2.2.2 Menschenrechte

Gemeinwohlorientierte Gesetzgebung auf Basis der Vorschläge der EU »High-Level-Expert Group on Artificial Intelligence«

Eric Hilgendorf

I Ethik und Rechtspolitik im Zeitalter von Digitalisierung und künstlicher Intelligenz

Die Digitalisierung und ihr derzeit meist diskutiertes Anwendungsfeld, die künstliche Intelligenz (KI), sind dabei, unsere gesamte Lebens- und Arbeitswelt umzugestalten. Die Corona-Pandemie hat der Digitalisierung einen zusätzlichen, so noch nie dagewesenen Schub verliehen. Gleichzeitig werden die Notwendigkeit und die Schwierigkeiten einer Regulierung der Digitalisierung in Wirtschaft, Staat und Gesellschaft in teilweise drastischer Weise beleuchtet und hervorgehoben.[1]

1 Die Literatur zur normativen Bewältigung von KI ist inzwischen nicht mehr überschaubar, vgl. nur Anderson, Michael/Anderson, Susan Leigh: Machine Ethics, Cambridge: Cambridge University Press 2011; Beck; Susanne/Kusche, Carsten/Valerius, Brian: Digitalisierung, Automatisierung, KI und Recht, Baden-Baden: Nomos 2020; Bendel, Oliver: Handbuch Maschinenethik, Wiesbaden: Springer 2019; Coeckelbergh, Mark: AI Ethics, Cambridge: The MIT Press 2020; Dignum, Virginia: Responsible Artificial Intelligence, Schweiz: Springer 2019; Hengstschläger, Markus: Digital Transformation and Ethics, Elsbethen: Ecowin 2020; Misselhorn, Catrin: Grundfragen der Maschinenethik, Ditzingen: Reclam 2018; Nida-Rümelin, Julian/Weidenfeld, Nathalie: Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz, München: Piper 2018. Zu Regulierungsfragen Schallbruch, Martin: Schwacher Staat im Netz. Wie die Digitalisierung den Staat in Frage stellt, Wiesbaden: Springer 2018 und zuletzt Nemitz, Paul/Pfeffer, Matthias: Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, Bonn: J.H.W. Dietz Nachf. 2020.

Den neuen Technologien wird eine in höchstem Maße innovative, ja geradezu »disruptive« Macht zugesprochen.[2] Es liegt auf der Hand, dass ein derart radikaler Veränderungsprozess zahlreiche normative Probleme, und damit ethische und rechtspolitische Herausforderungen aufwerfen muss. Die westlichen Gesellschaften (aber nicht nur sie) haben sich nach dem letzten Weltkrieg mit in Rechtsform gebrachten Menschenrechten einen normativen Rahmen gegeben, der die Staatsmacht bindet und sie verpflichtet, bei Verletzungen der Menschenrechte einzuschreiten. Die Basis der Menschenrechte bildet die Menschenwürde, die sich als ein Ensemble aus grundlegenden subjektiven Rechten des Individuums verstehen lässt.[3] Auf diese Weise wird das Individuum besonders wirksam geschützt.

Der Rückbezug auf die Menschenrechte ermöglicht es, Rechtspolitik und Ethik miteinander zu verknüpfen. Beide werden oft als gegensätzliche Tätigkeitsfelder beschrieben: Die Ethik, so meinen manche, habe es mit übergeordneten Werten und Maßstäben zu tun, die Rechtspolitik dagegen werde bestimmt durch kurzfristige Ziele und bloße »instrumentelle Vernunft«. Bei näherem Hinsehen zeigt sich indes, dass normative Ethik und Rechtspolitik schon deswegen eng aufeinander bezogen sind, weil es in beiden Bereichen um gut begründete oder erst zu begründende Normen geht, an denen sich die menschliche Praxis orientieren kann.

Dementsprechend muss gefragt werden: Welche normativen Vorgaben sollten die Rechtspolitik auf dem Gebiet der Digitalisierung leiten? Zunächst gilt, dass die ethische Analyse für eine rationale Rechtspolitik unverzichtbar ist. Es überrascht deswegen nicht, dass auf vielen politischen Handlungsfeldern ethische Expertise herangezogen wird, und zwar nicht als bloße Bemäntelung anderweitig gefundener Entscheidungen, Ethics-Washing[4] oder als Inspirationsquelle für Sonntagsreden, sondern als wichtiges und in vielen Bereichen sogar unverzichtbares Analyse- und Reflexionsangebot. Dies

2 Zu den Anwendungsfeldern der neuen Technologien eingehend Grunwald, Armin: Der unterlegene Mensch. Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern, Teil II, München: Riva 2019.

3 Dazu näher unter 2.1. am Anfang.

4 Darunter versteht man die Formulierung hochtrabender ethischer Prinzipen, um öffentlicher Kontrolle und einer wirksamen (weil verpflichtenden) rechtlichen Regulierung zu entgehen. Möglicherweise liegt hier eine Ursache für die auffällige Inflation ethischer Regeln für KI, dazu auch Jobin, Anna/Lenca, Marcello/Vayena, Effy: The global landscape of AI ethics guidelines, Berlin: Nature Machine Intelligence 1, 2019, S. 389-399.

betrifft neben Europa auch andere Länder und Großregionen, die sich anschicken, den technischen Fortschritt auf dem Gebiet von Digitalisierung und KI im Einklang mit ihren jeweiligen kulturellen und gesellschaftspolitischen Vorstellungen zu regulieren, insbesondere die USA und China.[5]

Allerdings scheint Europa auf dem Gebiet der Regulierung anderen Ländern und Großregionen einen Schritt voraus zu sein. In den *Ethics Guidelines for Trustworthy AI* der EU, die am 8. April 2019 veröffentlicht wurden,[6] hat die EU eine Basis für weitere Regulierungsmaßnahmen geschaffen. Die Ethik-Leitlinien stellen den Menschen und seine Bedürfnisse in den Mittelpunkt; immer wieder wird darin von einem »human-centric approach« gesprochen. Dieser Ansatz lässt sich am ehesten als »humanorientiert« oder »humanistisch« einstufen.[7] Die Vorteile der neuen Technologien sollen nicht in erster Linie einzelnen Großunternehmen zugutekommen oder die Macht des Staates mehren, sondern dem Gemeinwohl dienen, also dem Wohlergehen aller Menschen.[8]

---

5 Dementsprechend haben praktisch alle technisch fortgeschrittenen Länder auch entsprechende Ethik-Entwürfe für die Regulierung von KI vorgestellt, siehe den Überblick bei M. Coeckelbergh: AI Ethics, S. 150ff.; P. Nemitz/M. Pfeffer: Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, S. 314.

6 https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Insgesamt publizierte die HLEG AI vier Dokumente: (1) Die »Ethics Guidelines for Trustworthy AI«, welche die grundlegenden ethischen und rechtspolitischen Erwägungen enthalten, (2) die »Policy and Investment Recommendations«, welche sich mit Fragen der Umsetzung der Grundlagenerwägungen in Politik und Wirtschaft beschäftigen, (3) die »Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment«, durch die den beteiligten Unternehmen die Möglichkeit gegeben wurde, den eigenen Stand der Umsetzung zu überprüfen, und schließlich (4) die »Sectoral Considerations on the Policy and Investment Recommendations for Trustworthy Artificial Intelligence«, in denen vier ausgewählte Anwendungsgebiete, nämlich die industrielle Produktion, E-Government, die Rechtspflege und der Gesundheitsbereich, näher analysiert und potenzielle Anwendungsmöglichkeiten von künstlicher Intelligenz herausgearbeitet wurden.

7 Ausführlich Nida-Rümelin, Julian/Weidenfeld, Nathalie: Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz; Hilgendorf, Eric: Humanismus und Recht – Humanistisches Recht? Eine erste Orientierung, in: Groschopp, Horst (Hg.), Humanismus und Humanisierung, Aschaffenburg: Alibri 2014, S. 36-56.

8 Coeckelbergh, Mark: AI Ethics, S. 183f. hat zu Recht darauf hingewiesen, dass eine solche Humanorientierung schon deshalb nicht selbstverständlich ist, weil andere empfindungsfähige Lebewesen, also heute Tiere und in Zukunft vielleicht einmal Maschinen, ausgeschlossen bleiben.

Einige Beispiele für rechtspolitische Herausforderungen durch KI-Anwendungen

Warum ist die ethische und rechtspolitische Auseinandersetzung mit KI-Anwendungen überhaupt so wichtig? Die folgenden Beispiele machen deutlich, vor welchen Problemen wir stehen.

Drängend, aber ethisch und rechtspolitisch schwierig zu beantworten, sind zunächst Fragen im Zusammenhang mit der modernen digitalisierten Kommunikation, etwa in den sozialen Netzwerken, in denen offenbar zunehmend »autonome« Algorithmen Botschaften formulieren, über sogenannte Social Bots verbreiten und so das wahrgenommene Meinungsspektrum beeinflussen. Eine der größten ethischen wie rechtspolitischen Herausforderungen stellt daher die Frage nach der Reichweite von Meinungsfreiheit im digitalisierten Raum dar.[9] Es ist kaum zu übersehen, dass die Kommunikation im Internet mehr und mehr verroht und Hassrede zu einem Alltagsphänomen geworden ist. KI-gestützte Filtertechnologien können hier helfen, werfen aber das Problem auf, wer über die »herauszufilternden« Inhalte entscheiden soll – die Gesellschaft, Tech-Konzerne oder gar die KI selbst? Haftungsrisiken können für Plattform-Unternehmen ein wirkungsvolles Motiv sein, sich um die über sie verbreiteten Inhalte zu kümmern. Die Haftungsprivilegien der großen Internet-Provider erscheinen deshalb zunehmend als problematisch. Des Weiteren stellen sich interessante interkulturelle Fragen, die sich aus der globalen Reichweite moderner Kommunikation ergeben, etwa wenn Äußerungen, die nach westlichen Standards unbedenklich sind, in einer anderen Großregion (etwa in der arabischen Welt) als grob beleidigend oder gotteslästerlich angesehen werden.

Mit Blick auf die durch autonome Systeme gesteuerte automatisierte Produktion – die Industrie 4.0 – stellt sich die Frage nach dem normativen Rahmen neuer Formen von Arbeit, etwa wenn Menschen direkt mit maschinellen »Kollegen« zusammenarbeiten. Für die industrielle Produktion, die auf ein Zusammenspiel von Mensch und Maschine setzt, rücken Sorgfaltsstandards in den Mittelpunkt: Wie sicher muss die verwendete Technologie sein? Welche Schutzmaßnahmen hat der Arbeitgeber vorzuhalten? Hinzu treten Haftungsfragen: Wer trägt bei einem Unfall die Verantwortung und muss Schadensersatz leisten? Noch komplizierter werden die Haftungsfragen im Zusammen-

9 Umfassend Garton Ash, Timothy: Redefreiheit. Prinzipien für eine vernetzte Welt, München: Carl Hanser Verlag GmbH 2016 (deutlich angelsächsisch geprägte Sicht).

hang mit der Nutzung von Augmented Reality, etwa wenn Menschen über eine VR-Brille missverständliche Arbeitsanweisungen erhalten (und in der Folge ein Schaden entsteht) oder wenn der Kontakt mit dem Gegenüber nicht mehr von Mensch zu Mensch geschieht, sondern mittels Avataren in einer virtuellen Umgebung.[10]

Ein weiteres, ebenfalls bereits intensiv diskutiertes Problemfeld bilden neue Formen digital gestützter Mobilität, so etwa Dilemma-Probleme, die sich ergeben, wenn ein Fahrzeug bzw. dessen autonom agierender Kollisionsvermeide-Assistent zwischen der Verletzung oder gar Tötung von Menschen zu entscheiden hat.[11] Dürfen wir Entscheidungen über Leben und Tod an Maschinen übertragen? Welche Regeln sollen gelten, wenn eine solche Übertragung vorgenommen wurde? Welchen ethischen und rechtlichen Kriterien soll die Maschine folgen? Ebenfalls bislang nicht hinreichend thematisiert ist die Frage, inwieweit der Staat mithilfe technischer Mittel in den Straßenverkehr eingreifen darf, um Verhalten zu unterbinden, durch das sich die Fahrzeugführer selbst oder andere gefährden. Man könnte hier von den Herausforderungen, aber auch Chancen eines »technologischen Paternalismus« sprechen.[12]

Auch die digitalisierte Medizin, die mehr und mehr an Aufmerksamkeit gewinnt, wirft erhebliche ethische wie rechtspolitische Probleme auf, etwa solche der Verteilungsgerechtigkeit (sollten leistungsfähige und dementsprechend teure medizinische Technologien auch den Armen zur Verfügung gestellt werden?) oder ob Maschinen als »Gefährten« von hochbetagten oder geistig beeinträchtigten Menschen eingesetzt werden dürfen. Es droht die Abhängigkeit von außereuropäischen Monopolanbietern, wenn Europa unter Berufung auf den Datenschutz darauf verzichtet, konkurrenzfähige E-

10 In Würzburg kam es schon vor einigen Jahren zu einem Zwischenfall in einem VR-Testlabor, als der (weibliche) Avatar einer Studentin vom (männlichen) Avatar einer anderen Person massiv sexuell bedrängt wurde.

11 Zusammenfassend Hilgendorf, Eric: Dilemma-Probleme beim automatisierten Fahren. Ein Beitrag zum Problem des Verrechnungsverbots im Zeitalter der Digitalisierung, Berlin: Zeitschrift für die gesamte Strafrechtswissenschaft (ZStW) Bd. 130, 2018, S. 674-703. Es ist sehr bemerkenswert, dass das Dilemma-Problem sogar im neuen Gesetzentwurf zu einer Reform des Straßenverkehrsgesetzes (StVG) behandelt wird, vgl. Bundesrats-Drucks. 155/21 vom 12.2.2021 – »Entwurf eines Gesetzes zur Änderung des Straßenverkehrsgesetzes und des Pflichtversicherungsgesetzes – Gesetz zum autonomen Fahren«, S. 27.

12 Dazu näher unter III.6 sowie den Beitrag von Timo Radermacher und Erik Schilling in diesem Band.

Health-Angebote zu entwickeln. Darüber hinaus stellen sich Fragen, die gemeinhin mit dem Schlagwort »Enhancement« umschrieben werden. Dabei geht es letztlich um die »Verbesserung« von Menschen mit technischen Mitteln. Eine extreme, zugleich aber durchaus herausfordernde Position nimmt hier der Transhumanismus ein, der eine Fortentwicklung des Menschen mithilfe der Technik offen begrüßt.[13]

Alle genannten Problemfelder werden verknüpft durch die Herausforderungen, die die KI selbst aufwirft. Handelt es sich nur um ein Werkzeug von Menschen, sodass die Regulierung und Festlegung von Verantwortlichkeiten zum Beispiel bei Personen ansetzen sollte, die KI herstellen, vermitteln oder verwenden? Oder ist es zweckmäßiger, KI als eigenständigen Akteur und eigenständiges Verantwortungssubjekt zu sehen? Letzteres würde die überkommenen Mechanismen der Verantwortungszuschreibung und der Trennung zwischen Rechtssubjekten (die Rechte besitzen können) und bloßen Objekten (die als solche nicht »rechtsfähig« sind) auf eine harte Probe stellen.

II Die Ethischen Leitlinien für eine vertrauenswürdige KI und ihre Rezeption

2.1 Die Ethics Guidelines der EU High-Level Expert Group

Im Frühjahr 2018 hatte die EU-Kommission angekündigt, zur Förderung der KI-Forschung in Europa tätig werden zu wollen.[14] Zu diesem Zweck wurde im Herbst 2018 eine Kommission aus 52 Fachleuten zusammengestellt, die je zu einem Drittel aus der Industrie, der akademischen Welt und aus NGOs kamen: die EU High-Level Expert Group on Artificial Intelligence (HLEG AI). Die Gruppe erhielt den Auftrag, tragfähige und praktisch umsetzbare Regeln für die neue Welt der KI in Europa zu entwerfen. Bereits am 8. April 2019 veröffentlichte die HLEG AI ihre Vorschläge für eine vertrauenswürdige (*trust-*

13 Coeckelbergh, Mark: AI Ethics, S. 38ff.; ausführlich Hilgendorf, Eric: Menschenwürde und die Idee des Posthumanen, Menschenwürde und Medizin: Ein interdisziplinäres Handbuch, Berlin: Duncker & Humblot 2013, S. 1047-1067.

14 Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe, COM (2018) 237 final (vom 25.4.2018).

worthy) KI.[15] Danach gehören zu einer vertrauenswürdigen KI drei zentrale Elemente: Die KI muss (1) rechtmäßig sein, also den jeweiligen rechtlichen Vorgaben entsprechen, sie muss (2) ethisch akzeptabel sein und (3) (technisch wie sozial) robust. Zu Letzterem gehört auch und gerade der Gesichtspunkt der IT-Sicherheit gegenüber Angriffen, ein fundamental wichtiges Erfordernis, das in sämtlichen Anwendungszusammenhängen von KI zu beachten ist.

Von Anfang an wurde in der HLEG AI die praktische Umsetzbarkeit der zu entwickelnden Vorschläge mit bedacht.[16] Deshalb wurden die Industrie und Verbraucherschutzverbände gleich zu Beginn in die Arbeit eingebunden. Zudem wurde eine Testphase eingeführt, in der die Vorschläge in ausgewählten Unternehmen[17] praktisch auf ihre Umsetzbarkeit geprüft wurden. Dadurch unterscheiden sich die Empfehlungen der HLEG AI von vielen eher akademisch orientierten und nur in zweiter Linie auf konkrete Wirkung hin angelegten Regelwerken.

Menschzentrierter Ansatz

Eine leitende Idee der hier vorgestellten Regeln ist, dass die Entwicklung von KI am konkreten Menschen mit seinen faktisch vorfindbaren Bedürfnissen orientiert sein soll. Dies ist mit dem Konzept des *human-centric approach* gemeint. Als Leitwert verweist die HLEG AI ausdrücklich auf die Menschenwürde. Sie lässt sich als ein Ensemble von sieben grundlegenden subjektiven Rechten verstehen: (1) einem Recht auf ein materielles Existenzminimum, (2) dem Recht auf autonome Selbstentfaltung (minimale Freiheitsrechte), (3) dem Recht auf Freiheit von extremen Schmerzen (z. B. gegen Folter), (4) dem Recht auf Wahrung einer minimalen Privatsphäre, (5) dem Recht auf geistig-see-

15 Siehe oben Fn. 6. Zum Terminus »trustworthy« ebd. M. Coeckelbergh: AI Ethics, S, 152f.; zum Konzept einer »trustless technology« Nemitz, Paul /. Pfeffer, Matthias: Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, S. 168. Die HLEG AI beschäftigte sich mit sogenannter schwacher, also bereichsspezifischer KI, nicht mit »starker«, menschenähnlicher KI. Diese Grundentscheidung führte mehrfach zu scharfen Auseinandersetzungen in der Gruppe. Letztlich wurden die interessanten, aber eher theoretischen Fragen starker KI ausgespart, um sich auf die praktisch derzeit wesentlich relevanteren Fragen »schwacher« KI konzentrieren zu können.

16 Zur Bedeutung dieses Ansatzes ebd. Coeckelbergh, Mark: AI Ethics, S. 168ff.

17 Dazu gehört etwa die Firma BOSCH. Der Leiter der KI-Forschung von BOSCH, Christoph Peylo, wirkte in der HLEG AI als Experte mit.

lische Integrität, (6) dem Recht auf grundsätzliche Rechtsgleichheit und (7) dem Recht auf minimale Achtung.[18]

Es handelt sich nach dieser Konzeption um echte subjektive Rechte von Individuen auf Schutz ihrer basalen Interessen, nicht bloß um objektives Recht ohne unmittelbaren Individualbezug. Nur subjektive Rechte erlauben es der berechtigten Person, ihr Recht einzuklagen. Durch die Konzeption der Menschenwürde als Bündel von grundlegenden subjektiven Rechten wird also die Stellung der beziehungsweise des Einzelnen mithilfe des wirkungsvollsten verfügbaren Rechtsinstruments gestärkt, der Einräumung eines einklagbaren, relativ präzise umrissenen subjektiven Rechts. Die Menschenwürde schützt nach dieser Konzeption aber nur einen Kernbereich menschlicher Interessen; die oben genannten Rechte sind eng zu interpretieren. So ist nur der innerste Bereich der Privatsphäre (»Intimsphäre«) durch die Menschenwürde absolut geschützt; hier sind keinerlei Abwägungen zulässig. Dagegen ist die weitere Privatsphäre nicht primär durch die Menschenwürde geschützt, sondern durch das abgeleitete und fortentwickelte Recht auf informationelle Selbstbestimmung.[19] Auch bei den anderen Ausprägungen der Menschenwürde kann zwischen einem unantastbaren Kernbereich und einem Außenbereich unterschieden werden, in dem die entsprechenden Interessen zwar grundrechtlich geschützt sind (etwa durch die allgemeine Handlungsfreiheit oder den Gleichheitsgrundsatz), aber eben nicht mehr absolut.

Vier grundlegende ethische Prinzipien

Ausgehend von der Menschenwürdegarantie werden in den Ethik-Leitlinien vier grundlegende ethische Prinzipien identifiziert und daraus sieben Anforderungen abgeleitet, die vertrauenswürdige KI-Systeme erfüllen müssen. Zusätzlich hat die HLEG AI eine Reihe von Fragen und Kriterien formuliert, die dabei helfen sollen, die Leitlinien zu testen und ihre Anforderungen umzusetzen. Die in den Leitlinien festgehaltenen vier ethischen Prinzipien werden explizit auf die europäischen Grundrechtevorgaben gestützt, insbesondere auf

18 Hilgendorf, Eric: Problem Areas in the Dignity Debate and the Ensemble Theory of Human Dignity, in: Grimm, Dieter, Kemmerer, Alexandra und Möllers, Christoph (Hg.), Human Dignity in Context. Explorations of a Contested Concept, Baden-Baden: Nomos 2018, S. 325ff.

19 Art. 2 Abs. 1 i.V.m. Art. 1 Abs. 1 Grundgesetz.

die Charta der Grundrechte der EU, und so im geltenden Recht verankert.[20] Es handelt sich um die vier folgenden Grundsätze:

(1) Respekt vor der menschlichen Autonomie: KI-Systeme müssen so entwickelt werden, dass sie der Freiheit und der Autonomie von Individuen hinreichend Rechnung tragen,
(2) Schadensvermeidung: KI-Systeme dürfen Menschen nicht schädigen,
(3) Fairness: Zu diesem relativ unbestimmten Begriff soll unter anderem gehören, dass Anstrengungen unternommen werden, um individuelle oder Gruppenvorurteile zu verhindern, die zu Diskriminierung oder Stigmatisierung von Minderheiten führen könnten,[21]
(4) Erklärbarkeit: Die Transparenz und Kommunikationsfähigkeiten von KI-Systemen sollen verbessert werden, um ihre Entscheidungen nachvollziehen und kontrollieren zu können.

Sieben zu erfüllende Anforderungen

Aus diesen Grundsätzen ergeben sich sieben Anforderungen: (1) »human agency and oversight«, (2) »technical robustness and safety«, (3) »privacy and data governance«, (4) »transparency«, (5) »diversity, non-discrimination and fairness«, (6) »societal and environmental wellbeing« und (7) »accountability«.[22]

Der Grundsatz der »human agency and oversight« ruft dazu auf, menschliche Entscheidungsmacht und Kontrolle über KI zu wahren.[23] Dazu gehört auch die Möglichkeit, KI-gesteuerte Vorgänge zumindest grundsätzlich zu verstehen. Die Forderung nach »technical robustness and safety« beinhaltet, dass die Systeme so konzipiert sind, dass die Schädigung anderer vermieden

20 Dieser Punkt war in der Gruppe durchaus nicht unbestritten. Manche hätten sich eine stärker theoretisch ausgerichtete philosophische Begründung gewünscht. Derartige philosophische Vorgaben konnten aber nicht eindeutig identifiziert werden; außerdem entbehren sie, anders als die EU-Menschenrechte, der rechtlichen Verbindlichkeit. Dies bedeutet aber nicht, dass bei der Interpretation menschenrechtlicher Vorgaben nicht auch philosophische Erwägungen angestellt werden müssen.

21 Zur »Fairness« wird ferner eine gleiche Verteilung von Nutzen und Kosten gerechnet; es soll Möglichkeiten geben, Kompensationen für Schäden (Schadensersatz) zu erhalten.

22 Ethics Guidelines (oben Fn. 6), S. 14ff.

23 Beispiele für einen Übergang von »agency« vom Menschen auf die Maschine bei Köszegi, Sabine: The Autonomous Human in the Age of Digital Transformation, in: Digital Transformation and Ethics (Fn. 1), Heidelberg: Springer 2020, S. 60-84 (71f.).

wird. Für die HLEG AI gehört dazu auch die Sicherung gegen Angriffe von außen. »Privacy and data government« meint nicht bloß den Schutz des Rechts auf informationelle Selbstbestimmung, sondern darüber hinaus die Kontrolle über sämtliche »eigene« Daten.[24] Mit »transparency« ist die grundsätzliche Erklärbarkeit der Arbeitsweise und Arbeitsergebnisse der KI gemeint. »transparency« kann man als eine Voraussetzung der meisten Formen von »accountability« verstehen.[25] Der Punkt »diversity, non-discrimination and fairness« umfasst unter anderem die Sicherung gegenüber einer unfairen, vorurteilsbehafteten KI. Zum »societal and environmental wellbeing« zählt der Schutz der gesamten Gesellschaft, der Umwelt und auch anderer empfindungsfähiger Wesen. Die Forderung nach »accountability« von KI-Systemen schließlich soll sicherstellen, dass im Falle von Schädigungen durch KI angemessene Haftungs- und Verantwortungsmechanismen existieren.

Die Corona-Krise hat die Zusammenarbeit in der Gruppe seit Frühjahr 2020 stark beeinträchtigt.[26] Dennoch können sich die Arbeitsergebnisse sehen lassen. Die Reaktionen in Deutschland und anderer EU-Partner waren überwiegend positiv. Inzwischen haben die Empfehlungen außerdem Eingang in die wissenschaftliche Diskussion gefunden und teilweise auch in konkrete Policy-Vorschläge.[27] Die betroffenen Verbände stimmten den Vorschlä-

24 Der Fokus der Gruppe lag allerdings auf dem Umgang mit personenbezogenen Daten; technische und andere nicht-personenbezogenen Daten und Fragen nach einem möglichen »Dateneigentum« wurden nur am Rande diskutiert.

25 Eine reine Gefährdungshaftung käme wohl ohne Transparenz der KI aus.

26 Eine Erklärung dafür könnte sein, dass sich Kompromisse im persönlichen Miteinander eher finden lassen als in der Kommunikation online.

27 Siehe etwa die Texte von Coeckelburgh, Mark: AI Ethics, V. Dignum: Responsible Artificial Intelligence und S. Köszegi: The Autonomous Human in the Age of Digital Trasformation. Alle drei waren Mitglieder der HLEG. Vgl. ferner Heinz-Uwe Dettling/Stefan Krüger, Erste Schritte im Recht der Künstlichen Intelligenz. Entwurf der »Ethik-Leitlinien für eine vertrauenswürdige KI«, in: MMR 4/2019, S. 211-217 (mit interessanten Hinweisen auf parallele Fragestellungen im Arzneimittelrecht). Die Autoren orientieren sich am Arbeitsentwurf der Leitlinien vom 18.12.2018. Eine Parallele zum Arzneimittelrecht thematisieren auch P. Nemitz/M. Pfeffer: Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, S. 327f. Eine »Innenperspektive« bietet der vorzügliche Artikel von Nathalie A. Smuha, The EU Approach to Ethics Gudelines for Trustworthy Artificial Intelligence. A continuous journey towards an appropriate governance framework for AI, Computer Law Review International (Cri), 4/2019, S. 97-106. Smuha hat die Arbeit der Gruppe über fast zwei Jahre hinweg als Projektassistentin souverän betreut und gemanagt.

gen ganz überwiegend zu,[28] was zeigt, dass die HLEG AI eines ihrer wichtigsten Ziele erreicht hat, nämlich die elitären Zirkel universitärer Ethik-Debatten zu verlassen und praktisch wirksam zu werden. Die Mitglieder der HLEG AI haben ihre Arbeit übrigens von Anfang an als *work in progress* angesehen; es wäre naiv und vermessen zu meinen, man könne in vier knappen Dokumenten die normativen Grundlagen für ein so dynamisches Feld wie die künstlichen Intelligenz ein für alle Mal festschreiben.

Im Folgenden wird zunächst die Aufnahme der HLEG-Empfehlungen im neuen EU-Weißbuch zur künstlichen Intelligenz vorgestellt. Europäische Rechtsvorgaben werden in vielen Teilen der Welt beachtet und oft sogar kopiert. Gelegentlich spricht man gar von einem »Brüssel Effekt«.[29] Deshalb soll auch kurz dargelegt werden, welche Wirkung die EU-Vorschläge in China und den USA erzielen. Schon die Tatsache der Rezeption deutet übrigens darauf hin, dass die These vom »Brüssel Effekt« jedenfalls nicht völlig aus der Luft gegriffen ist. Umso wichtiger wird es, das EU-Regelwerk konstruktiv kritisch zu begleiten und auf die reale Bedeutung für das Gemeinwohl hin zu überprüfen.

2.2 Das EU-Weißbuch zur künstlichen Intelligenz

Am 19. Februar 2020 erschien das *EU White Paper on Artificial Intelligence: a European approach to excellence and trust*.[30] Darin werden die Ansätze der Ethik-Leitlinien und der Policy-Vorschläge weiterentwickelt. Europa sei in der Lage, die Entwicklung und Nutzung von KI an vorderster Stelle voranzutreiben.[31] Besonders bemerkenswert ist die große Bedeutung, die der Analyse und Nutzung von Daten zugemessen wird.[32] Auch wird hervorgehoben, wie wichtig es ist, entsprechende Kompetenzen in der Bevölkerung zu entwickeln.[33] Als Grundlage für das durch angemessene Regulierung zu schaffende »Ecosystem of Trust« werden die sieben Kernvoraussetzungen der oben erwähnten Ethik-

28 Siehe nur das gut durchdachte Positionspapier der Bitkom, https://www.bitkom.org/sites/default/files/2019-02/HLEG_Consultation_Bitkom.pdf

29 Bradford, Anu: The Brussels Effect. How the European Union Rules the World, Oxford: Oxford University Press 2020.

30 COM (2020) 65 final.

31 Ebd., S. 3.

32 Ebd., S. 4.

33 Ebd., S. 6f.

Leitlinien genannt.[34] Mehrfach wird außerdem auf die Gefahren »vorurteilsbehafteter« und diskriminierender KI hingewiesen.[35] Mit Blick auf eventuell erforderliche gesetzgeberische Änderungen erwähnt das Weißbuch insbesondere eine Anpassung des geltenden Produktsicherheits- und Produkthaftungsrechts.[36] Flankiert wird das Weißbuch durch einen detaillierten Bericht über die Sicherheit und Verantwortlichkeit für KI[37] sowie eine Zusammenfassung zur Europäischen Datenstrategie.[38] Im Oktober 2020 fand die 2. European AI Alliance Assembly statt, weitere sind geplant.[39]

2.3 Erste Reaktionen aus China und den USA

Die chinesische Regierung hat die Arbeit der EU HLEG AI nicht direkt kommentiert. Verschiedene offizielle Websites (einschließlich der Parteimedien)[40] haben jedoch über die Veröffentlichung der KI-Leitlinien berichtet oder diese nachgedruckt. Besonders hervorgehoben wird, dass Unternehmen, Forschungsinstitute und Regierungsbehörden die Leitlinien testen sollen. Interessant ist auch die Vermutung, dass die EU möglicherweise versuchen werde, mit den Leitlinien den technologischen Wettbewerb zwischen den USA und China zu unterlaufen, und eine regulative Führungsrolle anstrebe.[41]

---

34 Ebd., S. 9. »human agency and oversight«, »technical robustness and safety«, »privacy and data governance«, »transparency, diversity, non-discrimination and fairness«; »societal and environmental wellbeing«, und »accountability«

35 Ebd., S. 11f. und passim.

36 Ebd., S. 13ff.

37 Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee: Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM (2020) 64 final (vom 19.2.2020). Der Bericht stützt sich auf folgendes Dokument: »Liability for Artificial Intelligence and other Emerging Digital Technologies«, verfasst von der Expert Group on Liablity and New Technologies. New Technology Formation (2019).

38 Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen: Eine europäische Datenstrategie, COM (2020) 66 final (vom 19.2.2020).

39 https://fra.europa.eu/en/news/2020/second-ai-alliance-assembly

40 Für die Übersetzung dieser Seiten und der dort abgedruckten Dokumente und die Unterstützung bei ihrer Auswertung danke ich meinem Mitarbeiter Herrn Liu, Chang sehr herzlich.

41 Siehe oben Fn. 8 zum »Brussels Effect«.

Die Haltung der chinesischen Regierung seither lässt sich (aufgrund einer Reihe von Regierungsbeschlüssen und Reden) grob wie folgt zusammenfassen: Die chinesische Regierung ist der Ansicht, dass 1) eine ethische Regulierung und damit auch ethische Leitlinien notwendig seien, 2) es einige Grundprinzipien gebe, auf die sich unterschiedliche Länder einigen könnten, 3) die EU und die USA verschiedene Entwicklungsrichtungen aufweisen, wobei die EU den Schwerpunkt auf Gesetzgebung und Regulierung lege, während die USA bei der Regulierung einen liberaleren Ansatz verfolge und auf die Förderung technischer Innovation setze, 4) China seinen eigenen Weg zwischen beiden Positionen finden müsse und dabei weder die Innovation ersticken noch technologische und ethische Risiken unkontrolliert zulassen solle. Ethische Standards werden in China meist mit Sicherheitsstandards zusammengebracht, wobei die Entwicklung ethischer Standards insbesondere in Bereichen erfolgt, in denen spezifische ethische Fragen auftreten können, wie zum Beispiel medizinische Behandlung und Notfallmaßnahmen. Zwei Monate nach der Veröffentlichung der EU-Leitlinien legte die chinesische Regierung die *Governance Principles for the New Generation of AI – Developing Responsible Artificial Intelligence* vor.[42] Die darin formulierten Prinzipien unterscheiden sich kaum von den sieben Anforderungen, die durch die HLEG AI aufgestellt wurden.

Auch in den USA wurde über die Arbeit der EU HLEG AI berichtet, allerdings seltener und zurückhaltender als in China. Beispielsweise zieht ein Autor des *Forbes Magazine*[43] insgesamt eine positive Bilanz und betont, dass die EU-Leitlinien über die vielen ähnlichen, ebenfalls nicht verbindlichen ethischen Leitlinien in der Welt hinausgingen und zumindest einen detaillierten Rahmen böten, der Einfluss auf die Regulierungspraxis der Vereinigten Staaten haben könnte. Ein Autor von *The Verge*[44] verweist auch auf die Unver-

42 http://govt.chinadaily.com.cn/a/201906/17/WS5d08a7be498e12256565e009.html. Die acht Grundprinzipien lauten: Harmonie und Freundlichkeit (»Harmony and Human-friendly«), Fairness und Gerechtigkeit (»Fairness and Justice«) Integration und Teilen (»Inclusion and Sharing«), Respekt der Privatsphäre (»Respect for Privacy«), Sicherheit und Kontrollierbarkeit (»Safety and Controllability«), geteilte Verantwortung (»Shared Responsibility«), Offenheit und Kooperation (»Open and Collaboration«) und agile Regulierung (»Agile Governance«).

43 https://www.forbes.com/sites/washingtonbytes/2019/04/11/europes-quest-for-ethics-in-artificial-intelligence/

44 https://www.theverge.com/2019/4/8/18300149/eu-artificial-intelligence-ai-ethical-guidelines-recommendations

bindlichkeit und vermutet, dass die EU anstrebe, die Wettbewerbsfähigkeit der EU auf internationaler Ebene durch die Gestaltung ethischer und rechtlicher Normen zu gewährleisten, da Investitionen und Spitzenforschung nicht mit den USA und China konkurrieren könnten. Besonders positiv fiel die Resonanz von Microsoft aus.[45] Die Leitlinien werden als Meilenstein für die ethische und rechtliche Regulierung von KI bewertet. Der Konzern gibt an, die von der HLEG AI formulierten Werte zu teilen, und erklärt sich bereit, an entsprechenden Tests und weiteren Experimenten teilzunehmen. Bemerkenswerterweise unterzeichnete schließlich auch der damalige US-Präsident Donald Trump am 3. Dezember 2020 eine Executive Order zum Einsatz von KI durch die US-amerikanische Regierung, die schon im Titel den Bezug zu den EU-Leitlinien erkennen lässt, ohne sie aber im Text zu erwähnen. Dagegen sind die inhaltlichen Überschneidungen offensichtlich.[46]

III Legislative Herausforderungen – wie könnte ein europäischer Weg bei der Regulierung von KI aussehen?

Die Empfehlungen der HLEG AI sowie das EU-Weißbuch zur künstlichen Intelligenz bieten eine Grundlage für weitergehende Überlegungen, wie die neuen digitalen Technologien reguliert werden sollten. Dabei lassen sich zumindest folgende, besonders wichtige Themenbereiche unterscheiden:

3.1 Haftung und strafrechtliche Verantwortung

Wer für was, wann und wie Verantwortung im Internet oder bei digitalen Prozessen übernimmt, ist eine der drängendsten und schwierigsten ethischen und rechtspolitischen Fragen in der digitalen Welt. Das Thema *Haftung und Verantwortung für von Maschinen verursachte Schäden* wird bereits seit Län-

45 https://blogs.microsoft.com/eupolicy/2019/04/09/ethical-guidelines-trustworthy-ai/

46 Executive Order on Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, https://www.whitehouse.gov/presidential-actions/executive-order-promoting-use-trustworthy-artificial-intelligence-federal-government. Vgl. auch schon die Executive Order on Maintaining American Leadership in Artificial Intelligence, unter https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence, vom 11.2.2019.

gerem intensiv diskutiert.[47] Durch den Einsatz autonomer Systeme und KI drohen Haftungslücken und eine damit einhergehende Diffusion von Verantwortung, die mit einem sozialstaatlichen Werten verpflichteten und auf Schadensausgleich angelegten Gemeinwesen schwer vereinbar erscheinen.[48] Gerade am Arbeitsplatz sollte sichergestellt sein, dass bei der Verletzung von Menschen durch Maschinen ein angemessener Schadensausgleich erfolgt. Mit der überkommenen Verschuldenshaftung ist die Zuschreibung von Verantwortung in der Beziehung von Mensch und Maschine oft nur schwer möglich. Deshalb treten manche für die Einführung einer E-Person, also einer elektronischen Person, als Haftungssubjekt ein, was es ermöglichen würde, die Maschinen selbst auf Schadensersatz zu verklagen.[49]

Ein Beispiel mag dies verdeutlichen: Vor einigen Jahren stellte Microsoft den lernfähigen Chatbot »Tay« online, der mit Menschen Gespräche führen und so seine kommunikativen Fähigkeiten perfektionieren sollte. Unerkannt gebliebenen Hackern gelang es, das System so zu beeinflussen, dass es rassistische und frauenfeindliche Äußerungen abgab. Daraufhin musste »Tay« vom Netz genommen werden. Angenommen, durch die Äußerungen des Chatbots

47 Zusammenfassend Hilgendorf, Eric: Zivil- und Strafrechtliche Haftung für von Maschinen verursachte Schäden, in: Bendel, Handbuch Maschinenethik (Fn. 1), S. 437-452; vertiefend für das Zivilrecht jüngst Spindler, Gerald: Haftung für autonome Systeme – ein Update, in: Beck u.a. (Hg.), Digitalisierung, Automatisierung, KI und Recht (Fn. 1), S. 255-284; Zech, Herbert: Gutachten A zum 73. Deutschen Juristentag Hamburg 2020/Bonn 2022: Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?, München: C.H. Beck 2020; für das Strafrecht Joerden, Jan: Zur strafrechtlichen Verantwortlichkeit bei der Integration von (intelligenten) Robotern in einen Geschehensablauf, in: Beck u.a. (Hg.), Digitalisierung, Automatisierung, KI und Strafrecht (Fn. 1), Baden-Baden: Nomos 2020, S. 287-304; Schuster, Frank: Künstliche Intelligenz, Automatisierung und strafrechtliche Verantwortung ebd., Baden-Baden: Nomos 2020, S. 387-400.

48 Zum Problem der Verantwortungsdiffusion Hilgendorf, Eric: Verantwortungsdiffusion und selbstlernende Systeme in der Industrie 4.0 – ein Problemaufriß aus strafrechtlicher Perspektive, in: Gerrit Hornung (Hg.), Rechtsfragen der Industrie 4.0. Datenhoheit, Verantwortlichkeit, rechtliche Grenzen der Vernetzung, Baden-Baden: Nomos 2018, S. 119-137.

49 Die »E-Person« wäre juristisch als Sonderform einer juristischen Person anzusehen, von der man wie z.B. von einer GmbH oder einer anderen juristischen Person Schadensersatz verlangen könnte. Praktisch würde dies des Weiteren erfordern, haftungsfähigen Maschinen eine hinreichend große Vermögenssumme zuzuordnen, was aber z.B. über eine obligatorische Versicherung leicht zu erreichen wäre.

wäre ein finanzieller Schaden aufgetreten (etwa infolge eines behandlungsbedürftigen Traumas bei einer verbal attackierten Person) – wer wäre dann zum Schadensersatz verpflichtet gewesen? Die Hacker waren nicht zu belangen, der Hersteller beziehungsweise Programmierer konnte darauf verweisen, dass sein System fehlerfrei funktioniert habe. Demnach wäre das Opfer auf seinem Schaden sitzengeblieben. Hier könnte das Modell einer E-Person helfen, als neuer Form einer juristischen Person, die es Betroffenen ermöglichen würde, das Computersystem selbst zur Verantwortung zu ziehen.[50]

Das Konzept einer E-Person entstammt allerdings angelsächsischen Rechtsvorstellungen und ist mit der europäischen, zumal der deutschen, Rechtstradition nicht ohne Weiteres zu vereinbaren, obwohl die Schaffung eines E-Person gesellschaftsrechtlich wohl möglich wäre. Vorzugswürdig erscheint es deshalb, der Gefahr von Haftungslücken und Haftungsdiffusion zu begegnen, indem die (schuldunabhängige) Gefährdungshaftung ausgeweitet wird und etwa die Produkthaftung auch für unkörperliche Produkte wie Algorithmen gilt. Allerdings dürfte noch sehr viel gesetzgeberische Detailarbeit erforderlich sein, bis eine den Leitwerten Europas angemessene Verteilung der Haftungsrisiken im Zusammenhang mit KI erreicht ist.

Noch problematischer ist die Situation im Strafrecht, in dem wegen des (in Deutschland auch in der Verfassung festgeschriebenen) Schuldgrundsatzes (niemand kann für eine Tat bestraft werden, wenn ihn keine Schuld trifft) eine Verantwortung von Maschinen von vornherein ausgeschlossen ist. Überlegungen, Maschinen strafrechtlich zu belangen (etwa aus Gründen der Generalprävention), haben allenfalls den Charakter von (durchaus interessanten!) Gedankenexperimenten, könnten jedoch praktisch nicht ohne massive Verletzungen zentraler Basisannahmen rechtsstaatlichen Strafens in Europa umgesetzt werden. Bis auf Weiteres muss hier also die Möglichkeit von Strafbarkeitslücken akzeptiert werden; bislang scheint es übrigens kaum reale Fälle zu geben, in denen die Strafbarkeit einer Maschine (etwa »Tay«) sinnvoll wäre.[51]

50 Näher dazu Hilgendorf, Eric: Autonome Systeme, Künstliche Intelligenz und Roboter, in: Stephan Barton u.a. (Hg.), Festschrift für Thomas Fischer, München: C.H. Beck 2018, S. 99-113 (109f.).

51 Ebd. E. Hilgendorf, Autonome Systeme, Künstliche Intelligenz und Roboter, S. 110.

3.2 Schutz von Persönlichkeitsrechten – auch mittels des Strafrechts

Ein zweiter wichtiger Problemkreis ist der Schutz von Persönlichkeitsrechten, die im Internet, zumal den sozialen Netzwerken, offenbar zunehmend auch durch Bots und autonome Systeme angegriffen werden. Die Probleme, die sich hier stellen, sind außerordentlich vielschichtig und entsprechend schwierig zu lösen:

In den USA wird die in der Verfassung festgelegte Redefreiheit (*freedom of speech*) von den Gerichten so weit ausgedehnt, dass Persönlichkeitsrechte in aller Regel dahinter zurückzutreten haben.[52] Ein Beleidigungsstrafrecht in unserem Sinne, durch das ein Minimum an zwischenmenschlichem Respekt gesichert wird, existiert nicht. Dieses extrem weite Verständnis von Redefreiheit findet sich im Wesentlichen nur in den USA; im Rest der Welt, angefangen von Ländern des angelsächsischen Rechtskreises wie Großbritannien oder Kanada über Kontinentaleuropa bis hin zu Lateinamerika oder Ostasien, gelten Regelungen zum Schutz der Persönlichkeitsrechte. Da aber die Internettechnologie und insbesondere die sozialen Netzwerke von US-Anbietern dominiert werden, zählen dort grundsätzlich die US-amerikanischen Standards, die so über die USA hinaus Geltung beanspruchen, ohne nennenswerte Mitspracherechte der davon Betroffenen vorzusehen. Dies führt in vielen Ländern zu erheblichen Problemen bei der Umsetzung nationalen Persönlichkeitsschutzrechts, wie sich auch während den hitzigen Debatten um das inzwischen reformierte Netzwerkdurchsetzungsgesetz[53] zeigte.

Phänomene in den sozialen Netzwerken, wie Hassrede, Fake News, sexuelle Anzüglichkeiten und Cybermobbing, werden inzwischen auch in den USA als massives Problem empfunden. Der Verzicht auf eine gesetzliche Kontrolle von Hassrede,[54] der sich früher vor allem zulasten der afroamerikanischen Bevölkerung und anderer unterprivilegierter Minderheiten auswirkte,

52 Vgl. T. Garton Ash, Redefreiheit. Prinzipien für eine vernetzte Welt, S. 198ff. In fast allen anderen Ländern existiert ein Beleidigungsstrafrecht, das besonders drastische Verletzungen zwischenmenschlichen Respekts mit Strafe belegt. Siehe für Deutschland etwa Hilgendorf, Eric: Beleidigungsstrafrecht, in: Eric Hilgendorf, Hans Kudlich und Brian Valerius (Hg.), Handbuch des Strafrechts, Band 4, Heidelberg: C.F. Müller 2019, § 12.

53 Die Reform des NetzDG wurde Anfang April 2021 vom Bundespräsidenten unterzeichnet.

54 Zu Kompensationsphänomenen wie »political correctness« E. Hilgendorf, Beleidigungsstrafrecht (Fn. 52), Rn. 8.

hat im Zeitalter sozialer Netzwerke zu einer bisher nicht dagewesenen Spaltung der US-amerikanischen Gesellschaft geführt. Seit einigen Jahren wird darüber spekuliert, dass viele der besonders enthemmten Posts möglicherweise gar nicht von Menschen stammen, sondern von KI-gestützten Bots, die mehr oder weniger autonom, aber nicht ohne Ziel aktiv sind. Die rechtliche Begrenzung und Kontrolle derartiger Einsatzformen von KI[55] gehört weltweit zu den wichtigsten, neuen rechtspolitischen Aufgaben.

In diesen Zusammenhang gehört auch die Frage nach einer angemessenen zivilrechtlichen und strafrechtlichen Haftung von Providern und Intermediären. Bislang wurden Diensteanbieter im Internet europaweit von zivil- und strafrechtlicher Haftung weitgehend freigestellt.[56] Ziel der Regelung war es seinerzeit, der sich entwickelnden Internetwirtschaft und der Entfaltung des offenen Internets keine unnötigen Steine in den Weg zu legen und klarzustellen, dass ein Provider nicht ohne Weiteres für rechtswidrige Inhalte haftet, zu denen er technisch den Zugang eröffnet. Ob diese Haftungsprivilegien heute noch zeitgemäß sind, ist sehr zweifelhaft.

3.3 Diskriminierungsfreiheit

Eines der schwierigsten Themen bezüglich der Regulierung von KI stellt der Umgang mit potenziell vorurteilsgeprägter Technologie dar.[57] Ein Ausgangspunkt vieler Debatten ist die von US-Gerichten verwendete Software Correctional Offender Management Profiling for Alternative Sanctions (COMPAS). Damit wird unter anderem die Rückfallwahrscheinlichkeit von Straftätern eingeschätzt und so die richterliche Entscheidung unterstützt. Dem Hersteller wurde vorgeworfen, das System würde infolge selektiver Datenauswahl Menschen mit Afroamerikanischer Herkunft benachteiligen, es arbeite mit

55 Vgl. Grunwald, Arnim: Der unterlegene Mensch. Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern, S. 167ff. bezeichnet derartige Algorithmen geradezu als »Totengräber der Demokratie«.

56 Diese Privilegierung geht zurück auf die Richtlinie 2000/31/EG des Europäischen Parlaments und des Rates vom 8. Juni 2000 über bestimmte rechtliche Aspekte der Dienste der Informationsgesellschaft, insbesondere des elektronischen Geschäftsverkehrs, im Binnenmarkt (»Richtlinie über den elektronischen Geschäftsverkehr«) Amtsblatt Nr. L 178 vom 17/07/2000 S. 0001-0016. In Deutschland wurden diese Vorgaben im Telemediengesetz umgesetzt.

57 Vgl. Coeckelbergh, Mark: AI Ethics, S. 125ff. Siehe außerdem den Beitrag von Hustedt und Beining in diesem Band.

einem Bias gegenüber Schwarzen.[58] In anderen Studien wurde dem jedoch widersprochen;[59] auch das Oberste Gericht des US-Bundesstaates Wisconsin hielt die Verwendung von COMPAS für zulässig.[60]

KI entscheidet anhand der ihr zur Verfügung gestellten Daten. Schon darin kann man einen Vorteil gegenüber manchen menschlichen Entscheidungen sehen. Auch lassen sich Maschinen nicht durch Emotionen oder eigene Interessen beeinflussen. Allerdings hängt die Qualität maschineller Entscheidungen von der Qualität des Dateninputs ab, schlechte Daten bewirken schlechte Entscheidungen.[61] Daher wäre es verfehlt, Entscheidungen – oder entsprechende Vorschläge – durch Maschinen von vornherein für objektiver zu halten als menschliche Entscheidungen. Bei der Bewertung einer maschinellen Entscheidung sollte die Datenbasis stets mit geprüft werden.

Ein zweiter Aspekt tritt hinzu. Als Beispiel mag eine KI dienen, die anhand aller öffentlich verfügbaren Daten Entscheidungsvorschläge für die Besetzung von Vorstandsposten erarbeiten soll. Da in der Vergangenheit derartige Positionen ganz überwiegend von älteren Männern besetzt waren, schlägt die KI weiterhin in erster Linie Personen mit eben diesen Eigenschaften vor. Ähnliche Beispiele ließen sich für die Besetzung einer Stelle als Hebamme oder Fachkraft im Kindergarten bilden. Das wesentliche Problem liegt wohl im Folgenden: Auch wenn eine bestimmte Gruppe oder Menschen mit bestimmten Eigenschaften in der Vergangenheit stets mit einer bestimmten Stellung oder Tätigkeit in Verbindung gebracht werden konnten, ist es nicht ohne Weiteres zwingend und möglicherweise sogar problematisch, diesen Zustand in die Zukunft zu verlängern. Der auf früheren Daten aufgebaute maschinelle Entscheidungsvorschlag ist von einem Konservativismus geprägt, der seinerseits einer Begründung bedarf. Möglicherweise sprechen normative Gesichtspunkte dagegen, die Vergangenheit einfach in

58 Vgl. Angwin, Julia/Larson, Jeff u.a.: »Machine Bias« https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (23.5.2016), weiterführend und das Problem kontextualisierend Lobe, Adrian: Speichern und Strafen. Die Gesellschaft im Datengefängnis, 2019, S. 173ff. (Standardbestimmung durch KI), S. 186ff. (COMPAS).

59 Vgl. Flores, Anthony W./Lowenkamp, Christopher T./Bechtel, Kristin: False Positives, False Negatives, and False Analyses: A Rejoinder to »Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks.« https://www.crj.org/assets/2017/07/9_Machine_bias_rejoinder.pdf

60 Vgl. https ://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690

61 Informatiker nutzen oft die Formulierung »garbage in – garbage out«.

die Zukunft zu übertragen. Eine derartige »normative Kontrollebene« fehlt der maschinellen Intelligenz.

Allerdings sollte man es sich hier nicht zu einfach machen. In vielen Fällen wird eine auf feststehenden Daten beruhende maschinelle Einschätzung zutreffend und überzeugend sein. Nehmen wir an, ein KI-System soll anhand öffentlich verfügbarer Daten einen Vorschlag darüber unterbreiten, ob die Stelle eines Fahrrad-Pizzaboten an einen 20-jährigen Sportler oder eine 75-jährige Rentnerin vergeben wird. Unter Berücksichtigung der bisherigen Besetzung ähnlicher Positionen empfiehlt die Maschine den 20-Jährigen. Liegt darin eine ungerechtfertigte Diskriminierung? Die meisten würden das wohl verneinen, weil das Lebensalter und die damit normalerweise einhergehende körperliche Leistungsfähigkeit für die angebotene Stelle entscheidend sind. Das Alter ist hier also ein guter Grund für eine Ungleichbehandlung und seine Berücksichtigung stellt keine ungerechtfertigte Diskriminierung dar.

Damit wird deutlich, dass das Problem potenziell vorurteilsbehafteter KI eng mit dem wesentlich weitergehenden Problem verknüpft ist, welche Gesichtspunkte wir als zulässige Kriterien einer Ungleichbehandlung ansehen. Letzteres ist eine gesellschaftliche Frage, die nicht nur in zeitlicher Perspektive, sondern auch von Gruppe zu Gruppe unterschiedlich beantwortet wird. Rechtliche Vorgaben lassen sich in Artikel 3 Grundgesetz (Gleichheitsgrundsatz), aber auch im Allgemeinen Gleichbehandlungsgesetz finden; ihre Übertragbarkeit auf KI-generierte Entscheidungsvorschläge oder Entscheidungen ist aber im Detail noch ungeklärt. Die Forderung nach Transparenz darf sich jedenfalls nicht bloß auf Daten beziehen, sondern muss auch die zugrunde gelegten Kriterien umfassen. Die Analyse und systematische Aufarbeitung der damit zusammenhängenden Fragen stellt derzeit eine der wichtigsten Forschungsaufgaben und politisch-regulativen Herausforderungen im Zusammenhang mit KI dar.

3.4 Transparenz und Erklärbarkeit von KI-Entscheidungen

Das Thema explicabilty beziehungsweise transparency, also die Erklärbarkeit beziehungsweise Transparenz von KI-Systemen, spielte in den Diskussionen der HLEG AI eine erhebliche Rolle, da es sich um ein neues Prinzip handelt, das im traditionellen Diskurs über Menschenrechte so noch nicht vorkam.[62]

62 Eine gewisse Parallele lässt sich allerdings zu den in den 1990er und früher 2000er Jahren verbreiteten Vorstellungen ziehen, das seinerzeit neue Internet könne helfen,

Außerdem gab und gibt es erhebliche Meinungsunterschiede darüber, worauf explicability oder explainability genau abzielen.[63] Wegen seines innovativen Charakters und der daraus entstehenden Interpretationsoffenheit überschneiden sich explicability und transparency mit anderen Themen, etwa accountability oder responsibility, meines Erachtens, ohne dass bisher eine hinreichend klare Abgrenzung gelungen wäre.[64] In den Ethik-Leitlinien heißt es, explicability sei:

> »crucial for building user's trust in AI systems. This means that processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – explainable to those directly or indirectly affected. Without such information, a decision cannot be duly contested.«[65]

Es liegt auf der Hand, dass eine so verstandene »Erklärbarkeit« bei vielen Systemen an technische Grenzen stößt. Im Bereich des Deep Learnings scheint es sogar von vornherein ausgeschlossen zu sein, dass es zu erklären ist, warum die (sich selbstständig weiterentwickelnde) Maschine zu einem bestimmten Ergebnis gekommen ist. Man sollte überdies bedenken, dass Detailinformationen über Arbeitsabläufe im Computer und die Art und Weise der Ergebnisgenerierung in vielen Fällen Betriebsgeheimnisse sind, die die betroffenen Firmen weder offenbaren wollen noch (nach derzeitigem Gesetzesstand) müssen. Schließlich sei der Hinweis erlaubt, dass die entsprechenden Prozesse im menschlichen Gehirn ebenfalls im Dunklen liegen; warum jemand zu einer bestimmten Aussage oder Wertung gelangt, kann im Detail nicht nachvollzogen werden, auch wenn Ex-ante-Prognosen und Ex-post-Erklärungen möglich sind, die allerdings auch fehlerhaft sein können, wie die Lebenserfahrung zeigt.

---

politische Prozesse »transparenter« zu gestalten. Damals bezog sich die Transparenzforderung allerdings auf (menschliche) Entscheidungsprozesse, heute auf maschinell generierte Entscheidungen.

63 Guter Überblick bei M. Coeckelbergh, AI Ethics, S. 116ff.

64 Auch zum Grundsatz der »human agency« existieren Überschneidungen. So heißt es auf S. 16 der »Ethics Guidelines (Fn. 6): »Human agency. Users should be able to make informed autonomous decisions regarding AI systems. They should be given the knowledge and tools to comprehend and interact with AI systems to a satisfactory degree and, where possible, be enabled to reasonably self-assess or challenge the system. AI systems should support individuals in making better, more informed choices in accordance with their goals.«

65 Ethics Guidelines (Fn. 6), S. 13.

Andererseits erscheint es nicht bloß ethisch, sondern auch rechtlich erforderlich, KI-gestützte Entscheidungen und die zugrunde liegenden Kriterien transparent zu gestalten und bei Unklarheiten eine Erklärung fordern zu können. Ohne eine solche Transparenz würde die Klärung eventueller Haftungsfragen erheblich erschwert, von dem rechtsstaatlichen Erfordernis einer Erklärung KI-gestützter hoheitlicher Entscheidungen ganz zu schweigen. Es gehört deshalb zu den wesentlichen Aufgaben des KI-Rechts, die Konzepte *Erklärbarkeit* und *Transparenz* von KI-gestützten Entscheidungen zu vertiefen und ihnen einen Ort im überkommenen Rechtssystem, gerade im Verwaltungs- und Haftungsrecht, zuzuweisen.

3.5 Schutz der Privatsphäre und Datenhoheit[66]

Alle technisch fortgeschrittenen Gesellschaften der Erde, auch und gerade in Europa, entwickeln sich derzeit in Richtung auf das »panoptische Modell«, in dem einige wenige die Kontrolle über die Daten der überwältigenden Mehrheit der Menschen besitzen.[67] So ist offensichtlich, dass immer mehr Daten aus Europa von ausländischen, vor allem US-amerikanischen Großunternehmen erhoben und dort kommerziell genutzt werden, ohne dass erstere irgendein Mitspracherecht hätten oder gar an dem aus der Verwendung »ihrer« Daten entstehenden Profit irgendwie beteiligt wären. Im Mai 2018 ist die Europäische Datenschutz-Grundverordnung (DSGVO) in Kraft getreten. Damit wurde das traditionelle Datenschutzmodell noch einmal eindrucksvoll aktualisiert und revitalisiert. Die DSGVO liegt auch den entsprechenden Konzepten der HLEG AI zugrunde. Sie hat sich in den ersten drei Jahren ihrer Geltung als erfolgreicher erwiesen, als es anfangs vorausgesagt wurde. Dennoch ist nicht zu übersehen, dass sich der Datenschutz in Deutschland – und Deutschland steht hier pars pro toto für ganz Europa – in einer Krise befindet.

Schon die Bezeichnung »Datenschutz« ist irreführend, denn streng genommen werden nicht Daten geschützt, sondern das Grundrecht auf informationelle Selbstbestimmung, also das Recht, über die auf die eigene Person

66 Siehe auch den Beitrag von Nils Leopold in diesem Band.

67 Es sei die Anmerkung erlaubt, dass Jeremy Bentham, auf dessen Ideen die Vorstellung eines »panoptischen Modells« zurückgehen, das Problem »Wer kontrolliert die Kontrolleure?« gesehen und sogar eine bemerkenswerte Lösung dafür angeboten hat: die Öffentlichkeit! Bedauerlicherweise scheint Benthams Problembewusstsein heute verloren gegangen zu sein.

bezogenen Daten selbst zu bestimmen. Der Fokus auf personenbezogene Daten zeigt, dass es sich beim überkommenen Datenschutzrecht im Kern um eine Kommunikationsordnung handelt, die den Umgang mit personenbezogenen Daten regelt. Durch die gewaltigen Fortschritte der Digitalisierung in den letzten zwei Jahrzehnten hat sich die Problematik aber über den Schutz personenbezogener Daten hinaus hin zur Frage gewandelt, wie sich der Umgang mit Daten jeder Art regulieren lässt. Daten besitzen heute, anders als noch vor zwei Jahrzehnten, einen gewaltigen ökonomischen Wert, sie sind zu einem bedeutenden Wirtschaftsgut geworden. Es geht nicht mehr nur um Kommunikation, sondern um den Rohstoff für datengetriebene Geschäftsmodelle und damit letztlich um wirtschaftlichen Erfolg.

Umso fataler ist es, dass das europäische Datenschutzrecht nur personenbezogene Daten erfasst und der Umgang mit sowie der Schutz von Daten anderer Art, etwa technischer Daten, nicht angemessen reguliert werden. Da sie keine Sachqualität aufweisen, sind Daten nicht eigentumsfähig, die Rede von den »eigenen« Daten ist, wenn sie sich nicht auf eigene personenbezogene Daten bezieht, juristisch gesehen irreführend. Es existiert derzeit keine allgemein akzeptierte Möglichkeit, Daten originär eigentumsrechtlich zuzuordnen. Dies bedeutet unter anderem, dass nicht-personenbezogene Daten fast nach Belieben abgezogen und verwertet werden dürfen, ohne dass die Betroffenen dagegen Einspruch erheben können.[68]

Die meisten Menschen sind dem Datenschutz gegenüber sehr gleichgültig: Datenschutzverletzungen werden, sofern eine breitere Öffentlichkeit überhaupt davon erfährt, mehr oder weniger teilnahmslos hingenommen. Damit hängt ein paradoxes Phänomen zusammen: Auf der einen Seite fordern Viele vom Staat zu Recht ein hohes Maß an Privatsphäre und einen besonderen rechtlichen Schutz ihrer personenbezogenen Daten. Auf der anderen Seite ist der ganz überwiegende Teil der Bevölkerung ohne größere Bedenken bereit, seine Daten ausländischen (Quasi-)Monopolisten zur Verfügung zu stellen, wenn dafür (scheinbar) kostenfrei ein Dienst in Anspruch genommen oder eine App genutzt werden kann. Die Informationspflichten der Datenverarbeiter, die der beziehungsweise dem Einzelnen die Konsequenzen einer Einwilligungserteilung vor Augen führen sollen, werden (faktisch unbeanstandet) in Gestalt mehrseitiger Dokumente erfüllt, von

68 Vgl. Hilgendorf, Eric: Offene Fragen der neuen Mobilität: Problemfelder im Kontext von automatisiertem Fahren und Recht, Frankfurt a.M.: Recht – Automobil – Wirtschaft (RAW) 2018, S. 85-93 (89f.).

denen allgemein bekannt (und überdies häufig auch so gewollt) ist, dass sie nur von einem Bruchteil der Betroffenen gelesen werden. Dieses Phänomen setzt sich bei nicht-personenbezogenen Daten fort. Ein wirksames Datenschutzrecht müsste in puncto Information und Aufklärung mehr bieten als einfach zu erfüllende Informationspflichten, wenn es von mündigen Grundrechtsberechtigten wahrgenommen werden soll.

Die technische Entwicklung lässt außerdem zentrale Grundsätze des überkommenen Datenschutzrechtes als unzeitgemäß erscheinen. Prinzipien wie das der Datenminimierung (es sollen so wenig Daten wie möglich aufgenommen werden) und Zweckbindung (Daten sollen nur zu dem Zweck verwendet werden dürfen, zu dem sie erhoben wurden) sind im Zeitalter von Big Data und KI nahezu sinnlos geworden, da die neuen Geschäftsmodelle gerade voraussetzen, möglichst viele Daten einzusammeln, die dann zu beliebigen Zwecken im Data Mining verwendet werden können. Eine derart groß angelegte Datenanalyse ist keineswegs per se verwerflich, sie kann vielmehr durchaus sinnvoll und gesellschaftlich erwünscht sein. So lassen sich etwa in einem großen Fundus medizinischer Daten möglicherweise Muster erkennen und Korrelationen finden, die helfen, Krankheitsursachen zu identifizieren.

Um KI im medizinischen Sektor in großem Umfang einsetzen zu können, muss diese KI allerdings mit Daten, und zwar mit möglichst vielen Daten, trainiert werden. Dies ist mit dem geltenden Datenschutzrecht nicht ohne Weiteres vereinbar, da das Sammeln größerer Datenmengen von vornherein durch das Regelungsmodell »Verbot mit Erlaubnisvorbehalt« beschränkt wird. Der verbreitete und oft gänzlich unreflektierte »Datenschutz-Absolutismus«, der die gesetzlich verbürgten Einschränkungsmöglichkeiten beim Datenschutz[69] ignorieren zu können glaubt, verschärft die ohnehin große Gefahr einer Monopolbildung bei ausländischen Anbietern und läuft dem Gemeinwohl zuwider. Eine vorherige Anonymisierung der Daten ist schwer möglich, zumal sich heute mit hinreichendem Aufwand praktisch alle Daten wieder mit einem Personenbezug versehen lassen. Ohne Übertreibung

69 Kein Grundrecht, die Menschenwürde ausgenommen, gilt schrankenlos; vielmehr können alle Grundrechte eingeschränkt werden, um höher zu gewichtende Belange des Gemeinwohls zu verwirklichen. Dabei ist allerdings stets der Grundsatz der Verhältnismäßigkeit zu beachten. So enthält etwa die DSGVO in Art. 89 eine großzügige Öffnungsklausel für die Forschung; bei medizinischen Daten ist außerdem Art. 9 DSGVO zu beachten.

lässt sich sagen, dass durch den technischen Fortschritt bereits das Konzept des »anonymisierten Datums« als solches fragwürdig geworden ist.

Gerade im medizinischen Bereich, der zunehmend in den Mittelpunkt rückt, droht eine Monopolisierung der Daten und damit des verfügbaren Wissens. Es dürfte kein anderes Gebiet geben, in dem der Grundsatz *The winner takes it all* in dem Maße gilt wie in der Medizin.[70] Wenn es um die Gesundheit oder gar das Leben der eigenen Person oder naher Angehöriger geht, ist niemand bereit, sich mit der zweitbesten Lösung zufriedenzugeben, und kein Aufwand ist zu groß, wenn er nur Hilfe verspricht. Eine Abhängigkeit von kommerziell orientierten und nicht mehr regulierbaren außereuropäischen Mega-Unternehmen wäre hier fatal.

Der Befund lässt sich so zusammenfassen: Das Recht auf informationelle Selbstbestimmung ist heute wichtiger denn je. Gleichzeitig wird es aber so massiv bedroht, dass ein Überdenken des bisherigen Schutzansatzes dringend nötig geworden ist. Datenschutz ist kein Selbstzweck, sondern muss sich am Gemeinwohl orientieren. Ein neuer Ansatz ist auch deshalb zentral, weil das überkommene Datenschutzrecht lediglich personenbezogene Daten erfasst und nicht-personenbezogene Daten, etwa Daten technischer Art, außer Betracht lässt, obwohl gerade diese Daten inzwischen für zahlreiche Geschäftsmodelle eine besonders große Rolle spielen. Mit ihrer neuen »Datenstrategie«[71] hat die Bundesregierung einen großen Schritt in die richtige Richtung unternommen; ob die wohlklingenden Worte auch umgesetzt werden, bleibt abzuwarten.

3.6 Technologischer Paternalismus

Ein weiteres schwieriges Problemfeld eröffnet sich mit der Frage, ob beziehungsweise inwieweit durch technische Mittel wie KI Rechtsverstöße erschwert, ganz unmöglich gemacht oder zumindest automatisiert sanktioniert werden dürfen oder sollten. Man kann das Problem anhand eines Beispiels aus dem Straßenverkehr verdeutlichen: Statt Geschwindigkeitsüberschreitungen oder das Überfahren roter Ampeln zu verbieten und im

70 Vgl. Hilgendorf, Eric: Medizin und Digitalisierung, Freiburg: ContraLegem 2019, S. 274-282 (280), https://www.contralegem.ch/2019-2-l-medizin-und-digitalisierung-(e-health)

71 https://www.bundesregierung.de/breg-de/themen/digitalisierung/datenstrategie-1693546

Entdeckungsfall mit Bußgeldern zu ahnden, ließen sich Fahrzeuge mit autonomen und vernetzten Systemen von vornherein so gestalten, dass ein Verstoß gegen die Verkehrsordnung unmöglich wäre. Derartige Fahrzeuge könnten gar nicht mehr mit 150 Kilometern pro Stunde in einer Innenstadt unterwegs sein, weil eine mit der Verkehrsüberwachung betraute KI sie schon lange vor Erreichen dieser Geschwindigkeit abbremsen würde.[72]

Im angelsächsischen Schrifttum werden ähnliche Probleme gelegentlich unter dem Stichwort »impossibility structures« behandelt.[73] Dieser Begriff dürfte die Problematik allerdings kaum angemessen bezeichnen, denn es geht meist nicht darum, Fehlverhalten ganz unmöglich zu machen, sondern nur darum, es zu erschweren beziehungsweise zu dokumentieren. So wäre es im obigen Beispiel eines autonomen und vernetzten Fahrzeugs fatal, wenn der Wagen unter keinen Umständen mehr die vorgeschriebene Höchstgeschwindigkeit überschreiten könnte. In Notfällen, etwa bei einem Krankenhaustransport, muss es möglich sein, die von der KI gezogene Grenze zu überwinden. Allerdings sollten solche Fälle automatisch dokumentiert werden, sodass sie später (juristisch) auf ihre Berechtigung überprüft werden können. Dabei sollten die technischen Möglichkeiten einer Notfall-Übersteuerung durchaus unterschiedlich ausgestaltet sein; zum Beispiel sollte die Überwindung eines »Alkolocks« (Wegfahrsperre bei Alkoholisierung des Fahrers oder der Fahrerin) nur gelingen, wenn er oder sie durch Ausschalten einer entsprechenden Sicherung die eigene Fahrtauglichkeit bewiesen hat.

Wie die Beispiele zeigen, wird es also im Regelfall nicht um die 100-prozentige faktische Verhinderung von Fehlverhalten gehen, sondern um seine Erschwerung. Es handelt sich meist um »safety by default«-Einstellungen, die eine mehr oder weniger starke Präventionswirkung besitzen. Denkbar ist sogar, dass eine überwachende KI sich auf Informationen, Warnhinweise oder (mehr oder weniger stark ausgestaltete) Anreize zu korrektem Verhalten beschränkt. Gemeinsam ist allen diesen Fällen, dass das Fahrverhalten technisch überwacht und zum Wohle des Fahrers beziehungsweise der Fahrerin und Dritter gesteuert wird. Man kann die hier einschlägige Problemklasse deshalb als »technologischen Paternalismus« bezeichnen.[74]

---

72 Im Flugverkehr sind derartige Technologien schon seit Langem im Einsatz.

73 Siehe hierzu auch den Beitrag von Timo Rademacher und Erik Schilling in diesem Band.

74 E. Hilgendorf: Offene Fragen der neuen Mobilität: Problemfelder im Kontext von automatisiertem Fahren und Recht, S. 92.

Technologischer Paternalismus wirft eine Fülle von Problemen auf: Zwar vermag die Technik, Leben und andere wichtige Rechtsgüter zu schützen, technologischer Paternalismus impliziert aber eine weitreichende, unter Umständen dauerhafte Beobachtung von Personen in risikoträchtigen Situationen (etwa im Straßenverkehr oder bei chirurgischen Eingriffen). Dies führt, neben datenschutzrechtlichen Bedenken, zu grundsätzlichen Fragen nach unserem Freiheits- und Autonomieverständnis. Darüber hinaus geht es um die Belange des Gemeinwohls, das durch einen verstärkten Einsatz intelligenter Technik zur Risikokontrolle und Risikoverhinderung erheblich befördert werden könnte. Problematisch ist weiter, dass vernetzte Technik stets gehackt und manipuliert werden kann. Aus einer philosophisch-theologischen Perspektive ließe sich schließlich fragen, ob sittliches Verhalten nicht die faktische Möglichkeit von Fehlverhalten voraussetzt.

3.7 Private Quasi-Monopole und der Bedeutungsverlust des Staates

Europäische Werte werden sich in der digitalisierten und damit global vernetzten Welt nur dann durchsetzen lassen, wenn hinreichend viele Menschen sie nicht nur theoretisch befürworten, sondern auch im realen Leben praktisch unterstützen. Solange ohne viel nachzudenken auf die Angebote US-amerikanischer Quasi-Monopolisten zurückgegriffen wird, sind die Chancen, europäische (Wert-)Vorstellungen zur Geltung zu bringen, vom guten Willen der US-Anbieter abhängig, selbst wenn die EU zunehmend versucht, europäisches Recht dem entgegenzustellen. Auch in der digitalisierten Welt sind Monopole gefährlich.[75] Um die Orientierung der staatlichen Ordnung auf das Gemeinwohl zu wahren und unsere sozialen Werte zu verteidigen, wäre es oft sinnvoll, europäische Anbieter zu wählen, die den rechtlichen Vorgaben Europas uneingeschränkt unterworfen sind. Es geht also nicht darum, von Staats wegen Konkurrenzangebote zu den US-Tech-Giganten aufzubauen. Staatliche Stellen sollten aber genau prüfen, ob bestimmte Aufgaben nicht auch von einem europäischen Anbieter angemessen erfüllt werden könnten.[76]

Leider steht dem bislang unsere Bequemlichkeit entgegen. Selbst die Regierungen Europas sind in dieser Hinsicht vor Fehlern nicht gefeit. Ein

75 Ramge, Thomas: Mensch und Maschine. Wie Künstliche Intelligenz und Roboter unser Leben verändern, Ditzingen: Reclam 2018, S. 87f.

76 So dürfte das europäische Übersetzungsprogramm DeepL mit den meisten anderen einschlägigen Angeboten gut mithalten können.

gutes Beispiel ist die Corona-Warn-App, die ursprünglich als europäische Entwicklung geplant war. Nachdem Akzeptanz-Probleme auftraten, wurde in Deutschland eine enge Anbindung an die US-Tech-Riesen Apple und Google vollzogen, wobei als Argument auch ein besserer Datenschutz (!) genannt wurde. Die europäische Lösung wurde fallengelassen, und auch die Möglichkeit betriebssystemunabhängiger Lösungen nicht mit dem nötigen Nachdruck verfolgt. Die von der Regierung enorm gehypte Corona-Warn-App hat sich inzwischen als stumpfes Schwert im Kampf gegen die Pandemie erwiesen. Nicht nur ältere Smartphones konnten sie zunächst nicht nutzen, auch auf neueren Modellen des chinesischen Konkurrenzanbieters Huawei war die App nicht einsetzbar. Hier wird deutlich, wie die unkritische Orientierung an den marktbeherrschenden Unternehmen dazu führt, deren Standards und damit deren Dominanz weiter zu stärken.

Besonders problematisch ist, dass in der digitalisierten Welt immer mehr traditionell staatliche Aufgaben an private Anbieter übertragen werden, etwa die Bereitstellung und Sicherung der Kommunikationsinfrastruktur (wie E-Mail), Zahlungsdienste und digitale Währungen, die Sicherung von Dokumenten in Clouds, sogar hoheitliche Aufgaben, wie die Sicherung von (digitalen) Identitäten usw. Viele Menschen achten nur auf das Funktionieren der Angebote und übersehen, dass der Staat grundsätzlich als Anbieter in anderer Weise gebunden ist als Private: Der Staat besitzt einen Versorgungsauftrag, er ist für die Daseinsvorsorge zuständig und unterliegt demokratischer Kontrolle und Steuerung. Bei privaten Mega-Unternehmen, noch dazu solchen, die aus dem Ausland agieren, ist dies nicht so. Zwar lassen sich manche staatliche Bindungen auf Private übertragen, doch sind diese stets fragil und müssen oft erst durchgesetzt werden. Deshalb ist es umso wichtiger, dafür zu sorgen, dass der Staat in der digitalisierten Welt nicht noch weiter geschwächt wird.[77]

77 Grundlegend hierzu Schallbruch, Martin: Schwacher Staat im Netz. Wie die Digitalisierung den Staat in Frage stellt, der überzeugend herausarbeitet, wie staatliche Instanzen (gerade in Deutschland) oft selbst daran mitwirken, ihren Einfluss abzubauen; überaus kritisch gegenüber den US-Tech-Giganten auch ebd. P. Nemitz/M. Pfeffer: Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, S. 301ff. und passim.

IV Zusammenfassung

Die in den Jahren 2019 und 2020 von der HLEG AI vorgestellten Leitlinien für eine vertrauenswürdige KI gehen einen Mittelweg zwischen einem bloßen akademischen Ethik-Kodex und nur partikular geltenden betrieblichen Compliance-Regeln. Vielmehr wurde versucht, auf der Grundlage der Menschenwürdegarantie und der europäischen Menschenrechtskataloge ethische Leitlinien zu formulieren, die hinreichend konkret und praktikabel sind, um auch in Unternehmen tatsächlich angewendet zu werden. Auf diese Weise ist es gelungen, akzeptanzfähige Grundlagen für konkrete Regulierungsmaßnahmen in Europa zu formulieren. Die ethische und rechtliche Einhegung der KI muss allerdings fortlaufend neuen technischen Entwicklungen angepasst werden, um den Vorrang des Gemeinwohls gegen kommerzielle oder hegemoniale Interessen zu verteidigen. Die Leitlinien der EU zur Regulierung von KI sollte man deshalb nicht als den Abschluss von Regulierungsüberlegungen verstehen, sondern als ihren Ausgangspunkt.