Martin Eifert

Staatliche Verantwortung für KI-Infrastrukturen und Datensicherheit*

Vortrag im Rahmen der 63. Bitburger Gespräche
Mainz, 09./10.01.2020

I. Künstliche Intelligenz als transformative Herausforderung

1. Künstliche Intelligenz: Von der Projektion zum Gesellschaftsexperiment

Künstliche Intelligenz (KI) ist spätestens seit den 1960er Jahren gesellschaftlich breiter bekannt. In *Stanley Kubricks* epochalem Film „2001 – Odyssee im Weltraum" sind Erwartungen und Enttäuschungen über den Supercomputer HAL 9000 zentral bei der Reise der Menschheit auf den Jupiter. In der Informatik warnte der Pionier *Joseph Weizenbaum* etwas später vor „Programmen ohne Autor".[1] Und in der Rechtswissenschaft wurden schon in den ersten Ausarbeitungen über die Technisierung der Verwaltung auch „lernende Maschinen" angesprochen.[2] Immer war die künstliche Intelligenz hierbei eine Projektion in die Zukunft, die schaudern oder hoffen machte.[3] Vor diesem Hintergrund erschienen die damals konkret anstehenden Probleme der Technisierung zugleich relativ übersichtlich und leichter lösbar.

Jetzt stellen wir fest: Wir sind in dieser Zukunft von gestern angekommen. Weitreichende Anwendungsszenarien einer starken künstlichen Intelligenz, also einer menschlichen Fähigkeiten

---

* Um Fußnoten ergänzter und leicht erweiterter Vortrag. Für ihre wertvolle Unterstützung bedanke ich mich ganz herzlich bei Ass. iur. *Philipp Breuling* und cand. iur. *Julian Siefert*.

[1] *Weizenbaum*, Alptraum Computer, Die Zeit v. 21.1.1972, S. 43 ff., der die „Systeme ohne Autoren" als Alptraum beschreibt und selbst von der Beschwörung der letzten Autorität des „Es steht geschrieben" seines Vaters abgrenzt: "Aber da…konnte ich mir einen Autor vorstellen, konnte seine Wertmaßstäbe rekonstruieren und schließlich zustimmen oder ablehnen".

[2] Siehe nur *Bull*, Verwaltung durch Maschinen, 2. Aufl. 1964, S. 6, Rn. 2.2143; vgl. auch *Luhmann*, Recht und Automation in der öffentlichen Verwaltung, 1966, S. 68 („Lernmaschinen"); zuvor bereits in diese Richtung *Zeidler*, Über die Technisierung der Verwaltung, 1959, S. 30 ff. („Regierungsmaschine").

[3] Neben den vorgenannten Beiträgen auch *Steinmüller*, Informationstechnologie und Gesellschaft, 1993, S. 542.

vergleichbaren oder ihnen überlegenen Intelligenz, sind zwar auch heute noch Projektionen.[4] Sie speisen sich aber nicht allein aus künstlerischer Vorstellung oder wissenschaftlichen Visionen. Sie sind vielmehr durchaus realistische Fortschreibungen[5] der schon verfügbaren „schwachen" künstlichen Intelligenz, also intelligenter Systeme, die auf konkrete Anwendungsprobleme begrenzt sind. Massenhaft verfügbare Daten, die enorm gestiegenen Datenübertragungsraten und Rechenleistungen sowie die Kombinationen verschiedener Ansätze lernender Systeme haben die Entwicklung und Verbreitung immer intelligenterer, also komplexerer und lernfähigerer, Algorithmen ermöglicht und rasant beschleunigt.[6] Diese schwache KI durchdringt schon jetzt die Industrien und unser aller digitale Lebenswelten.[7] Und Prototypen wie Pflegeroboter und autonome Fahrzeuge verweisen sehr konkret auf die bevorstehende Verbreitung und Fortentwicklung in allen gesellschaftlichen Zusammenhängen.[8]

Dies macht den jahrzehntelang aufgeschobenen aktiven Umgang mit den Chancen und Risiken von KI nicht nur zwingend, sondern hebt ihn auch aus einer rein theoretischen Bearbeitung heraus. Die bereits begonnene Entwicklung erhält den Charakter eines gesellschaftlichen Selbstexperiments, in dem wir unter Realbedingungen und in Echtzeit unsere eigenen weitreichenden Transformationen begleiten.

In einem gewissen Umfang ist dies zwar der Normalzustand moderner, innovationsoffener Gesellschaften.[9] Dieses Experiment weist aber doch eine Kombination von Besonderheiten auf: Zu nennen sind vor allem hohe Dezentralität und wachsende Allgegenwärtigkeit, die hohe Entwicklungsgeschwindigkeit, ein Mix aus stufenlosem Einschleichen in bislang menschlich gesteuerte Prozesse und einer disruptiven Kraft für ganze Geschäftsmodelle sowie schließlich die Vielzahl beteiligter Akteure und ihre wechselseitigen Abhängigkeiten und Vernetzungen. Es ist also ein besonders umfassendes, schnell verlaufendes und vielfältiges Experiment.

2. Unsicherheit und Besorgnispotentiale

Die große Bandbreite der Zukunftsszenarien von utopischen bis dystopischen Verläufen und die große Zahl der Kommissionen – von der Datenethikkommission[10] über die Enquete-Kommission[11]

---

[4] Vgl. den Bericht des *National Science and Technology Council*, Preparing for the Future of Artificial Intelligence v. Oktober 2016, S. 7, abrufbar unter https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf <19.6.2020>; *Wischmeyer*, AöR 143 (2018), 1 (3 f., 15).

[5] Ähnliche Einschätzung bei *Gaede*, Künstliche Intelligenz – Rechte und Strafen für Roboter?, 2019, S. 24 ff. m. w. N.; siehe auch die Übersicht zu verschiedenen Expertenmeinungen in *Bostrom*, Superintelligence, 2014, S. 18 ff.; bereits frühzeitig in eine solche Richtung als ausdrückliche lineare Verlängerung damaliger Trends *Steinmüller* (Fn. 3), S. 323 ff., 326.

[6] Zur technologischen Entwicklung von KI und deren Voraussetzungen vgl. den Bericht des *National Science and Technology Council* (Fn. 4), S. 5 f.; *Buxmann/H. Schmidt*, in: dies. (Hrsg.), Künstliche Intelligenz, 2019, S. 3 ff.; *Wischmeyer* (Fn. 4), S. 9 f.

[7] Siehe nur die Analyse der *Fraunhofer-Allianz Big Data und Künstliche Intelligenz*, Zukunftsmarkt Künstliche Intelligenz – Potentiale und Anwendungen, 2017, abrufbar unter https://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/KI-Potenzialanalyse_2017.pdf <19.6.2020>; zum breiten Anwendungsbereich von schwacher KI vgl. auch *Wischmeyer* (Fn. 4), S. 2 ff.

[8] Siehe nur *Fraunhofer-Allianz Big Date und Künstliche Intelligenz* (Fn. 7), S. 17 ff. (zu Sozialen Robotern) und S. 21 ff. (zu autonomen Transportmitteln).

[9] Vgl. *Krohn/Weyer*, Soziale Welt 40 (1989), 349 ff., die früh den Begriff der „Realexperimente" und der „experimentellen Wissensgesellschaft" prägen; s. a. *Krohn*, in: Rammert/Bechmann (Hrsg.), Technik und Gesellschaft, Jahrbuch 9, 1997, S. 65 (71) zur modernen Gesellschaft als „Gesellschaft der Selbstexperimentation"; sowie aus neuester Zeit *Böschen/Groß/Krohn*, in: dies. (Hrsg.), Experimentelle Gesellschaft: Das Experiment als wissensgesellschaftliches Dispositiv, 2017, S. 7 ff.

[10] Siehe https://www.bmi.bund.de/DE/themen/it-und-digitalpolitik/datenethikkommission/datenethikkommission-node.html <19.6.2020>.

[11] Siehe Enquete-Kommission des Deutschen Bundestags „Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale",

bis zur Kommission Wettbewerbsrecht 4.0[12] – verdeutlichen gleichermaßen das Bewusstsein um die Bedeutung von KI sowie die Unsicherheit über die richtige Strategie im Umgang mit dieser.

Dabei werden die weitreichenden Möglichkeiten der KI immer deutlicher.[13] Die Besorgnispotentiale aber auch. Im Zentrum der Besorgnisse stehen die Intransparenz der Entscheidungsfindung[14] und die beispiellose Konsequenz der Zweckverfolgung[15]. Beide weichen deutlich von menschlichen Interaktionen ab. Die Intransparenz führt zu Unsicherheiten und Kontrollverlusten über den Entscheidungsvorgang und löst Sorgen um die Qualitätssicherung und mögliche versteckte Diskriminierungen aus.[16] Die beispiellose Konsequenz der Zweckverfolgung vergrößert zunächst das Schädigungspotential bei Fehlern. Sie versieht aber auch selbst sozial erwünschtes Verhalten (wie die Rechtsdurchsetzung) oder sozial toleriertes Verhalten mit einer neuen bedrohlichen Komponente. Wir wollen ja in Wirklichkeit gar nicht, dass z. B. jede Geschwindigkeitsüberschreitung beim Autofahren immer sanktioniert wird, und wir tolerieren zwar punktuelle Schlitzohrigkeit, wollen aber keine konsequente und systematische Manipulation.

3. Grundrechte als Rahmen und Gestaltungsauftrag

Dass unser begonnenes Gesellschaftsexperiment angesichts der bestehenden Risiken ein kontrolliertes Experiment bleibt, dafür sind Grundrechte und Rechtsetzer verantwortlich – auf nationaler wie europäischer Ebene. Die Grundrechte bilden den bewährten Rahmen gesellschaftlicher Entwicklung. Sie sichern die Freiheit als Innovationsgaranten und weisen mit dem Schutzauftrag für die Rechte und Rechtsgüter aller Mitglieder der Gesellschaft zugleich dem Staat eine Verantwortung für die regulatorische Begleitung der Innovationen zu.[17] Das verfassungsrechtliche Schutzversprechen ist dabei nur am Ziel eines im Ergebnis hinreichenden Schutzes orientiert.[18] Es ist vor allem Auftrag an den Gesetzgeber, die rechtlichen Instrumente zu entwickeln, mit denen dieser Schutz sichergestellt werden kann, ohne dabei die Freiheit unverhältnismäßig zu verkürzen.[19] Dem Gesetzgeber steht hierbei ein Gestaltungsspielraum[20] zu,

---

https://www.bundestag.de/ausschuesse/weitere_gremien/enquete_ki <19.6.2020>.

12 Siehe https://www.bmwi.de/Redaktion/DE/Artikel/Wirtschaft/kommission-wettbewerbsrecht-4-0.html <19.6.2020>.

13 Vgl. nur den Beitrag von *Achatz* in diesem Band.

14 Siehe nur *Martini*, Blackbox Algorithmus, 2019, S. 28 ff.; *Wischmeyer* (Fn. 4), S. 8, 42 ff.; *Kreutzer/Sirrenberg*, Künstliche Intelligenz verstehen, 2019, S. 12 f.

15 Hierzu *Rademacher*, AöR 142 (2017), 366 (397 f.); vgl. auch *Hoffmann-Riem*, AöR 142 (2017), 1 (33 ff.); *Wischmeyer* (Fn. 4), S. 42.

16 Zum Diskriminierungspotenzial von selbstlernenden Algorithmen siehe *Martini* (Fn. 14), S. 47 ff.; vgl. auch *West/Whittaker/Crawford*, Discriminating Systems: Gender, Race, and Power in AI, AI NOW Institute, 2019, abrufbar unter https://ainowinstitute.org/discriminatingsystems.html <19.6.2020>.

17 Vgl. insgesamt näher *Hoffmann-Riem*, Innovation und Recht – Recht und Innovation, 2016, S. 28 ff. und die Beiträge in *Eifert/Hoffmann-Riem*, Innovationsverantwortung, 2009.

18 Vgl. nur BVerfGE 92, 26 (46); aus der Literatur statt aller *Isensee*, in: ders./Kirchhof (Hrsg.), HbStR IX, 2011, § 191 Rn. 293 ff.; zur besonderen strukturellen Offenheit der aus den grundrechtlichen Schutzpflichten folgenden Gebote – insb. mit Blick auf die regelmäßige Verfügbarkeit verschiedener (geeigneter) Schutzhandlungen und der in der Abwägung zu berücksichtigenden verschiedenen Effektivitätsgrade – siehe *Alexy*, Theorie der Grundrechte, 8. Aufl. 2018, S. 420 ff.

19 Vgl. nur die knappe Zusammenfassung bei *Dreier*, in: Dreier-GG, 3. Aufl. 2013, Vorb. Rn. 101 ff. und die ausführlichere Darstellung bei *Calliess*, in: Merten/Papier (Hrsg.), HGR II, 2006, § 44.

20 Siehe nur BVerfGE 125, 39 (78) (st. Rspr.); aus der Literatur nur die Rekonstruktion bei *Alexy*, VVDStRL 61 (2001), 7 (14 f.), der zeigt, dass der Gesetzgeber zwischen der abwehrrechtlichen Eingriffsgrenze einerseits und der Schutzpflicht andererseits i. d. R. strukturelle und epistemische Spielräume für Gestaltung hat und dass eine daraus folgende Rahmenordnungsidee die Idee einer verfassungsrechtlichen Grundordnung nicht ausschließt.

der sich bei einer Unsicherheit über die zukünftigen Entwicklungen und Gefährdungen – wie sie in vielen Bereichen der KI-Entwicklung besteht – nochmals vergrößert[21].

Wir müssen die Entwicklungen deshalb sicherlich im Lichte der zu schützenden verfassungsrechtlichen Werte diskutieren. Die informationstechnische Entwicklung wurde ja auch schon immer unter starker Bezugnahme auf die Menschenwürde, die Diskriminierungsverbote und den Persönlichkeitsschutz diskutiert.[22] Wir sollten aber zurückhaltend sein, aus diesen Verfassungsnormen unmittelbare rechtliche Gestaltungsvorgaben abzuleiten. Die staatliche Innovationsverantwortung ist vor allem politische Gestaltungsaufgabe.

Die Politik beginnt hierbei freilich nicht bei Null. Weil der Umgang mit Innovationen grundsätzlich ein Normalfall moderner Gesellschaften ist, haben der deutsche wie europäische Gesetzgeber bereits eine Vielzahl an Instrumenten für die Umhegung von Innovationsrisiken entwickelt. Ich möchte mit Ihnen deshalb zunächst einen Blick in diesen rechtlichen Instrumentenkasten werfen. Es wird dabei deutlich werden, wie gut er grundsätzlich bestückt ist und wo für die Herausforderungen der KI besondere Schwerpunkte der Aufmerksamkeit liegen sollten oder doch noch Entwicklungsbedarfe bestehen (II.).

Anschließend möchte ich den Blick aber auf eine Besonderheit der digitalen Transformation lenken. Sie kann nicht nur zentrale ethische und verfassungsrechtlich geschützte Werte gefährden, sondern fordert auch dazu heraus, diese in vielen Fällen neu zu durchdenken und gegebenenfalls neu zu justieren. Das Oszillieren der Prognosen zwischen Dystopien und Utopien ist auch Ausdruck dieser tiefergehenden Unsicherheit (III.).

II. Die rechtliche Toolbox des Technik- und Risikorechts – Was passt wie gut?

Lassen Sie uns den Blick in den Instrumentenkasten zunächst bei den Sicherungen einer Innovationsoffenheit des Rechts beginnen.

1. Innovationsoffenheit: Erhöhung regulatorischer Flexibilität durch kontrollierte Experimente

Staat und Verwaltung haben auf den gesellschaftlichen Innovationsdruck mit eigenen Experimentierdesigns reagiert.

Seit den 1970er Jahren sind Experimentalgesetze ein bekanntes Mittel der Innovationsbegleitung; sie wurden nicht nur in gesellschaftspolitischen Reformfeldern wie Schule und Hochschule, sondern z. B. im Medienbereich gerade zur Verarbeitung des technologischen Wandels eingesetzt[23] – auch wenn aus heutiger Perspektive Bildschirmtextversuche und Kabelpilotprojekte nur noch Relikte vergangener Zeiten sind und mehr mit Sentimentalität als mit Zukunftsoffenheit verbunden werden.

---

[21] Zur Einschätzungsprärogative des Gesetzgebers bei Entscheidungen unter ungesicherter Tatsachengrundlage und Prognoseentscheidungen siehe statt vieler *Ossenbühl*, in: Festg. BVerfG, 1976, S. 458 (501 ff., insb. 502 [„Prognosekontrolle nach Maßgabe der Rationalität"] u. 504 ff.).

[22] Siehe nur die frühe Diskussion über die Zulässigkeit von automatisierten Verwaltungsentscheidungen bei *Bull* (Fn. 2), S. 92 ff. (zu Art. 1 und Art. 2 GG), S. 108 ff. (zu Art. 3 GG).

[23] Siehe die Übersicht zu den Anwendungsfeldern experimenteller Gesetzgebung bei *Hoffmann-Riem*, in: FS Thieme, 1993, S. 55 (57 ff.); zu Experimentalgesetzen s. a. *Horn*, Experimentelle Gesetzgebung unter dem Grundgesetz, 1989; *Maaß*, Experimentierklauseln für die Verwaltung und ihre verfassungsrechtlichen Grenzen, 2001.

Rechtlich nutzen diese Gesetze politische Entscheidungsspielräume innerhalb der für Experimentiergesetze intensiv diskutierten, aber letztlich unverhandelbaren verfassungsrechtlichen Grenzen.[24] Auch wenn diese Experimentalgesetze nicht nur unschuldig Erfahrungen sammeln sollten, sondern zugleich Instrumente der Innovationsdurchsetzung bildeten, waren sie im Grunddesign auf eine begleitende wissenschaftliche Forschung hin angelegt, deren Ergebnisse in die Politik und Rechtsetzung zurückgespielt werden sollten.[25]

Staatlich initiierte, wissenschaftlich durchgeführte oder von unabhängigen Beobachtern begleitete Experimente sind auch im Bereich der Digitalisierung und des Einsatzes von KI wichtig und werden staatlich eingerichtet.[26] Hauptanwendungsfälle hierfür dürften komplexe soziale Zusammenhänge mit einer Vielzahl von Akteuren sein. Derartige Beispiele sind die Experimentierräume für digitale und agile Arbeitsformen im Geschäftsbereich des Bundesministeriums für Arbeit und Soziales[27] oder die Reallabor-Strategie des Bundesministeriums für Wirtschaft und Energie u. a. für digitale und KI-Fragen der Mobilitätsgestaltung, der Energiewende oder der Gesundheitsversorgung[28].

Die in diesen Experimentaldesigns vorgesehenen zeitlich gestreckten Verschleifungen von Politik und Wissenschaft werden der Entwicklungsgeschwindigkeit und der Heterogenität der Anwendungen auf den Dienstleistungs- und Produktmärkten mit KI aber regelmäßig nicht gerecht. Hier bedarf es eher der Synchronisation von technologischer Entwicklung und staatlicher Aufsicht – gegebenenfalls mit zusätzlicher wissenschaftlicher Begleitung. Als Beispiel für einen darauf eingestellten Rahmen können die sogenannten Sandboxes[29] genannt werden, die insbesondere in einer Reihe europäischer Staaten im Finanzdienstleitungssektor[30], aber auch im Datenschutzrecht[31] eingeführt wurden.

In diesen Sandboxes können etwa innovative Finanzdienstleister neue Produkte und Geschäftsmodelle – z. B. eine algorithmenbasierte Anlageberatung (sog. Robo-Advice)[32] – unter Realbedingungen und enger Begleitung der zuständigen Behörden testen. Im Grundmodell[33]

---

[24] Vgl. *Hoffmann-Riem* (Fn. 23), S. 63 ff.; *Maaß* (Fn. 23), S. 70 ff. und *Horn* (Fn. 23), S. 233 ff.

[25] Zu diesen verschiedenen vorder- und hintergründigen Funktionen der Experimentalgesetze auch *Hoffmann-Riem* (Fn. 23), S. 56 f., 61 ff.

[26] Vgl. überblicksweise *Hill,* DÖV 2018, 497 (498) und *ders.*, DÖV 2016, 493 ff.

[27] Vgl. https://www.experimentierraeume.de/start/ <19.6.2020>.

[28] Vgl. https://www.bmwi.de/Redaktion/DE/Dossier/reallabore-testraeume-fuer-innovation-und-regulierung.html <19.6.2020>.

[29] Die Europäische Bankenaufsichtsbehörde (EBA) definiert „regulatory sandboxes" als „regulatorische Erprobungszonen", die „Finanzinstituten und Nichtfinanzunternehmen einen kontrollierten Raum [bieten], in dem sie innovative FinTech-Lösungen mit Unterstützung einer Behörde für begrenzte Zeit testen können, sodass sie ihre Geschäftsmodelle in einem sicheren Umfeld validieren und erproben können", siehe das Diskussionspapier der *EBA*, On the EBA's approach to financial technology (FinTech) v. 4. April 2017, EBA/DP/2017/02, S. 7 Fn. 8, abrufbar unter https://eba.europa.eu/sites/default/documents/files/documents/10180/1919160/7a1b9cda-10ad-4315-91ce-d798230ebd84/EBA%20Discussion%20Paper%20on%20Fintech%20%28EBA-DP-2017-02%29.pdf <19.6.2020>. Vgl. auch die ausführliche Beschreibung dieser Regulierungsstrategie in *Basel Committee on Banking Supervision*, Sound Practices – Implications of fintech developments for banks and bank supervisors v. Februar 2018, S. 41, abrufbar unter https://www.bis.org/bcbs/publ/d431.pdf <19.6.2020>.

[30] Für eine Übersicht zu den das Sandbox-Verfahren bereits einsetzenden Staaten siehe den Bericht des *Joint Committee of European Supervisory Authorities*, Fintech: Regulatory sandboxes and innovation hubs, JC 2018/74, S. 16 f., abrufbar unter https://esas-joint-committee.europa.eu/Publications/Reports/JC%202018%2074%20Joint%20Report%20on%20Regulatory%20Sandboxes%20and%20Innovation%20Hubs.pdf <19.6.2020>; siehe auch *Krimphove/Rohwetter*, BKR 2018, 494 (495); *Zetzsche/Barberis/Buckley/Arner*, Fordham Journal of Corporate & Financial Law Vol. 23 (2017), 31 (64 ff.).

[31] Zum datenschutzrechtlichen Sandbox-Verfahren des britischen Information Commissioner's Office siehe https://ico.org.uk/for-organisations/the-guide-to-the-sandbox-beta-phase/ <19.6.2020>; zum Sandbox-Verfahren nach dem indischen Entwurf eines Datenschutzgesetzes aus 2019 siehe *team Trilegal*, Cri 2020, 1 (6).

[32] Zu Robo-Advice und deren aufsichtsrechtlicher Einordnung siehe *Baumanns*, BKR 2016, 366 ff.; vgl. auch die Verbraucherinformation des BaFin zu Robo-Advice unter https://www.bafin.de/DE/Verbraucher/Finanzwissen/Fintech/RoboAdvice/robo_advice_node.html <19.6.2020>.

[33] Siehe die vergleichende Analyse der von verschiedenen EU-Mitgliedstaaten in der Finanzmarktregulierung eingesetzten Sandbox-Verfahren im Bericht des *Joint Committee of European Supervisory Authorities*, Fintech: Regulatory sandboxes and innovation hubs (Fn. 30), S. 18 ff.,; vgl. auch die Beschreibung des britischen

bewerben sich Anbieter mit einem Produkt oder Geschäftsmodell unter Darlegung des innovativen Charakters und der damit verbundenen Risiken bei der Aufsichtsbehörde. Bei Zulassung zur Sandbox vereinbaren sie mit dieser das Testverfahren, die Parameter und eine eventuelle Exit-Strategie und beginnen dann mit dem zeitlich befristeten Test, der über Berichtspflichten und eine intensivierte Aufsicht abgesichert ist. Nach einer Evaluation wird schließlich über das weitere Vorgehen entschieden. Zentral ist, dass in diesen Experimenten zwar allgemeine regulatorische Anforderungen wie Zulassungsvorbehalte abgesenkt werden können, aber regelmäßig alle Kundenschutzvorschriften einzuhalten sind und ggf. Sicherungen für den Kundenschutz wie Warnhinweise oder Beschränkungen bestehen. Auch erfolgt der Test oft nicht im allgemeinen Markt, sondern mit akkreditierten Kunden.[34]

Ausweislich eines Berichts des Zusammenschlusses der europäischen Aufsichtsbehörden sehen die meisten Behörden in diesen Verfahren die Gelegenheit, ein besseres Verständnis der Innovationen und eingesetzten Technologien zu erhalten.[35]

Sandboxes dürften darüber hinaus wichtige politische und wirtschaftliche Funktionen haben. Politisch könnten sie das drohende Aufeinanderprallen disruptiver Geschäftsmodelle und überkommener Wirtschafts- und Aufsichtsstrukturen in einen kooperativen Entwicklungsprozess überführen und wirtschaftlich erleichtern sie kleineren Unternehmen den Markteinstieg.[36] Es erscheint sinnvoll, über ihren verstärkten bereichsspezifischen Einsatz für KI in regulierten Branchen nachzudenken und sicherzustellen, dass das verbesserte Wissen der Regulierer auch in die Rechtsetzungsprozesse eingespeist wird.

Darüber hinaus wäre zu überlegen, ob nicht auch eine öffentliche Testinfrastruktur für zentrale allgemeine Aspekte der KI eingerichtet werden könnte. Hierüber könnte ebenfalls der trade-off zwischen regulatorischen Anforderungen einerseits und leichtem Markteinstieg auch für kleinere Unternehmen andererseits abgemildert werden.

Sandboxes und erst recht eine öffentliche Testinfrastruktur benötigen allerdings zusätzliche öffentliche Ressourcen. Ohne diese dürfte eine angemessene Begleitung der digitalen Transformation aber ohnehin nicht zu leisten sein.

2. Risikosteuerung: Risikoorientierte Kontrolle und beobachtende Begleitung

Staat und Verwaltung haben aber nicht nur Experimentierdesigns, sondern auch zahlreiche rechtliche Instrumente zur dauerhaften Wahrnehmung der staatlichen Innovationsverantwortung entwickelt.

Sandbox-Verfahrens im Bericht der britischen Finanzmarktaufsichtsbehörde *FCA,* Regulatory sandbox v. November 2015, S. 7 ff., abrufbar unter https://www.fca.org.uk/publication/research/regulatory-sandbox.pdf <19.6.2020> und in *FCA*, Regulatory sandbox lessons learned report v. Oktober 2017, S. 4, abrufbar unter https://www.fca.org.uk/
publication/research-and-data/regulatory-sandbox-lessons-learned-report.pdf <19.6.2020>; dazu auch *Krimphove/*
Rohwetter (Fn. 30), S. 495 f.; *D.-F. Lange*, in: FS Schwintowski, 2017, S. 331 (336 ff.); *H. J. Allen*, The George Washington Law Review Vol. 87 (2019), 579 (596 ff.).

[34] Zu diesen verschiedenen Ansätzen des Kundenschutzes siehe den Bericht der *FCA*, Regulatory sandbox (Fn. 33), S. 9 f.

[35] Vgl. Bericht des *Joint Committee of European Supervisory Authorities*, Fintech: Regulatory sandboxes and innovation hubs (Fn. 30), S. 33; zu den Bedenken einiger Aufsichtsbehörden ebendort, S. 35 f. Zur entsprechenden Zielsetzung dieser Ansätze die Mitteilung der *Europäischen Kommission*, FinTech-Aktionsplan: Für einen wettbewerbsfähigeren und innovativeren EU-Finanzsektor, COM (2018) 109 final v. 8.3.2018, S. 9 ff., abrufbar unter https: //ec.europa.eu/transparency/regdoc/rep/1/2018/DE/COM-2018-109-F1-DE-MAIN-PART-1.PDF<19.6.2020>.

[36] So auch die Einschätzung im Bericht des *Joint Committee of European Supervisory Authorities*, Fintech: Regulatory sandboxes and innovation hubs (Fn. 30), S. 33 f.; zu den Vorteilen der Sandbox-Verfahren auch *Krimphove/Rohwetter* (Fn. 30), S. 496 f.

Die zentrale Schmiede für die rechtlichen Instrumente bildete das Technik- und Umweltrecht. Dort wurden die Grenzen des traditionellen Ordnungsrechts als erstes deutlich[37] und wurden in der Folge auch zentrale Herausforderungen bearbeitet, die mit jeder technologischen Innovation verbunden sind: Dies sind namentlich nicht oder kaum abschätzbare Risiken[38] und ein Wissensvorsprung der privaten Innovatoren gegenüber der staatlichen Regulierung sowohl hinsichtlich der Risikoermittlung als auch hinsichtlich effizienter risikominimierender Maßnahmen[39]. Die Ergebnisse der Anstrengungen im Umwelt- und Technikrecht sind nach Risiken differenzierte Anzeige- und Zulassungsregime[40] sowie vielfältige kooperative Mechanismen der Risikoermittlung, der Standardsetzung und der Kontrolle[41]. Die Kontrollregime umfassen u. a. Direktübertragungen von Daten[42] und Auditierungen durch akkreditierte Stellen[43]. Hinzu treten Transparenzpflichten[44], Informationszugangsrechte für Dritte[45] und Verbandsklagen[46]. Und im Zivilrecht haben sich hier die Gefährdungshaftung und die Haftung für Produktrisiken entwickelt.[47]

Neben dem allgemeinen Umwelt- und Technikrecht steht für die Informationstechnik das Datenschutzrecht. Es ist den spezifisch informationstechnischen Gefährdungen mit dem begrenzten Fokus auf personenbezogene Daten frühzeitig entgegengetreten.[48] Dabei entwickelte es im Wechselspiel mit der allgemeinen technikrechtlichen Diskussion[49] immer vielfältigere Ansätze. Allgemeine in ihren je konkreten Ausprägungen teilweise „risikobasierte“, also auf die Höhe des jeweiligen Risikos abgestimmte, Anforderungen an die Datenverarbeitung[50]

---

[37] Vgl. nur *Winter*, Das Vollzugsdefizit im Wasserrecht, 1975; *Bohne*, Der informale Rechtsstaat, 1981, S. 22 ff.; *Wahl/Appel*, in: Wahl (Hrsg.), Prävention und Vorsorge, 1995, S. 1 (29 ff.); siehe auch die frühzeitige Forderung einer dogmatischen Fortentwicklung des Gefahren- und Risikobegriffs im Umweltrecht unter Integration der Aufgabe der Wissensgenerierung bei *Ladeur*, Das Umweltrecht der Wissensgesellschaft – Von der Gefahrenabwehr zum Risikomanagement, 1995, S. 69 ff.

[38] Hierzu grundlegend *Ladeur* (Fn. 37), S. 9 ff.

[39] Vgl. bereits *Bohne* (Fn. 37), S. 77 ff.; näher zur Dezentralisierung moderner Wissensbestände in komplexen, ausdifferenzierten Gesellschaftsordnungen *Wollenschläger*, Wissensgenerierung im Verfahren, 2009, S. 34 ff., 40 ff.; dazu auch *Vesting*, in: Hoffmann-Riem/Schmidt-Aßmann/Voßkuhle (Hrsg.), GVwR II, 2. Aufl. 2012, § 20 Rn. 50.

[40] Siehe Überblick bei *Eifert*, in: Schoch (Hrsg.), Besonderes Verwaltungsrecht, 2018, Kap. 5 Rn. 84 ff.; *Kahl/Gärditz*, Umweltrecht, 11. Aufl. 2019, S. 89 ff.; ausführlicher *Rehbinder*, in: ders./Schink (Hrsg.), Grundzüge des Umweltrechts, 5. Aufl. 2018, S. 145 ff.

[41] Zu den verschiedenen Kooperationsformen im Umweltrecht siehe *Shirvani*, Das Kooperationsprinzip im deutschen und europäischen Umweltrecht, 2005, S. 132 ff.; zu den kooperativen Mechanismen der Kontrolle siehe Überblick bei *Eifert* (Fn. 40), Rn. 112 ff. und breiter *ders.*, in: Hoffmann-Riem/Schmidt-Aßmann/Voßkuhle (Hrsg.), GVwR I, 2. Aufl. 2012, § 19.

[42] So kann die Behörde im Rahmen des § 31 Abs. 5 S. 2 BImSchG auch die computergerechte und fortlaufende Übermittlung der Messergebnisse an die Behörde anordnen (BVerwG, NVwZ 1997, 998 f. zu § 31 S. 2 BImSchG a.F.).

[43] Zum freiwilligen Umweltaudit-System nach der Verordnung (EG) Nr. 1221/2009 (EMAS-VO), dem Audit-System der ISO 14001 und dem für große Unternehmen verpflichtenden Energieaudit nach EDL-G siehe *Eifert* (Fn. 40), Rn. 127 ff.; *Kahl/Gärditz* (Fn. 40), S. 166 ff.; vgl. auch die zwingende Zertifizierung von Erstbehandlungsanlagen für elektronische Altgeräte gem. § 21 ElektroG.

[44] Siehe nur §§ 289b, 289c Abs. 2 Nr. 1 HGB (Lagebericht von großen börsennotierten Kapitalgesellschaften muss Erklärung zu Umweltbelangen enthalten; in Umsetzung der RL 2013/34/EU [CSR-Richtlinie]); zu Kennzeichnungspflichten und der freiwilligen Nutzung von Umweltzeichen vgl. *Eifert* (Fn. 40), Rn. 153.

[45] Siehe das UIG des Bundes, die Umweltinformationsgesetzte der Länder und die speziellen Unterrichtungspflichten in §§ 46a und 47d Abs. 3 BImSchG, § 28a GenTG; zum Zugang zu Umweltinformationen als umweltrechtliches Mittel und Ausgangspunkt für gesellschaftliches Engagement vgl. statt vieler *Klein*, Umweltinformationen im Völker- und Europarecht: Aktive Umweltaufklärung des Staates und Informationszugangsrechte des Bürgers, 2011.

[46] Siehe das UmwRG (in Umsetzung der RL 2003/35/EG [Öffentlichkeitsbeteiligungs-RL] und der Aarhus-Konvention); vgl. hierzu den instruktiven Überblick bei *Schlacke*, NVwZ 2017, 905 ff.

[47] Siehe nur §§ 32 ff. GenTG; § 89 WHG und die Regelungen des UmweltHG.

[48] So war das 1. Hessische Datenschutzgesetz vom 30.9.1970 das erste Datenschutzgesetz weltweit. Auf Bundesebene wurde das 1. BDSG am 1.2.1977 verkündet. Zur Geschichte der Datenschutzgesetzgebung siehe *Simitis/Hornung/Spiecker gen. Döhmann*, in: dies. (Hrsg.), Datenschutzrecht, 2019, Einl. Rn. 1 ff.

[49] Zur jüngeren geschichtlichen Entwicklung des Technikrechts unter Einschluss des Datenschutzrechts siehe *Vec*, Kurze Geschichte des Technikrechts, in: Schulte/Schröder (Hrsg.), Handbuch des Technikrechts, 2. Aufl. 2011, S. 3 (79 ff.).

[50] So hängen bestimmte Anforderungen der Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung [DS-GVO]) von der Bewertung des Risikos ab, welches von der Datenverarbeitung für die Grundrechte der betroffenen Personen ausgeht (siehe insb. ErwG 74 ff., 83 ff., 89 ff.; Art. 24 Abs. 1, 25 Abs. 1, 27 Abs. 2 lit. a, 30 Abs. 5,

einschließlich der Pflicht zur Dokumentation ihrer Einhaltung[51] wurden gepaart mit einem starken Individualschutz der Betroffenen u. a. über Auskunfts- und Löschungsansprüche;[52] und neben die allgemeinen Regelungen traten spezifische Anforderungen für besondere Gefährdungslagen[53]. Diese beinhalten auch erste, wenngleich noch (zu) eng gefasste Regelungen über automatisierte Entscheidungen[54] und besondere Anforderungen für den Umgang mit diskriminierungsanfälligen sensiblen Daten[55].

Dieser Instrumentenkasten als gesellschaftlicher Erfahrungsschatz im Umgang mit technologischen Innovationen wird nun – wenig verwunderlich – auch für den Umgang mit KI und Datensicherheit eingesetzt. Das IT-Sicherheitsrecht des Bundes ist im Wesentlichen eine Anwendung der umwelt- und technikrechtlichen Regulierungsstrategien auf IT-Risiken in kritischen Bereichen.[56] Es strukturiert das Monitoring über Sicherheitsrisiken und -vorkehrungen und sichert den systematischen Wissensaufbau über Informations-, Melde- und Unterrichtungspflichten mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI) als zentralem Knotenpunkt.[57] Es schreibt insbesondere für die erfassten kritischen Infrastrukturen die Einhaltung des Standes der Technik vor[58] und eröffnet eine differenzierte Konkretisierung des Standards über ein kooperatives Vorgehen von Branchenverbänden und BSI[59]. Und schließlich setzt es für die Einhaltung und Kontrolle der Standards in erheblichem Maße auf Zertifizierungen[60] und Auditierungen[61].

Ganz parallel greifen jetzt auch die Vorschläge zur Regulierung der Algorithmen regelmäßig explizit (z. B. Martini)[62] oder implizit (z. B. Datenethikkommission)[63] auf diesen Instrumentenkasten

---

32 Abs. 1 u. 2, 33 Abs. 1, 34 Abs. 1, 35 Abs. 1, 36 Abs. 1 und 37 Abs. 1 DS-GVO); vgl. auch *Albrecht*, CR 2016, 88 (93 f.); für eine frühe Systematisierung der risikobasierten Ansätze in den Entwürfen des EU-Parlaments und des Rats zur DS-GVO siehe *Veil*, ZD 2015, 347 ff.; kritisch *Gellert*, International Data Privacy Law, 2015 5/1, S. 3 (6 ff., 12 ff.).

[51] Vgl. die Vorschriften zu den erforderlichen Nachweisen in Art. 5 Abs. 2, 24 Abs. 1 S. 1 u Abs. 3, 25 Abs. 3, 32 Abs. 3 DS-GVO; siehe auch Art. 30 Abs. 1 lit. g und die Dokumentationspflicht in Art. 33 Abs. 5 DS-GVO.

[52] Siehe Art. 13 f. DS-GVO i. V. m. §§ 32 f. BDSG (Informationspflichten); Art. 15 DS-GVO i. V. m. § 34 BDSG (Auskunftsanspruch); Art. 16 DS-GVO (Berichtigungsanspruch); Art. 17 DS-GVO i. V. m. § 35 BDSG (Löschungsanspruch).

[53] Siehe z. B. die besonderen Anforderungen an die Rechtmäßigkeit der Verarbeitung besonderer Kategorien personenbezogener Daten nach Art. 9 DS-GVO (i. V. m. §§ 22 ff. BDSG) und von Daten über strafrechtliche Verurteilungen nach Art. 10 DS-GVO, die besondere Regelung zur Videoüberwachung öffentlich zugänglicher Räume nach § 4 BDSG und zum Schutz des Wirtschafsverkehrs bei Scoring und Bonitätsauskünften nach § 31 BDSG.

[54] Siehe Art. 22 DS-GVO i. V. m. § 37 BDSG.

[55] Siehe Art. 9 Abs. 1 u 2 DS-GVO i. V. m. §§ 22 ff. BDSG.

[56] Vgl. dazu, dass sich das IT-Sicherheitsrecht neben klassisch ordnungsverwaltungsrechtlichen insbesondere auch Regulierungsansätzen aus dem Infrastruktur- und Technikrecht bedient, W*ischmeyer*, Die Verwaltung 50 (2017), 155 (157, 161, 169 ff.); s. a. monographisch zu den Pflichten des IT-Sicherheitsrechts *Freimuth*, Die Gewährleistung der IT-Sicherheit Kritischer Infrastrukturen, 2018.

[57] Siehe §§ 3 Abs. 1 S. 2 Nr. 2, 8a Abs. 3 S. 3 f., 8b, 8c Abs. 4 S. 1 Nr. 1 BSIG und insb. die Meldepflichten in §§ 8b Abs. 4, 8c Abs. 3 BSIG, § 44b AtG, § 11 Abs. 1c EnWG, § 109 Abs. 5 TKG; für einen Überblick über die bestehenden Meldepflichten siehe *Hornung*, NJW 2015, 3334 (3337); *Schneider*, Meldepflichten im IT-Sicherheitsrecht, 2017, S. 131 ff.; *Freimuth* (Fn. 56), S. 313 ff., 348 ff., 359 ff. Die Regelung von Meldepflichten für kritische Infrastrukturen und für bestimmte Anbieter digitaler Dienste wird auch von der Richtlinie (EU) 2016/1148 (NIS-Richtlinie) eingefordert, s. Art. 14 Abs. 3 bzw. Art. 16 Abs. 3 NIS-RL.

[58] Siehe § 8a Abs. 1 S. 2 BSIG; vgl. auch § 109 Abs. 1 S. 2, Abs. 2 S. 3 TKG. Auch insoweit besteht inzwischen mit Art. 14 Abs. 1 u 2 bzw. Art. 16 Abs. 1 u 2 NIS-RL eine europarechtliche Pflicht, nationale Regelungen zur Einhaltung bestimmter Mindestsicherheitsanforderungen durch die Infrastrukturbetreiber zu treffen.

[59] Siehe § 8a Abs. 2 BSIG und die abweichenden Vorgaben in § 109 TKG, § 11 Abs. 1a u. 1b EnWG; die bloß lückenhaft normierte prozedurale Einrahmung der kooperativen Standardsetzung kritisieren *Hornung* (Fn. 57), S. 3336; *Wischmeyer* (Fn. 56), S. 169; ausführlich hierzu *Freimuth* (Fn. 56), S. 279 ff.

[60] Siehe §§ 9, 2 Abs. 7 BSIG und § 8a Abs. 3 S. 2, Abs. 5 BSIG; dazu *Wischmeyer* (Fn. 56), S. 169 f. Die am 28.6.2019 in Kraft getretene VO (EU) 2019/881 (sog. Cybersecurity Act) führt einen Europäischen Zertifizierungsrahmen für Cybersicherheit ein, in dem nun europäische Zertifizierungsschemata für IKT-Produkte, -Dienste und -Prozesse ausgearbeitet und mittels Durchführungsakt der Europäischen Kommission festgelegt werden können, s. Art. 46 ff. VO (EU) 2019/881.

[61] Siehe §§ 8a Abs. 3 S. 2, Abs. 5 BSIG; vgl. dazu auch *Freimuth* (Fn. 56), S. 303 ff.; kritisch zur Unbestimmtheit der gesetzlichen Regelungen über erforderliche Nachweise *Roßnagel*, DVBl 2015, 1206 (1209).

[62] *Martini* (Fn. 14), S. 114 ff. bzw. 158 ff.

[63] Gutachten der Datenethikkommission, 2019, S. 173 ff., abrufbar unter https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-

zurück. Besonders anschaulich wird dies im Vorschlag der Datenethikkommission der Bunderegierung für eine „Kritikalitätspyramide“.[64] Mit steigendem Schädigungspotential soll die Anwendung algorithmischer Systeme steigenden Anforderungen unterworfen werden. Bei Anwendungen mit allenfalls geringem Schädigungspotential sind keine besonderen Anforderungen vorgesehen. Bei einem gewissen Schädigungspotential werden Anforderungen wie Transparenzpflichten, die Veröffentlichung von Risikofolgenabschätzungen sowie Kontrollverfahren wie Auditierungen oder Offenlegungen gegenüber Aufsichtsbehörden ausgelöst. Mit weiter steigendem Schädigungspotenzial werden dann auch Zulassungsverfahren und eine Live-Schnittstelle zu Aufsichtsbehörden gefordert, bis schließlich unvertretbare Anwendungen verboten werden.

Insgesamt liegen die vorgeschlagenen Regulierungsansätze zur KI damit im Wesentlichen auf dem Pfad des bewährten risikorechtlichen- und datenschutzrechtlichen Instrumentariums.

3. Unterbeleuchtete Faktoren effektiver Risikosteuerung

Die Tatsache, dass dort Instrumente bereit liegen, heißt jedoch noch lange nicht, dass sie auch wirksam sind. Wir können schnell der Illusion erliegen, die Probleme der KI mit der Einführung bekannter Instrumente des Risikorechts gelöst zu haben, wenn wir nicht genau hinschauen, ob die Instrumente denn in ihrem ursprünglichen Bereich überhaupt effektiv waren und ob im Bereich der KI überhaupt Rahmenbedingungen bestehen, die das jeweilige Instrument auch hier wirksam werden lassen.

Zwei mir besonders wichtig erscheinende Problemfelder möchte ich hier exemplarisch ansprechen.

a) Breitere Anwendung von Beobachtungspflichten

Die diskutierte Risikosteuerung setzt meist beim Betreiber der KI-Systeme an. Dies liegt nahe, denn die Betroffenen werden ja auch durch dessen KI-Verwendung einem Risiko ausgesetzt. Die Risikosteuerung folgt hier dem Muster der Regulierung gefährlicher Anlagen.[65] Der Betreiber wird aber oft die KI nur teilweise oder gar nicht selbst entwickelt haben, sondern quasi blind als Modul in seine Anwendungen einbauen. Die Leistungsfähigkeit der lernenden Systeme besteht wohl oft gerade aus der Kombination verschiedener Bausteine, so dass Probleme, die in Software-Module eingeschrieben sind, zwar auch durch den Verwender zu verantworten sind, aber nicht auf dessen Verwendung beschränkt bleiben. Die multifunktionale Verwendung, die dezentrale Entwicklung und die Kombination von Modulen mit vielstufigen Wertschöpfungsketten bei vielen KI-Systemen legen auch Beobachtungspflichten für Softwareentwickler entsprechend den Modellen des Produktrechts nahe. Im Chemikalienrecht gibt es etwa auch Informationspflichten entlang der Wertschöpfungsketten[66] und in vielen Bereichen des Produktrechts müssen die Hersteller auch nach der Markteinführung beobachten, ob sich später bisher unbekannte Risiken realisieren oder neue Verwendungen neue Risiken hervorrufen[67]. Hier sollte die Risikosteuerung für die KI stärker Parallelen auch zum Produktrecht einbeziehen.

---

datenethikkommission.pdf?__blob=publicationFile&v=6 <19.6.2020>.

[64] Siehe hierzu näher Gutachten der Datenethikkommission (Fn. 63), S. 177 ff.

[65] Vgl. *Martini* (Fn. 14), S. 158 ff., 225 ff.

[66] Vgl. Art. 1 Abs. 3 S. 1, Art. 31 ff. VO (EG) Nr. 1907/2006 (REACH-VO); dazu *Führ/Bizer*, in: Eifert/Hoffmann-Riem (Hrsg.), Innovationsverantwortung, 2009, S. 303 ff. und *Kloepfer*, Umweltrecht, 4. Aufl. 2016, § 19 Rn. 49 f., insb. Rn. 143 ff. zum Informationsmanagement in der REACH-VO und im ChemG.

[67] Dazu näher und systematisierend *Eifert*, in: ders. (Hrsg.), Produktbeobachtung durch Private, 2015, S. 9 ff. sowie die jeweils fachgebietsspezifischen Beiträge ebenda; als ein gesetzliches Beispiel siehe § 6 Abs. 3 ProdSG.

b) Bereichsspezifische Operationalisierung der Maßstäbe und Koordination der Beteiligten

Insgesamt hält das Risikorecht auf der gesetzlichen Ebene zunächst formale Instrumente bereit, die in den konkreten inhaltlichen Maßstäben regelmäßig sehr unspezifisch sind. Dies gilt im Umweltrecht für zentrale Begriffe wie „schädliche Umwelteinwirkungen“ oder „erhebliche Nachteile“[68] ebenso wie im Datenschutzrecht für die durchlaufenden allgemeinen Abwägungsklauseln, nach denen zu prüfen ist, ob schutzwürdige Interessen oder Grundrechte der Betroffenen überwiegen[69] oder angemessen gewahrt werden[70]. Eine Operationalisierung dieser Maßstäbe oder auch gehaltvolle Standardsetzungen benötigen Konkretisierungen. Dies wird im Auftrag zur Förderung von konkretisierenden Verhaltensregeln gem. Art. 40 DS-GVO und den zahllosen untergesetzlichen Standardsetzungen des Umweltrechts[71] aufgegriffen. Solche Konkretisierungen müssen sich auf Branchen und Anwendungskontexte beziehen, wie etwa die konkreten Emissionsstandards im Umweltrecht[72], oder müssen als Differenzierungsansatz Fallgruppen für typisierte Anwendungen entwickeln, wie etwa hinsichtlich einer Datenverarbeitung für Werbezwecke[73]. Bei der Regulierung von KI-Anwendungen und den dort ebenfalls breit geforderten Standardsetzungen und Selbstregulierungen[74] werden solche Konkretisierungen und Differenzierungen dann gleichermaßen notwendig.

Während die differenzierten Standardisierungen im Umweltrecht unter Beteiligung von Branchen- und Berufsverbänden sowie Normungsorganisationen in vielen Bereichen gut gelingen,[75] lässt sich für das Datenschutzrecht (etwa hinsichtlich konkretisierender Verhaltensregeln) bislang keine vergleichbare Erfolgsgeschichte ausmachen[76]. Für den Bereich algorithmischer Systeme und KI ist eine von selbst erfolgende, erfolgversprechende Selbstorganisation nun mindestens so unwahrscheinlich wie im Datenschutz. Dies liegt an einer Reihe von Faktoren. Die Welt der Software-Programmierung scheint von einer weniger institutionell orientierten Kultur geprägt, viele Anwendungen entstehen gerade bei Neueinsteigern mit weniger etablierten Vernetzungen, das ökonomische Leitbild ist eher an Disruption als an institutioneller Verfestigung ausgerichtet und die Akteurslandschaft ist von vornherein hochinternational.[77] Es dürfte zentral für eine gelingende

---

[68] Vgl. nur § 5 Abs. 1 Nr. 1 u. Nr. 2 i. V. m. § 3 Abs. 1 BImSchG.

[69] Vgl. nur Art. 6 Abs. 1 S. 1 lit. f, Art. 21 Abs. 1 S. 2, Art. 49 Abs. 1 S. 2 DS-GVO; vgl. ähnlich konkretisierungsbedürftig auch die Parameter für eine wirksame Einwilligung in Art. 7 DS-GVO.

[70] Vgl. nur Art. 22 Abs. 3 u. 4, Art. 32 Abs. 1 u. 2 DS-GVO.

[71] Siehe grundlegend dazu *Marburger*, Die Regeln der Technik im Recht, 1979; aktuell zusammenfassend *Kloepfer/Durner*, Umweltschutzrecht, 3. Aufl. 2020, § 2 Rn. 23 ff.

[72] Vgl. v. a. die verschiedenen Bundesimmissionsschutz-Verordnungen, z. B. 39. BImSchV (Verordnung über Luftqualitätsstandards und Emissionshöchstmengen vom 2.8.2010, zuletzt geändert am 18.7.2018), die häufig der Umsetzung europäischer Richtlinien dienen, z. B. RL 2008/50/EG (Luftqualität und saubere Luft für Europa vom 11.06.2008) und RL 2001/81/EG (nationale Emissionshöchstmengen für bestimmte Luftschadstoffe vom 27.11.2001).

[73] Vgl. etwa die Fallgruppenbildung bei *Schulz*, in: Gola (Hrsg.), DS-GVO, 2. Aufl. 2018, Art. 6 Rn. 68 ff. (Werbescoring, Bestandskundenwerbung, Werbung für fremde Angebote, [Adresshandel,] Online-Marketing) und bei *Ehmann*, in: Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.), Datenschutzrecht, 2019, Anh. 3 zu Art. 6 DS-GVO, Rn. 38 ff. (Bestandskundenwerbung, Empfehlungswerbung, nutzungsbasierte Online-Werbung, Lettershop-Verfahren, Werbung mittels des Facebook-Modells).

[74] Vgl. nur Gutachten der Datenethikkommission (Fn. 63), S. 76; *Martini* (Fn. 14), S. 358 ff.

[75] Vgl. *Eifert* (Fn. 40), Kap. 5 Rn. 83 und oben Fn. 72.

[76] In mehr als 20 Jahren sind in Deutschland lediglich zwei anerkannte Verfahrensregeln entstanden: Der Datenschutzkodex des Gesamtverbands der Deutschen Versicherungswirtschaft vom 7.9.2012 (zusammen mit Beitrittsliste der Versicherungsunternehmen abrufbar unter https://www.gdv.de/de/ueber-uns/unsere-services/daten-schutz-ko-dex---code-of-conduct---15544 <19.6.2020>) und der „Geo-Business Code of Conduct“ des Selbstregulierung Informationswirtschaft (SRIW) e.V. und der Kommission für Geoinformationswirtschaft (GIW) vom 13.1.2015 (abrufbar unter https://www.bmwi.de/Redaktion/DE/Publikationen/Geobusiness/publikation-code-of-conduct.html <19.6.2020>); vgl. *Roßnagel*, in: Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.), Datenschutzrecht, 2019, Art. 40 DS-GVO Rn. 4, und *Reifert*, ZD 2019, 305 ff., die sich von Art. 40 DS-GVO eine Verstärkung dieses Instruments der „regulierten Selbstregulierung“ erhofft.

[77] Vgl. die kurze Marktbeschreibung bei *Hill* (Fn. 26), S. 500; s. a. *Downes/Nunes*, Harvard Business Manager 6/2013, 64 ff.

Operationalisierung der vorgeschlagenen Instrumente des Risikorechts im Bereich der KI sein, die Bereitschaft zur Selbstorganisation anzureizen und die Fähigkeiten dazu zu stärken. Hierauf sollte ein besonderes Augenmerk liegen.

III. Jenseits der Toolbox: Diskussion über Werte und Justierungen ermöglichen

Bis hierher haben wir KI als neue Technologie betrachtet, deren Chancenverwirklichung eine Flexibilisierung des Rechts erfordert[78] und deren Risikoumhegung auf den Erfahrungsschatz des Umgangs mit anderen riskanten Technologien zurückgreifen kann. Das wird ihr jedoch nur teilweise gerecht. Denn in vielen Bereichen, in denen KI zu erheblichen Veränderungen führt, sind wir unsicher über die Schutzbedürftigkeit und das gewünschte Schutzniveau bestehender Werte oder müssen überhaupt erst herausfinden, welche Werte wir schützen wollen. Dies ist etwa ein Unterschied zu Innovationen, die „nur" gesundheitliche Risiken mit sich bringen und bei denen wir eine relativ klare Vorstellung davon haben, was die Gesundheit ist, die wir vor den neuen Risiken schützen wollen (nachfolgend 1.). Wegen der Unsicherheit über die Schutzgüter ist bei KI-Anwendungen übrigens auch die risikotechnologische Einordnung hinsichtlich ihres Schädigungspotentials oft sehr schwierig. Um diese Unsicherheiten abbauen zu können, muss eine angemessene Diskussion über die Werte und gewünschten Schutzniveaus gesichert werden (nachfolgend 2.).

1. Unklarheiten über Schutzbedarfe und Schutzniveaus

Drei Schlaglichter sollen zunächst die Unsicherheit etwas ausleuchten.

a) Wert menschlicher Interaktion

KI drängt menschliche Interaktion und Kommunikation zurück. Der menschliche Kontakt war über lange Zeit aber so selbstverständlich, dass seine Vor- und Nachteile für die unterschiedlichen Interaktionen nur begrenzt reflektiert wurden[79] und entsprechend die Konsequenzen eines Verlusts auch nur begrenzt absehbar sind. Sehr markant ist dies etwa im Bereich des Legal Tech.[80] Die Bewältigung von Rechtstreitigkeiten ist einerseits in vielen Aspekten Routinegeschäft, das in seinen gleichförmigen Abläufen schon äußerlich stark an Automatisierung erinnert.[81] Andererseits sind Rechtstreitigkeiten bekanntlich so vielfältig wie das Leben und oft hoch emotionalisiert. Die bestehenden Verfahren sind über die menschliche Abwicklung mit jederzeitigen und vielfältigen Möglichkeiten der Kontextualisierung von Sachverhalten, der feinsinnigen Unterscheidung von Konstellationen, der Fortentwicklung rechtlicher Maßstäbe und mit Resonanzräumen für Empathie

[78] Vgl. auch die allgemeinen Befunde zu Flexibilisierungsmöglichkeiten im geltenden Recht, insb. Experimentierklauseln, Ermessens- und Abwägungsregeln bei *Hill* (Fn. 26), S. 497, 501 ff.

[79] Eine hohe Sensibilität für die Bedeutung der kommunikativen Dimension im Technisierungsprozess aber frühzeitig bei *Garstka*, JbRsoz 7 (1980), 233 (236 ff.).

[80] Vgl. einführenden Überblick bei *Buchholtz,* JuS 2017, 955 ff.

[81] Vgl. *Engel*, JZ 2014, 1096 (1097 f.) mit der Einschätzung, dass eine mechanische Rechtsprüfung zwar nicht alle, aber doch manche juristischen Fragen befriedigend beantworten könne; kritisch hierzu *Kotsoglou*, JZ 2014, 451 ff.

und Akzeptanzsicherung ausgestattet.[82] Wir müssen uns für die Nutzung von KI darüber klar werden, in welchen Fällen wir diese selbstverständlichen Möglichkeiten erhalten wollen, in welchen Fällen wir Funktionen technisch nachbilden können[83] und wollen und in welchen Fällen diese Aspekte verzichtbar oder sogar unerwünscht sind und die Bewältigung von Rechtstreitigkeiten nur unnötig verschlechtern sowie verteuern.

b) Grenzen zulässiger Manipulation

KI perfektioniert auch die Anwendung verhaltensökonomischer Erkenntnisse. E-Commerce Angebote und Videoplattformen zeigen die Richtung ebenso eindrücklich an wie Computerspiele. Dies eröffnet qualitativ neue Möglichkeiten der Manipulation.[84] Unser verfassungsrechtlich bis auf die Menschenwürde rückführbares normatives Leitbild ist demgegenüber Autonomie.[85] Manipulation steht sicher in einer unfreundlichen Spannung zur Autonomie. Es ist aber offenkundig, dass die Autonomie nur Leitbild ist, also ein Orientierung gebendes Ideal, nicht ein feststehender absoluter Wert. Wir leben ganz selbstverständlich mit unseren vielfältigen täglichen Beeinflussungen und Manipulationen. So lassen wir es zu, dass die billigen Waren im Supermarkt ganz unten im Regal stehen und eröffnen unsere Mitarbeitergespräche erst einmal mit einem Lob, weil dies die konstruktive Atmosphäre für die anschließende Kritik schafft. Wir leben mit diesen kleinen Manipulationen unter anderem, weil wir die Erfahrung gemacht haben, dass sie jedenfalls keinen großen Schaden anrichten und letztlich auch durchschaubar sind. Die bestehenden rechtlichen Grenzen des Manipulationsschutzes beruhen maßgeblich auf solchen Erfahrungsschätzen und auf den immanenten Grenzen der Manipulationsmechanismen. Die systematische, perfektionierte Manipulation schätzen wir offenkundig anders ein. Sie führt zu jener Besorgnis des Ausgeliefertseins, die auch viele Dystopien antreibt.[86] Wo aber angesichts dieser Möglichkeiten die Grenze zwischen zulässiger Nutzung von Verhaltensmechanismen und notwendigem Schutz bewusster, eigenverantwortlicher Entscheidungen verlaufen soll, wissen wir noch nicht. Die Diskussion um das „Nudging“, das eine systematische Nutzung verhaltensökonomisch optimierter Anreizstrukturen regulatorisch einsetzen möchte, hat gleichermaßen im Autonomiedenken wurzelnde Störgefühle wie bestehende Unsicherheiten über die rechtliche Verarbeitung solcher Ansätze offen gelegt.[87] Bei KI-Nutzungen werden diese Fragen nochmals erheblich verschärft.[88]

c) Art und Qualität von erforderlichen Begründungen

KI arbeitet damit, Muster zu erkennen. Entscheidungen werden vor dem Hintergrund von Ähnlichkeiten der behandelten Phänomene getroffen, nicht als logische Ableitung aus

---

[82] Vgl. *von Graevenitz*, ZRP 2018, 238 (240 f.) zur besonderen Eignung menschlicher Entscheider hinsichtlich komplexer Wertungsentscheidungen mit Offenheit für Erkenntnisse aus Nachbardisziplinen und einem Blick für längerfristige Entscheidungswirkungen.

[83] Zur (eingeschränkten) technischen Machbarkeit der Verwendung von KI bei der Automatisierung komplexer gerichtlicher Entscheidungen, siehe z. B. *Dreyer/Schmees*, CR 2019, 758 ff.

[84] Zum Manipulationspotential im Zivilrechtsverkehr siehe *Wagner/Eidenmüller*, ZfPW 2019, 220 ff.; die Beeinflussung von Wahlen durch sog. Behavioral Microtargeting thematisiert *Hill*, in: ders./Kugelmann/Martini (Hrsg.), Digitalisierung in Recht, Politik und Verwaltung, 2018, S. 47 ff.; vgl. allgemein zur „digitalen Technosteuerung von Verhalten“ *Hoffmann-Riem* (Fn. 15), S. 11 ff.

[85] Statt vieler *Dreier*, in: Dreier-GG, Band I, 3. Aufl. 2013, Vorb. Rn. 6 und Art. 1 I Rn. 42 und die Beiträge in *Bumke/Röthel* (Hrsg.), Autonomie im Recht, 2017.

[86] Kritisch zu Manipulationen durch Big Data-Applikationen statt vieler *Ebers*, MMR 2018, 423 (424, 428); *Mik*, Law, Innovation and Technology 8 (2016), 1 (12 ff.).

[87] Siehe nur zum Konzept *Thaler/Sunstein*, Nudge, 2008; *Sunstein*, Why Nudge?, 2014; und die konträren Positionen mit Blick auf die Autonomie von *Rebentato*, Taking Liberties, 2012 und *Conly*, Against Autonomy, 2013. Zur Diskussion in Deutschland nur die Beiträge in Kemmerer/Möllers/Steinbeis (Hrsg.), Choice Architecture in Democracies, 2017; *Eifert*, TUP 2015, 178 ff.

[88] Siehe nur *M. Hildebrandt*, Smart Technologies and the End(s) of Law, 2015, S. 80 ff., 226; *Stalder*, Kultur der Digitalität, 2016, S. 199 ff., 218 ff.

vordefinierten Entscheidungsprämissen, die auf einen Sachverhalt angewendet werden. Den Entscheidungen zu Grunde liegen also vor allem Korrelationen, nicht bestehende oder angenommene Kausalitäten.[89] Unabhängig vom bekannten Problem der Intransparenz[90] verändert dies elementar die Begründungsstruktur.[91] Es lässt sich zwar die Qualität der Entscheidungen im Sinne einer Zielgenauigkeit der eingesetzten KI ermitteln.[92] Es lassen sich in Zukunft vielleicht auch statistische Faktoren und ihre Gewichtungen bestimmen.[93] Aber gegenwärtig kann man nur die statistische Bestimmung der Zielgenauigkeit des Gesamtoutputs gegenüber den Vorgaben und allenfalls einen Ausschluss evident sachwidrigen Inputs in der Trainingsphase der KI sicher bestimmen. Die verfügbaren Informationen weichen damit deutlich von den überkommenen sozialen Mustern der Begründungen ab, an welche auch rechtliche Begründungszusammenhänge anknüpfen.

Welche Anforderungen genau an Begründungen zu stellen sind, ist nicht ganz klar. Auch hier haben wir es mit selbstverständlichen und überkommenen Elementen zu tun, die nicht voll reflektiert sind. Tim Miller hat allerdings jüngst in einer Meta-Studie unter Auswertung von 250 Publikationen aus den Gesellschafts- und Geisteswissenschaften zentrale Elemente von „Erklärungen" identifiziert, die einen ersten Eindruck vermitteln, auch wenn „Begründungen" sicher nicht einfach mit „Erklärungen" gleichgesetzt werden können, sondern Elemente sowohl von „Argumenten" als auch von „Erklärungen" beinhalten dürften[94]. Nach Miller[95] sind Erklärungen „contrastive", stellen also den Kontext zu ausdrücklich anders gelagerten Fällen her, sind „selected", indem sie auf die zentralen Ursachen abstellen, überzeugen in Wahrscheinlichkeitsaussagen nur, wenn sie durch kausale Erklärungen begleitet werden und sind schließlich sozial, müssen also kommunikativ oder interaktiv vermittelt werden. Und Douglas Walton[96] arbeitet den Streit um die beste Erklärung als Wettbewerb konkurrierender Geschichten heraus, die jeweils aus einer raumzeitlichen Abfolge von Ereignissen und Handlungen bestehen und damit ebenfalls Kausalverläufe zentral stellen. Rechtliche Begründungserfordernisse, wie sie vor allem für staatliche Entscheidungen und hier beispielhaft in § 39 VwVfG bestehen, sollen vor allem über die Herstellung der Entscheidung informieren, durch deren Erklärung Akzeptanz schaffen und die Aussichten für mögliche Rechtsbehelfe abschätzbar machen.[97] Sie zielen damit auf Erklärung (Akzeptanz) und (kommunikative) Bestreitbarkeit (insbes. Rechtsschutz und seine Befriedungsfunktion[98]). Dies legt es nahe, eine Begründungsstruktur zu fordern, die an die voranstehenden Grundmuster von Erklärungen anknüpft.

---

[89] Zur Funktionsweise von lernenden Algorithmen, insb. zur Musterbildung auf Grundlage aufgefundener Korrelationen, siehe nur *Wischmeyer* (Fn. 4), S. 11 ff. m. w. N.

[90] Siehe die Nachweise oben, bei Fn. 14.

[91] Dieser Gesichtspunkt wird häufig nicht klar genug vom allgemeinen Transparenzproblem getrennt; undeutlich insoweit *Wischmeyer* (Fn. 4), S. 42 ff., der aber in Fn. 48 die Frage aufwirft, ob man auf lange Sicht in größerem Umfang statistische Korrelationen als Erklärungen akzeptieren werde und in Fn. 90 für den strafrechtlichen Schuldvorwurf eine auf bloße Korrelationen gestützte Tatsachenfeststellung aus prinzipiellen Gründen ablehnt; vgl. auch *Rademacher* (Fn. 15), S. 390, der ein „Recht auf Plausibilität" andenkt; und *Burrell*, Big Data & Society 3:1 (2016), 1 (2 f., 10), die auf einen „mismatch between mathematical procedures and machine learning algorithms and human styles of semantic intepretation" hinweist; zu den möglichen Auswirkungen der neuen Begründungsstrukturen von Big Data auf die Voraussetzungen personaler Identitätsbildung *M. Hildebrandt*, Philosophy & Technology 24 (2011), 371 ff.

[92] Siehe den Überblick zu möglichen Qualitätssicherungsmaßnahmen bei *Martini/Nink*, NVwZ-Extra 10/2017, 1 (12 f.); zum sog. black-box-testing siehe nur *Busch*, Algorithmic Accountability, ABIDA Gutachten 2018, S. 65 f., insb. Fn. 287 m. w. N.; *Zweig*, Wo Maschinen irren können, 2018, S. 30 f.; zum Einsatz von Kontrollalgorithmen *Martini*, JZ 2017, 1017 (1022).

[93] Zu den technischen Bemühungen um sog. „Explainable AI" siehe auch *Martini* (Fn. 14), S. 193 ff.; *Wischmeyer* (Fn. 4), S. 61 f. jeweils m. w. N.

[94] Vgl. zu dieser unscharfen Unterscheidung näher *Walton*, Argument Structure: A Pragmatic Theory, 2019, S. 26 ff., 42 ff.

[95] *T. Miller*, Artificial Intelligence 267 (2019), 1 ff.

[96] *Walton*, Argument Evaluation and Evidence, 2016, S. 64 ff.

[97] Vgl. statt aller *Stelkens*, in: Stelkens/Bonk/Sachs (Hrsg.), VwVfG Kommentar, 9. Aufl. 2018, § 39 Rn. 1; daneben dienen die Begründungen selbstverständlich auch der Selbstkontrolle (der Verwaltung).

[98] Grundlegend *Luhmann*, Legitimation durch Verfahren, 1983, S. 129 ff.

Wenn beim KI-Einsatz Kausalitäten durch Korrelationen ersetzt werden und die Entscheidungslogik nicht mehr als Abfolge aufeinander aufbauender Argumentationsschritte dargestellt werden kann, entfallen zentrale Merkmale überkommener Begründungen und die typischen Ansatzpunkte für den Streit um deren Tragfähigkeit.[99] Die Kausalität ist eine selbstverständliche Zentralkategorie des Rechts und wenn wir über Begründungen streiten, erfolgt dies regelmäßig, indem wir unterstellte Kausalverläufe bestreiten oder bei Prognosen die Annahmen über zukünftige Kausalverläufe in Frage stellen. Diese Begründungsmuster sind nicht beliebig austauschbar, denn sie spiegeln geteilte Erfahrungen. Wir müssen deshalb beim Einsatz von KI auch entscheiden, was auf welche Weise begründet werden können muss und wo uns welche Begründungsstrukturen wichtig sind.

Viel spricht dafür, im staatlichen Bereich grundsätzlich, zumindest aber bei bedeutsamen Fragen auf den traditionellen Begründungsstrukturen zu beharren. Zugleich scheinen wir in einzelnen Bereichen problemlos davon abzusehen, wenn wir etwa an KI-basierte Verkehrsleitsysteme[100] denken, die ja auch zu Geschwindigkeitsvorgaben für Einzelne führen können. Bei geringen individuellen Freiheitsbeeinträchtigungen, die überdies Tätigkeiten mit hoher wechselseitiger Abhängigkeit betreffen und bei denen deshalb der Einsatz von KI diese Beeinträchtigungen überkompensieren könnte, wird diese Abweichung offenbar normativ hingenommen. Deutlich wird damit jedenfalls die Notwendigkeit, über die Begründungsstrukturen selbst im staatlichen Bereich neu nachzudenken.

d) Grenzen gruppenbasierter Zuschreibungen

Im Bereich der Diskriminierung schützen wir rechtlich die Individualität der Einzelnen, indem wir diese vor Zuschreibungen schützen, die aus ihrer bloßen Zugehörigkeit zu einer Gruppe folgen.[101] Dabei konzentrieren wir uns auf die erfahrungsgemäß besonders wirkmächtigen Stereotype, etwa ethnische Zugehörigkeit oder Geschlecht.[102] KI arbeitet mit Mustererkennung und deshalb prinzipiell auch mit Gruppenbildung.[103] Auch hier muss angesichts der Perfektionierung von Zuschreibungstechniken nicht nur gefragt werden, wie wir den bestehenden Diskriminierungsschutz absichern[104], sondern auch, ob und wenn ja welche gruppenbasierten Zuschreibungen wir vielleicht jenseits dessen neu in den Schutz einbeziehen wollen.[105]

2. Sicherungen angemessener Wertediskussionen

---

[99] Die Bedeutung von Bestreitbarkeit gegenüber IT-basierten Entscheidungen im Rahmen des Rechtsstaatsprinzips betont auch *M. Hildebrandt* (Fn. 88), S. 100 ff.; als grundlegendes System-Feature fordern es *Mulligan/ Kluttz/Kohli*, in: Werbach (Hrsg.), After the Digital Tornado: Networks, Algorithms, Humanity, 2020 (im Erscheinen).

[100] Zu solchen Projekten nur *Djeffal*, DVBl. 2017, 808 (809 f.); https://www.bast.de/BASt_2017/DE/Projekte/fp-laufend-v5.html?nn=1819560; https://www.bast.de/BASt_2017/DE/Verkehrstechnik/Fachthemen/v5-verkehrsbeein flussungsanlagen.html <19.6.2020>.

[101] Vgl. nur *Britz*, Einzelfallgerechtigkeit versus Generalisierung, 2008, S. 2 f.

[102] Siehe die in § 1 AGG aufgelisteten verpönten Differenzierungsmerkmale (Rasse, ethnische Herkunft, Geschlecht, Religion oder Weltanschauung, Behinderung, Alter, sexuelle Identität); siehe auch Art. 3 Abs. 3 GG (Geschlecht, Abstammung, Rasse, Sprache, Heimat und Herkunft, Glauben, religiöse oder politische Anschauung, Behinderung).

[103] Zum Diskriminierungspotential intelligenter Systeme siehe *Martini* (Fn. 14), S. 47 ff.; *Wischmeyer* (Fn. 4), S. 26 ff. jeweils m. w. N.; zur Diskriminierung von Frauen im Zuge des Einsatzes eines Algorithmus durch den Arbeitsmarktservice Österreich siehe auch *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, Verfassungsblog v. 26.10.2018, abrufbar unter https://verfassungsblog.de/koennen-algorithmen-diskriminieren/ <19.6.2020>.

[104] Dazu siehe nur *Martini/Nink* (Fn. 92), S. 10 ff.

[105] Siehe näher Gutachten der Datenethikkommission (Fn. 63), S. 28, 43, 167 ff.

Wenn wir also in vielen Fällen unsicher sind, welche Werte wir mit welcher Intensität wo schützen sollen, dann müssen wir sicherstellen, dass die notwendigen Diskussionen von Werten und Verhandlungen von Interessen angemessen geführt werden können. Die dafür notwendige allgemeine Grundsicherung ist eine freie individuelle und öffentliche Meinungsbildung. Völlig zu Recht sind wir deshalb über deren gegenwärtigen Gefährdungen besonders alarmiert und fordern auch für den Einsatz von KI in diesem Bereich besondere Anforderungen[106]. Hier sollen aber drei andere Punkte herausgestellt werden.

a) Erfahrungsgrundlagen schaffen

Eine angemessene gesellschaftliche Diskussion benötigt nicht nur Informationen, sondern auch Erfahrungsgrundlagen. Die gegenwärtig diskutierten Transparenzanforderungen an den Einsatz von Algorithmen und KI stehen immer im Kontext besonderer Risikopotentiale.[107] Sie sind als Instrumente der Risikobegrenzung und -kontrolle gedacht. Mit Blick auf die notwendigen Wertediskussionen ist der Informationsbedarf der Nutzer aber nicht auf Risikoinformationen beschränkt. Es bedarf gerade breiterer, erfahrungsgesättigter Kenntnis von Anwendungsbereichen und deshalb gerade auch Transparenz über den Einsatz in wenig riskanten Zusammenhängen.[108] Pflichten zur Kennzeichnung von Verwendungen mit starkem KI-Anteil sollten mit Blick auf diese Funktion breit gefasst werden. Für ihren Anwendungsbereich sollte weniger danach gefragt werden, ob sie zur Risikokontrolle erforderlich sind, sondern nur danach, ob sie für die Anbieter ausnahmsweise unzumutbar werden.

b) Alternativen eröffnen

Bloße Transparenz ermöglicht zwar Erfahrung, hat aber jedenfalls in Kontexten ohne Verhaltensalternativen auch schnell einen zynischen Beigeschmack. Für das gesellschaftliche Selbstexperiment mit KI ist es besonders hilfreich, wenn die einzelnen Nutzerinnen selbst mit dem Gebrauch experimentieren können. Dies setzt Alternativangebote voraus. Im Datenschutz wird schon länger darüber diskutiert, ob Anbieter datengetriebener Dienste den Nutzern als Alternative zur Datenauswertung auch eine Bezahlvariante zur Verfügung stellen sollten.[109] Im Bereich des E-

---

[106] Zur Bedrohung der freien öffentlichen Meinungsbildung durch Algorithmen und zu Regulierungsvorschlägen siehe nur *Drexl*, ZUM 2017, 529 ff. Speziell zur Beeinflussung der politischen Willensbildung durch Social Bots siehe die TA-Vorstudie des Büros für Technikfolgen-Abschätzung beim Deutschen Bundestag: *Kind/Jetzke/Weide/ Ehrenberg-Silies/Bovenschulte*, Social Bots, 2017, S. 40 ff.; dazu auch *Löber/Roßnagel*, MMR 2019, 493 (494 f.). Zur Beeinflussung durch sog. Microtargeting siehe *Hill* (Fn. 84), S. 47 ff.; *Söbbing*, InTeR 4/18, 182 ff.

[107] So wird eine Kennzeichnungspflicht bisher insbesondere für den Einsatz von Algorithmen in persönlichkeitssensiblen Feldern (vgl. *Martini* [Fn. 14], S. 177, 178 f.; siehe auch *Busch* [Fn. 92], S. 58 m. w. N.) und für Social Bots zum Schutz der freien öffentlichen Meinungsbildung (*Löber/Roßnagel* [Fn. 106], S. 494 ff.; *Rößner*, in: Hill/Kugelmann/Martini [Hrsg.], Digitalisierung in Recht, Politik und Verwaltung, 2018, S. 55 [56, 61]) gefordert. Siehe auch die geplante Pflicht zur Kenntlichmachung von Social Bots in Netzwerken nach § 18 Abs. 3 (i. V. m. § 93 Abs. 4) des Entwurfs eines neuen Medienstaatsvertrags in der Beschlussfassung der Konferenz der Regierungschefinnen und Regierungschefs der Länder vom 5.12.2019, abrufbar unter https://www.rlp.de/fileadmin/rlp-stk/pdf-Dateien/Medienpolitik/ModStV_MStV_und_JMStV_2019-12-05_MPK.pdf <19.6.2020>.

[108] So auch der Wunsch in der Bevölkerung: In einer im Dezember 2017 durch das Kompetenzzentrum Öffentliche IT durchgeführten repräsentativen Bevölkerungsfrage mit 1.008 Befragten forderten mehr als 9 von 10 Befragten eine klare Kennzeichnungspflicht automatisiert erstellter Entscheidungen und Inhalte (Umfrage abrufbar unter https://www.oeffentliche-it.de/umfragen?entry=ki-gestalten <19.6.2020>).

[109] So sollte schon nach dem sog. Koppelungsverbot alter Rechtslage (§ 28 Abs. 3b BDSG a. F.) der Vertragsabschluss dann nicht von einer Einwilligung des Betroffenen abhängig gemacht werden dürfen, wenn dem Betroffenen ein anderer Zugang zu gleichwertigen vertraglichen Leistungen ohne die Einwilligung nicht oder nicht in zumutbarer Weise möglich war; siehe dazu ausführlich *Rogosch*, Die Einwilligung im Datenschutzrecht, 2012, S. 84 ff.; vgl. zur neuen Rechtlage nach der DS-GVO auch *Albrecht/Jotzo*, Das neue Datenschutzrecht der EU, 2017, Teil 3 Rn. 44, die die Zurverfügungstellung einer Bezahlvariante sogar als notwendige Voraussetzung der Freiwilligkeit der Einwilligung nach Art. 7 Abs. 4 DS-GVO ansehen; diff. *Krohm/Müller-Peltzer*, ZD 2017, 551 (553 ff.).

Commerce nutzen die Anbieter unter Einsatz von KI ihre Kenntnisse über die Verbraucher zunehmend zur Preisdiskriminierung.[110] Je höher die Zahlungsbereitschaft der Interessenten (ausgedrückt im sog. Reservationspreis) eingeschätzt wird, umso höher ist der Preis, der ihnen angezeigt wird. Auch hier beginnt eine Diskussion darüber, ob den Verbrauchern nicht die Möglichkeit eröffnet werden soll, ein Angebot ohne Preisdiskriminierung nachzufragen. Ins Gespräch gebracht wurde dieser Vorschlag übrigens von zutiefst liberal denkenden Wissenschaftlern vor dem Hintergrund einer ökonomischen Analyse.[111] Und wenn wir noch einmal an die Bewältigung von Rechtstreitigkeiten denken, sollten algorithmengetriebene Angebote zunächst vor allem als Alternative ausgestaltet werden und in der näheren Zukunft die überkommenen Angebote noch nicht ersetzen.

Hier geht es zunächst um den generellen Punkt. Die Eröffnung von Alternativen bietet einen geeigneten Rahmen, um auf der Grundlage konkreter Erfahrungen über die Neujustierung von Werten zu diskutieren. Alternativen fordern und eröffnen ist ein zentrales Element der Diskussionssicherung. Sie sichert darüber hinaus zugleich individuelle Autonomie.

c) KI „erklärbar“ machen

KI hat gerade wegen der ihr eigenen Arbeitsweisen so viel Leistungspotential. Die Mustererkennung zeigt auf Basis der Korrelationen Zusammenhänge auf, die uns bislang verborgen blieben. Die Wertediskussion ist hier nicht auf blinde Entscheidungen darüber beschränkt, in welchen Zusammenhängen wir welche Arbeitsweisen fordern sollen. Denn auch wenn die Arbeitsweisen von KI recht neu und gegenwärtig im fortgeschrittensten Bereich der neuronalen Netze sehr intransparent sind, können wir versuchen sie besser verstehbar zu machen. Explainable AI (XAI) heißt der Forschungsansatz, der sich genau darauf konzentriert, die intransparenten Prozesse der Entscheidungsfindung von KI besser aufzuhellen und die KI selbst über diese Rechenschaft ablegen zu lassen, so dass sie besser kontrollierbar sind und wir eine bessere Entscheidungsgrundlage dafür bekommen, wie wir sie wo zum Einsatz kommen lassen sollten.[112] Die Förderung dieser Forschungen ist eine weitere staatliche Aufgabe im Rahmen seiner Innovationsverantwortung.

IV. Fazit

Welches Fazit lässt sich aus diesen Überlegungen ziehen? Der Philosoph *Nick Bostrom*, Director des Future of Humanity Institute an der Oxford University, wurde in einem Interview gefragt, ob die Entwicklung von KI zur Superintelligenz entweder die Vernichtung der Menschheit oder das Paradies bedeuten würde. Er antwortete nur: „Ja. Auf lange Sicht scheint mir das plausibler als irgendwas dazwischen.“[113] Weil das Paradies bekanntlich nicht auf Erden zu suchen ist, bleibt es Aufgabe des Staates, die Entwicklung in diesem angeblich so unplausiblen „Dazwischen“ zu halten – durch fortlaufende Reflexion der Werte und dauerhafte Umhegung der Risiken.

---

[110] Vgl. *Wagner/Eidenmüller* (Fn. 84), S. 224 f.

[111] Vgl. *Wagner/Eidenmüller* (Fn. 84), S. 227 ff., insb. 228 f.

[112] Vgl. *Kreutzer/Sirrenberg* (Fn. 14), S. 12 f.; *Rademacher* (Fn. 15), S. 377; *Wischmeyer* (Fn. 4), S. 61 f.; siehe auch Nachweise oben, bei Fn. 93.

[113] *Bostrom* im Interview mit ZEIT Campus, abrufbar unter https://www.zeit.de/campus/2015/03/kuenstliche-intelligenz-roboter-computer-menschheit-superintelligenz/seite-2 <19.6.2020>.