöste Schwierigkeit ergibt sich zudem durch eine jüngere Rechtsprechung des EuGH, in der dieser eine gemeinsame Verantwortlichkeit nach Art. 26 DSGVO (siehe dazu im Verhältnis zur Künstlichen Intelligenz unter Abschn. 6.4.3) und damit auch grundsätzlich eine gemeinsame Haftung für alle Verstöße gegen die DSGVO annahm, obwohl einer der Partner faktisch keinerlei Einflussmöglichkeit auf die – beanstandete – Datenverarbeitung des anderen Partners hatte. Der EuGH ließ es genügen, dass der geringere Partner von der Datenverarbeitung des bestimmenden Partners profitierte (EuGH, Urt. v. 05.06.2018 – C-210/16). Diese Situation dürfte sich auch bei Kooperationen regelmäßig stellen: Nicht alle Partner sind in gleicher Weise an allen Forschungsanliegen beteiligt, profitieren aber insgesamt – und sei es durch Förderzuwendungen oder Veröffentlichungen – von der Datenverarbeitung der anderen Partner. Verwendet also ein Beteiligter Daten unter Verstößen gegen die DSGVO, sind davon auch die sonstigen Datenverarbeitungen infiziert und jeder Partner muss fürchten, für die Verstöße anderer haftbar gemacht zu werden. Da Künstliche Intelligenz davon profitiert, dass Daten aus verschiedenen Quellen stammen, muss auf dieser Basis mangels rechtlicher Regelung wenigstens vertraglich sichergestellt sein, dass strikte Datenschutz- und IT-Sicherheitsvorkehrungen getroffen werden. Sind in anderen Staaten allerdings weitreichende Zugriffsrechte des Staates ohne besondere Zweckbindungen zulässig, ist dies schwierig (vgl. EuGH, Urt. v. 16.07.2020 – C-311/18).

Schließlich ist damit umzugehen, in welcher Weise sich Rechte und Pflichten nicht nur, aber insbesondere aus der DSGVO verändern, wenn ein Wissenschaftler seine institutionelle Anbindung wechselt, erst recht, wenn dies in einer Kooperation der Fall ist. Unklar ist, ob und wie der Datenzugriff dann weiterhin möglich ist, wer also Zugriffsrechte behält (siehe auch WissRat 2020, S. 33) und wie eine Kontrolle über die Daten gewährleistet werden kann, durch wen also beispielsweise Pflichten auf Löschung oder Berichtigung nachzukommen ist. Es entstehen darüber möglicherweise Datenpools bei Institutionen, die losgelöst von den individuellen Forschern existieren und damit – entgegen des Zweckbindungsgrundsatzes – zweckfrei für allgemeine Forschung vorgehalten werden.

6.6 Fazit und Ausblick

Wie bereits kenntlich gemacht, sind die Regulierungsanstrengungen im Bereich des Einsatzes Künstlicher Intelligenz bisher begrenzt; im Bereich der Forschung mit und an Künstlicher Intelligenz gilt dies erst recht. Am ehesten lässt sich noch der

Konflikt zwischen Wissenschaftsfreiheit und Datenschutzrecht als vom geltenden Recht beeinflusst herausarbeiten. Er prägt die Nutzung künstlicher Intelligenz im Bereich der Forschung erheblich mehr als vielen Anwendern dieser Technologie bisher bewusst ist, denn Künstliche Intelligenz basiert auf der weitreichenden und oftmals zweckentfremdenden Nutzung von Datenmengen und dazu gehören gerade auch personenbezogene Daten. Angesichts fast gänzlich fehlender Regelungen wäre es wünschenswert, wenn es Klarstellungen für Forschende gäbe. Dies kann in den hochschulrechtlichen Bestimmungen geschehen, dies kann und muss auch im Beamten- und Arbeitsrecht festgelegt werden. Denn Forschung findet eben nicht allein im Schutzzumfeld öffentlich geförderter Institutionen statt, sondern ganz wesentlich auch im privaten Bereich und unter Kooperationen beider.

Das bedeutet nicht, dass nicht eine Reihe von weiteren Rechtsfragen aufgeworfen werden können und aufgeworfen werden. Dies betrifft beispielsweise urheberrechtliche Fragestellungen: Wer ist Urheber eines Werks, das mit Hilfe von Künstlicher Intelligenz gefertigt worden ist – greift der Werkbegriff hier überhaupt? Am eindringlichsten mag dieses Problem sich illustrieren lassen, wenn Künstliche Intelligenz eigene Texte verfasst.

Eng damit verknüpft sind Fragen der Ergebniskontrolle und der Sicherung der Qualität von Forschung; so sind z. B. Publikationsorgane dazu übergegangen, von den Autoren/innen zu verlangen, ihre Datensätze zur Verfügung zu stellen, um eine Nachkontrolle – auch z. B. im Peer Review Verfahren durch unbekannte Dritte – zu ermöglichen. Dies schafft allerdings auch Möglichkeiten der Weiterverwertung dieser Daten für eine Vielzahl weiterer Personen und Institutionen über den unmittelbaren Forschungsbereich hinaus.

Will man solche weitreichenden Einflussnahmen auf Forschung durch Künstliche Intelligenz wegen der mangelnden Determination des Ergebnisses nicht vollständig ausschließen, kommt auch die Frage nach der Rechtsstellung schnell auf. Wird in absehbarer Zeit der/die Wissenschaftler/in womöglich der/diejenige sein, der/die Forschungsergebnisse der Künstlichen Intelligenz nur noch aufbereitet und präsentiert und die eigenständige Forschungsleistung maschinell erbracht?

Das ruft Fragen auf, die im Kontext Künstlicher Intelligenz ohnehin diskutiert werden, nämlich nach Nachvollziehbarkeit, Transparenz und Begründung/Rationalisierung des Einsatzes dieser Technologie. Sie sind im Forschungsbereich etwas anders gestellt und sie werfen zum Teil auch eigenständige Folgefragen auf – etwa, wenn zur Qualitätssicherung von Forschungsergebnissen und zur unabhängigen Kontrolle die vollständigen Datensätze und ein Duplikat des algorithmischen Systems an ein Publikationsorgan geliefert werden sollen und darüber – z. B. in Peer Review Verfahren – unmittelbare Konkurrenten Zugriff auf diese Daten und Forschungsansätze erhalten.

Auch der Transfer von Forschung mit und an Künstlicher Intelligenz in der Anwendung bleibt problematisch: Schon jetzt wird diskutiert, wie sichergestellt werden kann, dass und wie eine algorithmisch unterstützte Entscheidungsfindung weiterhin verantwortlich ausgestaltet werden kann. Dies trifft auch die Forschung selbst und es trifft sie zusätzlich, wenn sie den Transfer von der Wissenschaft in die Innovation und Marktfähigkeit bewältigen muss.

Letztlich treten damit die großen Fragen des Einsatzes von Künstlicher Intelligenz bereits in der Entwicklung im Umfeld der Forschenden auf: Wie kann die Verantwortung für die Durchführung und die Ergebnisse einer Forschung mit und von Künstlicher Intelligenz sichergestellt werden, wie kann eine Qualitätskontrolle gesichert werden und wie können Fehler und Ungenauigkeiten nachgewiesen und in eine Haftung für die Auswirkungen und den Einsatz überführt werden?

Die Forschungsfreiheit verlangt für den der Anwendung und Innovation vorgelagerten Bereich noch einmal besondere Freiheiten, gleichzeitig aber auch besondere Verantwortlichkeiten. Insoweit stellen sich für den Einsatz Künstlicher Intelligenz die grundlegenden Fragen der Regulierung von Forschung einmal mehr und einmal anders, aber nicht gänzlich neu. Es bleibt dabei, dass die Auswertung von Daten, gerade in großem Umfang, vielfältige Ergebnisse generieren kann. Wofür diese eingesetzt, weiterverfolgt, weiterentwickelt und letztlich verwendet werden, kann durch die Forschung oftmals selbst gar nicht bestimmt werden. Diese Weiterverwendungs- und Zugangsmöglichkeiten sind aber mitzudenken und sie verlangen eine nicht nur beiläufige beständige datenschutzrechtliche Begleitung bis hin zur Weiterentwicklung auf technischer Ebene. Künstliche Intelligenz verändert also nicht das Verständnis von Forschung, aber es führt dazu, dass – jedenfalls aus datenschutzrechtlicher Sicht – einige Fragen mit neuer Stoßrichtung und anderer Präzision gestellt werden und in der Konsequenz Künstliche Intelligenz auch anders ausgerichtet werden muss

Literatur

- Bernsdorff N (2019) Art. 13. In: Meyer J, Hölscheidt S (Hrsg) Charta der Grundrechte der Europäischen Union, 5. Aufl. Nomos, Baden-Baden
- Calliess C (2016) Art. 4 AEUV. In: Ruffert M (Hrsg) EUV, AEUV: das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta. Kommentar, ders, 5. Aufl. C.H. Beck, München
- Caspar J (2019) Art. 98. In: Simitis S, Hornung G, Indra Spiecker genannt Döhmann (Hrsg) Kommentar Datenschutzrecht. DSGVO mit BDSG, 1. Aufl. Nomos, Baden-Baden
- Datenethikkommission (2019) Gutachten der Datenethikkommission. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf;jsessionid=F8BF886A15BAEEF0C4117E68CCEACABD.2_cid373?__blob=publicationFile&v=6. Zugegriffen am 18.02.2021
- Dederer H-G (2018) Letale Autonome Waffensysteme (LAWS). Z Rechtswiss Forsch 9(4):380–413
- Ernst S (2021) Art. 4. In: Paal BP, Pauly DA (Hrsg) Kommentar Datenschutz-Grundverordnung. Bundesdatenschutzgesetz, 3. Aufl. C.H. Beck, München
- Fröhlich W, Spiecker genannt Döhmann I (2018) Können Algorithmen diskriminieren? <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/>. Zugegriffen am 18.02.2021
- Gausling T (2019) Künstliche Intelligenz im digitalen Marketing. Datenschutzrechtliche Bewertung KI-gestützter Kommunikations-Tools und Profiling-Maßnahmen. Z Datenschutz 9(8):335–341
- Gesmann-Nuissl D, Wünsche K (2012) Neue Ansätze zur Bekämpfung der Internetpiraterie – ein Blick über die Grenzen. Gewerblich Rechtsschutz Urheberrecht Int 61(3):225–234
- Hamer M (2020) Informatisierung des Verwaltungsprozesses: Berücksichtigung der Möglichkeiten moderner Informations- und Kommunikationstechniken bei der Gestaltung und Aus-

- legung der Normen des Verwaltungsprozessrechts. In: Schriften zum Prozessrecht 167. Duncker und Humblot, Berlin
- High Level Expert Group (2019) Ethik-Leitlinien für eine vertrauenswürdige KI. <https://www.demographie-netzwerk.de/site/assets/files/5064/ethicsguidelinesfortrustworthyai-depdf.pdf>. Zugegriffen am 17.02.2021
- Hoeren T (2016) Thesen zum Verhältnis von Big Data und Datenqualität. Erstes Raster zum Erstellen juristischer Standards. Z IT-Recht Recht Digital 19(1):8–11
- Hornung G, Spiecker genannt Döhmman I (2019) Einleitung, Rn. 155–273. In Kommentar Datenschutzrecht. DSGVO mit BDSG, hrsg. Spiros Simitis und dies, 1. Aufl. Nomos, Baden-Baden
- Huston M (2018) AI researchers allege that machine learning is alchemy. <https://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy>. Zugegriffen am 16.02.2021
- Jarass H (2021) Art. 13. In Charta der Europäischen Union unter Einbeziehung der sonstigen Grundrechtsregelungen des Primärrechts und der EMRK, hrsg. ders, 4. Aufl. C.H. Beck, München
- Karg M (2019) Art. 4. In: Simitis S, Hornung G, Spiecker Genannt Döhmman I (Hrsg) Kommentar Datenschutzrecht. DSGVO mit BDSG, 1. Aufl. Nomos, Baden-Baden
- Kühling J, Buchner B (2018) Art. 1. In Datenschutz-Grundverordnung/BDSG: Kommentar, hrsg. dies, 2. Aufl. C.H. Beck, München
- Nettesheim M (2004) Die Kompetenzordnung im Vertrag über eine Verfassung für Europa. Europarecht 39(4):511–546
- Paal BP, Pauly DA (2021) Einleitung. In Kommentar Datenschutz-Grundverordnung. Bundesdatenschutzgesetz, hrsg. dies, 3. Aufl. C.H. Beck, München
- Petri T (2019) Art. 26. In: Simitis S, Hornung G, Spiecker genannt Döhmman I (Hrsg) Kommentar Datenschutzrecht. DSGVO mit BDSG, 1. Aufl. Nomos, Baden-Baden
- Rauch A (2020) Autonome Waffensysteme und Völkerrecht: Stand der wissenschaftlichen Diskussion, offene Forschungsfragen und aktuelle Regelungsanstrengungen im internationalen System. In C. Tietje (Hrsg) Beiträge zum Europa- und Völkerrecht 20. https://tietje.jura.uni-halle.de/sites/default/files/BeitraegeEVR/Heft%2020_Rauch.pdf. Zugegriffen am 17.02.2021
- Richter E, Spiecker genannt Döhmman I (2019) Rechtliche Gestaltung von Verwaltungskooperationen. In: Durner W, Reimer F, Spiecker genannt Döhmman I, Wallrabenstein A (Hrsg) *Das sinnvoll Denkbare denken, das davon Machbare machen, Gedächtnisschrift für Arndt Schmehl*. Duncker und Humblot, Berlin, S 179–198
- Roßnagel A (2019a) Datenschutz in der Forschung. Die neuen Datenschutzregelungen in der Forschungspraxis von Hochschulen. Z Datenschutz 9(4):157–164
- Roßnagel A (2019b) Art. 5. In: Simitis S, Hornung G, Spiecker genannt Döhmman I (Hrsg) Kommentar Datenschutzrecht. DSGVO mit BDSG, 1. Aufl. Nomos, Baden-Baden
- Ruffert M (2016) Art. 13 GRCh. In EUV, AEUV: das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta. Kommentar, hrsg. Christian Calliess und ders, 5. Aufl. C.H. Beck, München
- Ruscheimer H (2020) 9. Speyerer Forum zur digitalen Lebenswelt: Regulierung Künstlicher Intelligenz in der Europäischen Union zwischen Recht und Ethik. NVwZ 39(7):446–448
- Sachverständigenkommission für den Dritten Gleichstellungsbericht der Bundesregierung (2021) Digitalisierung geschlechtergerecht gestalten. Gutachten für den Dritten Gleichstellungsbericht der Bundesregierung. Berlin: Geschäftsstelle Dritter Gleichstellungsbericht. www.drittergleichstellungsbericht.de/gutachten3gleichstellungsbericht.pdf. Zugegriffen am 18.02.2021
- Schantz P (2020) Art. 5. In: Brink S, Wolff H A (Hrsg) BeckOK Datenschutzrecht, 34. Aufl. C.H. Beck, München
- Schirmer J-E (2016) Rechtsfähige Roboter? JuristenZeitung 13:660–666. <https://doi.org/10.1628/002268816X14615987983565>
- Simitis S, Hornung G, Spiecker genannt Döhmman I (2019) Einleitung, Rn. 1-154. In Kommentar Datenschutzrecht. DSGVO mit BDSG, hrsg. dies, 1. Aufl. Nomos, Baden-Baden
- Sommerer LM (2020) Personenbezogenes Predictive Policing Kriminalwissenschaftliche Untersuchung über die Automatisierung der Kriminalprognose. In: Schriften zur Kriminologie 19, 1. Aufl. Nomos, Baden-Baden

- Specht L, Herold S (2018) Roboter als Vertragspartner? Gedanken zu Vertragsabschlüssen unter Einbeziehung automatisiert und autonom agierender Systeme. *Z IT-Recht Recht Digital* 21(1):40–44
- Spiecker genannt Döhmman I (2010) Wissensverarbeitung im Öffentlichen Recht. *Z Rechtswiss Forsch* 3:247–282
- Spiecker genannt Döhmman I (2017) Big und Smart Data: Zweckbindung zwecklos? *Spektrum Wiss* 1:56–62
- Spiecker genannt Döhmman I (2018) Fragmentierungen. Kontexte der Demokratie – Parteien, Medien, Sozialstrukturen. *VVDStRL* 77:9–65
- Spiecker genannt Döhmman I (2020) „Ich habe nichts zu verbergen“ – wirklich nicht?: Unterscheidung versus Diskriminierung durch KI, Algorithmen und die digitalen Dienste. *Forsch Frankf* 37(1):38–41
- Spiecker genannt Döhmman I (2021) Staatliche Entscheidungen unter Unsicherheit. Mohr Siebeck (i.E.), Tübingen
- Stevens J (2020) Datenqualität bei algorithmischen Entscheidungen. Überlegungen aus Anlass des Gutachtens der Datenethikkommission. *Comput Recht* 36:73–79
- Stewart M (2018) The machine learning crisis in scientific research. <https://towardsdatascience.com/the-machine-learning-crisis-in-scientific-research-91e61691ae76>. Zugegriffen am 16.02.2021
- Teubner G (2018) Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten. *Archiv civilis Prax* 218:155–205
- Wissenschaftsrat (2012) Strategische Weiterentwicklung des Hoch- und Höchstleistungsrechnens in Deutschland. Positionspapier. https://www.wissenschaftsrat.de/download/archiv/1838-12.pdf?__blob=publicationFile&v=1. Zugegriffen am 17.02.2021
- Wissenschaftsrat (2020) Zum Wandel in den Wissenschaften durch datenintensive Forschung. Positionspapier. https://www.wissenschaftsrat.de/download/2020/8667-20.pdf?__blob=publicationFile&v=5. Zugegriffen am 18.02.2021

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

