

Herbert Zech

Haftung für künstliche Intelligenz: Spagat zwischen Akzeptanz und Innovationsförderung

Vortrag im Rahmen der 63. Bitburger Gespräche

Mainz, 09./10.01.2020

I. Einführung

Der vorliegende Beitrag beschäftigt sich mit dem Haftungsrecht als einem privatrechtlichen Instrument der Risikosteuerung.¹ Das besondere Potential des Haftungsrechts, und insbesondere von Gefährdungshaftungsregelungen, besteht darin, dass sie auch den Einsatz neuartiger, noch in Entwicklung begriffener Technologien steuern und so die mit ihnen verbunden Risiken beeinflussen können.² Damit hat es aktuell auch für die Informationstechnologie Bedeutung, deren Entwicklungen momentan vielfach unter dem Stichwort der künstlichen Intelligenz diskutiert werden.

Im Hintergrund der Diskussion über die Risiken künstlicher Intelligenz steht als zentraler Aspekt ein erhebliches Wissensproblem. Bei den aktuellen Entwicklungen der Informationstechnologie handelt es sich um neuartige Technologien, über deren Risiken unterschiedliche Akteure (Entwickler, Anwender, Staat, Betroffene) zum einen unterschiedliches Wissen haben. Zum anderen ist das verfügbare Wissen über diese Technologien bislang überhaupt begrenzt bzw. noch

¹ Eine ausführliche Untersuchung des bestehenden Rechtsrahmens und möglicher neuer Regelungen zur Haftung für digitale Systeme wird als Gutachten A zum 73. Deutschen Juristentag „Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?“ veröffentlicht werden.

² Zur Funktion des Haftungsrechts als Instrument der Risikosteuerung *Wagner*, *VersR* 1999, 1441 ff.; *Zech*, *JZ* 2013, 21, 23 ff.; *ders.*, in: *Jahrbuch SGHVR* 2016, S. 17 ff. Zur Haftung für digitale Systeme *Wagner*, in: *Faust/Schäfer* (Hrsg.), *Zivilrechtliche und rechtsökonomische Probleme des Internet und der künstlichen Intelligenz*, 2019, S. 1, 6 ff. Speziell zur Haftung bei automatisierten Fahrzeugen *Borges*, *CR* 2016, 272, 277 ff.; *Hey*, *Die außervertragliche Haftung des Herstellers autonomer Fahrzeuge bei Unfällen im Straßenverkehr*, 2019, S. 150 ff.; *Eidenmüller*, *Machine Performance and Human Failure*, <https://www.law.ox.ac.uk/business-law-blog/blog/2018/11/law-and-autonomous-systems-series-machine-performance-and-human> (zuletzt aufgerufen am 10.12.2019).

in der Entwicklung begriffen, beispielsweise im Hinblick auf die Erklärbarkeit lernfähiger Systeme. Die Gefährdungshaftung kann sich in diesem Kontext als sinnvolles Instrument der Techniksteuerung unter Ungewissheit erweisen.³ Durch Gefährdungshaftungsregelungen kann staatliches Handeln der technischen Entwicklung gewisse Leitplanken vorgeben, privates Risikowissen nutzbar machen und auf die Wahl des optimalen Risikolevels hinwirken. So kann neben der Technikakzeptanz auch die Weiterentwicklung der Technik (insbesondere im Hinblick auf ihre Sicherheit) gefördert werden. Das heißt, Gefährdungshaftungsregelungen können nicht nur eine technikbegrenzende, sondern gerade auch eine technikfördernde Wirkung entfalten.

Eine grundlegende Weichenstellung für die haftungsrechtliche Beurteilung von informationstechnischen Systemen stellt die Frage dar, ob diese Systeme als Werkzeuge, oder aber als Akteure zu betrachten sind.⁴ Hier sollen sie nicht als eigenständige Akteure, sondern immer noch als Werkzeuge aufgefasst werden. Dies erscheint nach wie vor sinnvoll. Zwar handelt es sich dabei um immer komplexere und mächtigere Werkzeuge, dennoch lässt sich ihr Einsatz rechtlich beurteilen und gestalten, ohne dass die Perspektive des Werkzeuggebrauchs aufgegeben werden müsste.

II. Technische Entwicklung als Herausforderung für das Haftungsrecht

1. Drei wichtige technische Entwicklungen: Lernfähigkeit, Robotik, Vernetzung

Aus haftungsrechtlicher Sicht lassen sich aktuell drei technische Entwicklungen identifizieren, die wesentlich dazu beigetragen haben, dass Informationstechnologie bzw. digitale Systeme, die Informationen verarbeiten, heute sehr viel mächtiger – und damit sehr viel nützlicher, teilweise aber auch riskanter – sind als noch vor zehn Jahren: Lernfähigkeit, Robotik und Vernetzung.⁵

a) Lernfähigkeit

Das Stichwort Lernfähigkeit bzw. Machine Learning kennzeichnet den Umstand, dass digitale Systeme nicht mehr vollständig programmiert werden müssen, sondern auch selbstständig dazulernen können, genauer gesagt durch den Input von Daten aus der Außenwelt ihre

³ *Spiecker gen. Döhmann*, in: Hill/Schliesky (Hrsg.), *Die Vermessung des virtuellen Raums*, 2012, S. 137, 152. Ebd., S. 153: „Daher sind Haftungsregeln Ausdruck einer Unsicherheitsregel, die zwar einen grundsätzlichen Ausgleich der Positionen von Risiko und Chance wahr, aber in der Tendenz der positiven Bewertung den Vorrang einräumt.“ Zum Begriff Verantwortung als „Ausgleich einer Handlung in Unsicherheit“ *Linardatos*, ZIP 2019, 504, 506. Vgl. *Teubner*, AcP 218 (2018), 155, 175.

⁴ Zu der Frage, ob digitalen Systemen Rechtspersönlichkeit bzw. Rechtsfähigkeit zukommen kann *Teubner*, Zeitschrift für Rechtssoziologie 27 (2006), 5 ff.; *ders.*, in: Becchi/Graber/Luminati (Hrsg.), *Interdisziplinäre Wege der Grundlagenforschung*, 2007, S. 1 ff.; *ders.*, AcP 218 (2018), 155, 160 ff.; *Beck*, JR 2009, 225, 229 f.; *dies.*, AJP 2017, 183 ff.; *Gruber*, in: Beck (Hrsg.), *Jenseits von Mensch und Maschine*, 2012, S. 133 ff.; *ders.*, in: *ders./Bung/Ziemann* (Hrsg.), *Autonome Automaten*, 2. Aufl. 2015, S. 191 ff.; *Eidenmüller*, *The Rise of the Robots and the Law of the Humans*, SSRN Nr. 2941001, S. 11 ff.; *Kluge/Müller*, in: Taeger (Hrsg.), *Smart World – Smart Law?*, 2016, S. 989, 993; *Spindler*, JZ 2016, 805, 816; *Klingbeil*, JZ 2019, 718, 721 f.; *Schirmer*, JZ 2019, 711, 712 ff.; *Wagner*, in: Faust/Schäfer (Hrsg.), *Zivilrechtliche und rechtsökonomische Probleme des Internet und der künstlichen Intelligenz*, 2019, S. 1, 29 ff.; *Riehm/Meier*, in: DGRI Jahrbuch 2018 (erscheint demnächst), S. 14 ff. Vgl. auch *Sorge*, *Softwareagenten*, 2006, S. 33 ff.; *Borges*, NJW 2018, 977, 979.

⁵ Dazu *Zech*, *Risiken digitaler Systeme: Robotik, Lernfähigkeit und Vernetzung als aktuelle Herausforderungen für das Recht*, Weizenbaum Series #2, 2020, https://www.weizenbaum-institut.de/media/News/Weizenbaum_Series/Weizenbaum_Series_2_Zech_070220.pdf.

Verhaltensweise ändern können. Das bedeutet zunächst, dass die Verhaltensweise eines Systems nicht mehr komplett von einem Programmierer vorweggedacht werden muss, er nicht mehr alle möglichen Zustände, die das System einnehmen kann, schon im Vorhinein kennen muss. Zugleich impliziert dies, dass der Programmierer einen geringeren Einfluss auf das Verhalten des Systems ausübt. Die Lernfähigkeit digitaler Systeme ist sicherlich der Aspekt, der die aktuelle Diskussion über die Chancen und Risiken künstlicher Intelligenz und die rechtliche Verantwortung für diese Risiken am stärksten bestimmt. Ganz allgemein gesprochen geht es hier um die Frage der Zurechnung von Folgen des Einsatzes von künstlicher Intelligenz, welche sowohl negativer als auch positiver Natur sein können – sei es im Vertragsrecht, im Haftungsrecht oder etwa auch im Immaterialgüterrecht. Allerdings ist die Lernfähigkeit nicht der einzige Aspekt, der die Risiken und Chancen informationstechnischer Systeme bestimmt.

b) Robotik

Im Haftungsrecht ist die rechtliche Verantwortung für die Folgen des Einsatzes informationstechnischer Systeme auch schon vor der Verbreitung lernfähiger Systeme diskutiert worden, und zwar zunächst mit Blick auf Software. Fahrt aufgenommen hat die Diskussion dann aber vor allem durch die Entwicklungen in der Robotik. Robotik kann als Kopplung digitaler Systeme mit physischen Sensoren und Aktuatoren definiert werden. Das ist nicht selbstverständlich; zwar benötigt Software immer eine Hardware, auf der sie ausgeführt wird. Klassischerweise kennen wir digitale Systeme aber als Systeme, die Informationen letztlich an Menschen ausgeben. Menschen nutzen dann diese Informationen bzw. reagieren darauf mit einem bestimmten Verhalten. Durch die Robotik entsteht hingegen ein neuer Schädigungspfad, bei dem digitale Systeme physische Rechtsgüter wie Leib und Leben oder Eigentum ohne menschliches Dazwischentreten schädigen können. Es entstehen mithin neue Risikoquellen für bekannte Rechtsgüter.

c) Vernetzung

Als dritter Aspekt, der bislang eher unter dem Stichwort der kritischen Infrastrukturen im Sicherheitsrecht und weniger im Haftungsrecht diskutiert wurde, kommt schließlich immer mehr die Vernetzung digitaler Systeme hinzu. Während im klassischen Internet die informationstechnische Vernetzung den Austausch von Informationen zwischen Menschen betraf, findet nun mit dem Internet of Things eine weitergehende Vernetzung von Geräten statt.⁶ Roboter, also digitale Systeme mit unmittelbarer Hardwaresteuerung, sind in der Folge direkt untereinander vernetzt. Wichtigstes Beispiel dürften vollautomatisierte Fahrzeuge sein. Ein Auto kann sich nur dann selbstständig durch den Verkehr bewegen, wenn es mit einer Vielzahl von anderen Fahrzeugen gleichzeitig vernetzt ist (eventuell auch über eine Plattform), von denen es laufend Daten bekommt und daraus erst sein konkretes Verhalten errechnet. Hinzu kommt, dass immer mehr Alltagsgegenstände mit informationsverarbeitenden Systemen ausgerüstet sind, wie etwa der „intelligente“ Kühlschrank. Vorreiter ist der industrielle Bereich als Industrie 4.0, also die Vernetzung in der industriellen Produktion. Die Vernetzung führt zu zusätzlichen neuen

⁶ Das Internet of Things (IoT) kann definiert werden als „global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies“ (International Telecommunication Union, Recommendation ITU-T Y.2060, 3.2.3). Zum Begriff *Hornung/Hofmann*, in: Hornung (Hrsg.), Rechtsfragen der Industrie 4.0, 2018, S. 9 f. Vgl. auch *Sosnitza*, CR 2016, 764, 765. Eine Definition cyber-physikalischer Systeme im Zusammenhang mit der rechtlichen Diskussion findet sich bei *Reichwald/Pfisterer*, CR 2016, 208: „Sie sind als Systeme definiert, die physische Komponenten mit Software verflechten, wobei jedes CPS in unterschiedlichen zeit- und räumlichen Bereichen operieren kann, unterschiedliche Modalitäten und Verhaltensweisen aufweist und in vielfältiger Weise mit anderen Systemen kontextabhängig kommunizieren und interagieren kann.“ Vgl. *Möller*, Guide to Computing Fundamentals in Cyber-Physical Systems, 2016, S. 81 ff. Ebd., S. 141: „networked interconnection of everyday objects“.

Herausforderungen, nicht nur im Bereich der Sicherheit im regulären Betrieb, sondern auch im Bereich der Sicherheit vor Angriffen.

2. Herausforderungen für das Haftungsrecht

a) Neue Risiken

Die skizzierten technischen Entwicklungen stellen aus haftungsrechtlicher Sicht eine Herausforderung dar. Als wichtigster Punkt ist zunächst das Entstehen neuer Risiken zu nennen, wobei Risiko hier als die Möglichkeit eines Schadenseintritts verstanden wird, die sich als Produkt von Schadenshöhe und Eintrittswahrscheinlichkeit mit einer bestimmten Genauigkeit und Zuverlässigkeit quantifizieren lässt.⁷ Die Risiken aktueller digitaler Systeme können auch als „KI-Risiken“ bezeichnet werden. Dabei geht es um neue Schadensszenarien bzw. Schädigungswege infolge des Einsatzes digitaler Systeme, die neue Möglichkeiten der Schadensentstehung bedingen. Lernfähigkeit kann zu unvorhergesehenem Verhalten digitaler Systeme führen. Bei einer unmittelbaren Hardwaresteuerung können dadurch Schäden in der physischen Welt auftreten. Daher stellen lernfähige Roboter, insbesondere wenn sie etwa als humanoide Roboter in den Alltag eingebunden sind, ein besonderes Risiko dar. Durch die Vernetzung kann auch die Übertragung fehlerhafter Daten zur Schadensquelle werden, etwa bei selbstfahrenden Fahrzeugen.

b) Verschiebung der Risikobeherrschung

Daneben kann der Einsatz digitaler Systeme für Aufgaben, die bislang durch Menschen oder anderweitig ohne solche Systeme erfüllt wurden, dazu führen, dass nunmehr andere Akteure Einfluss auf die Schadensentstehung gewinnen und damit bekannte Risiken anderen Akteuren zugeordnet werden müssen. So führt etwa bei Fahrzeugen die zunehmende Automatisierung dazu, dass Programmierer mehr Einfluss auf das Fahrverhalten und mithin auf ein etwaiges Schadensgeschehen haben, während der Einfluss der Fahrer abnimmt. Bei lernfähigen Systemen wiederum verringert sich der Einfluss der Programmierer zugunsten der Einflussmöglichkeiten von Trainern. Das heißt, es treten hier weitere Akteure neben die Programmierer, die ebenfalls Risikobeherrscher sein können und insofern auch Adressaten von Regelungen.

⁷ Zum Risikobegriff *Seiler*, *Recht und technische Risiken*, 1997, S. 38 ff.; *Lepsius*, *VVDStRL* 63 (2004), 264 ff.; *Dietz*, *Technische Risiken und Gefährdungshaftung*, S. 5 ff. Ebenso für das Datenschutzrecht *Martini*, in: *Paal/Pauly* (Hrsg.), *DS-GVO, BDSG*, 2. Aufl. 2018, Rn. 15a f. Zur Steuerung von Technikrisiken *Zech*, *JZ* 2013, 21 ff.; *ders.*, in: *Jahrbuch SGHVR* 2016, S. 17, 19 f.

c) Komplexe Kausalverläufe, Beweisbarkeitsprobleme

Als dritter Aspekt ist zu nennen, dass es infolge zunehmender Vernetzung zu Szenarien kommen wird, bei denen es mit Blick auf die (zahlreichen) beteiligten Akteure immer schwerer nachzuweisen sein wird, dass sie schadensursächlich geworden sind.⁸ Es wird also durch komplexe Kausalverläufe zu Beweisproblemen kommen. Zwar lässt sich bei einer Schädigung durch physische Maschinen bzw. Roboter immer die schädigende Maschine identifizieren, der auch Hersteller und Betreiber bzw. Anwender zugeordnet werden können. Die Vielzahl weiterer Akteure, die Einfluss auf die konkrete Schädigung gehabt haben könnten, wird sich voraussichtlich aber schwerer identifizieren lassen, etwa, wenn Daten anderer Fahrzeuge oder von Plattformen schadensursächlich werden. Herstellung und Betrieb bzw. Einsatz der schädigenden Maschine sind in solchen Fällen zwar auch kausal für eine Schädigung, eine daran anknüpfende Haftung kann aber zu Fehlsteuerungen führen.

3. Akteure

a) Herstellerseite: Entwicklung und Herstellung

Bei den beteiligten Akteuren ist zwischen der Hersteller- und der Anwenderseite zu unterscheiden. Auf der Herstellerseite stehen am Anfang der Wertschöpfungskette die Entwickler. In der Regel löst die bloße Entwicklung einer Technologie noch keine haftungsrechtliche Verantwortung aus. Allerdings sind Fälle denkbar, in denen bereits die bloße Verbreitung einer besonders gefährlichen Technologie als pflichtwidrig und damit im Falle einer mittelbaren Schadensverursachung auch als haftungsauslösend angesehen werden könnte. Als nächstes ist der Hersteller zu nennen, der häufig zugleich Entwickler ist. Bei lernfähigen Systemen kann der Hersteller Programmierer und Trainer sein, das Training kann aber auch von der Anwenderseite übernommen werden.

b) Anwenderseite: Betrieb und Benutzung

Auf der Anwenderseite kann zwischen Betreibern (welche die gesamte Benutzung eines Systems steuern) und bloßen Benutzern digitaler Systeme unterschieden werden (etwa Halter und Fahrer bei selbstfahrenden Fahrzeugen). Übernimmt ein Betreiber das Training lernfähiger Systeme, kommt ihm eine größere Risikobeherrschung zu als bei klassischen digitalen Systemen, wo es für das Verhalten des Systems allein auf die herstellerseitige Programmierung ankam. Zudem wird sich auf der Anwenderseite bei komplexen digitalen Systemen auch die Kluft zwischen Betreibern mit besonderem Risikowissen, insbesondere professionellen Betreibern, und solchen ohne besonderes Risikowissen auf tun. Unternehmen, deren Hauptgegenstand der Betrieb digitaler Systeme wie etwa automatisierte Fahrzeugflotten darstellt, kommt damit eine besondere Stellung zu. Insgesamt bedeutet besseres Technikwissen stärkere Risikobeherrschung.

⁸ *Spiecker gen. Döhmman*, CR 2016, 689, 700 f.: „Untergang des Handelnden“ durch „Teilnahme einer unüberschaubar großen Zahl an weiteren Teilnehmern mit unterschiedlicher bis höchst geringer Einflussnahme auf die Ausgestaltung und Durchführung“; *Teubner*, AcP 218 (2018), 155, 202: „Abgrenzung der Haftungssubjekte praktisch unmöglich“.

III. Steuerung von „KI-Risiken“ durch Haftungsrecht

Das Haftungsrecht spielt mit Blick auf die Entwicklung und den Einsatz künstlicher Intelligenz eine wichtige Rolle als Instrument der indirekten Risikosteuerung. Ziel ist es, ein aus ökonomischer Sicht optimales Risikoniveau beim Einsatz der neuen Technologien zu finden. Dies bedeutet nicht Schadensvermeidung um jeden Preis, sondern nur in einem Umfang, der in der Zusammenschau mit den Chancen riskanter Handlungen den größten volkswirtschaftlichen Nutzen bringt.⁹ Im Folgenden soll ein Überblick über drei verschiedene rechtliche Grundprinzipien der Risikosteuerung gegeben werden, die sich nach ihrer Anreizwirkung unterscheiden: die Verschuldenshaftung, die Gefährdungshaftung und die haftungsersetzende Versicherung.

1. Verschuldenshaftung

Die Verschuldenshaftung stellt mit § 823 Abs. 1 BGB den Grundfall der deliktischen Haftung dar. Im Rahmen der Verschuldenshaftung ist der zurechenbare Verursacher nur verantwortlich, wenn er sich pflichtwidrig verhalten hat, was sich insbesondere aus der Verletzung von Verkehrspflichten ergibt.¹⁰ Im Kontext des Einsatzes digitaler Systeme begründet die Verschuldenshaftung Anreizeffekte sowohl für das Verhalten der Hersteller und Anwender dieser Systeme, als auch für das Verhalten potentiell Betroffener.

a) Steuerung des Sorgfaltsniveaus auf der Seite der Technikanwender

Die Verschuldenshaftung dient insbesondere der Steuerung des Sorgfaltsniveaus.¹¹ Sie greift nur ein, wenn sich ein Schädiger rechtswidrig bzw. verkehrspflichtwidrig verhalten hat. Damit wird für die Teilnehmer am Rechtsverkehr ein Anreiz gesetzt, die erforderliche Sorgfalt zu wahren, mehr jedoch nicht. Was erforderlich ist, um sich im Rahmen der erforderlichen Sorgfalt zu bewegen, richtet sich nach den Verkehrspflichten, deren Bestimmung typischerweise durch die Judikative erfolgt. Ihr liegt eine Abwägung zugrunde, bei der der potentielle gesellschaftliche Nutzen eines Verhaltens ebenso wie dessen Risiken berücksichtigt werden. Im Ergebnis wird also richterrechtlich festgesetzt, welches Risikolevel als tragbar erscheint. Ein optimales Risikolevel kann auf diesem Weg aber nur dann festgelegt werden, wenn die Judikative über ausreichendes Risikowissen verfügt. Obwohl die richterrechtliche Bestimmung des zulässigen Risikolevels flexibler ist als es eine gesetzliche Bestimmung sein könnte, bleibt das Problem, dass staatliche Akteure grundsätzlich ein geringeres Technik- und damit Risikowissen haben dürften als die Entwickler und Hersteller einer gegebenen Technik. Bei einer Technik, die überhaupt erst in Entwicklung begriffen ist, sollte deshalb für die Bestimmung, wo ein Optimum an Chancen und Risiken liegt, das Wissen dieser privaten Akteure fruchtbar gemacht werden. Eine Möglichkeit dazu liegt, wie noch zu zeigen sein wird, in einer verschuldensunabhängigen Haftung.

⁹ *Mazzini*, A System of Governance for Artificial Intelligence through the Lens of Emerging Intersections between AI and EU Law, SSRN Nr. 3369266, S. 13: „a question of determining the normatively acceptable risk level when balanced against the expected benefits of a certain product, activity or technology“. Zu den Steuerungszielen des Haftungsrechts *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, 5. Aufl., 2012, S. 145 ff.; *Kötz/Wagner*, Deliktsrecht, 13. Aufl. 2016, S. 30 ff.

¹⁰ Nach zutreffender Ansicht ist immer nach der Verletzung von Verkehrspflichten zu fragen, die bei der unmittelbaren Verletzung nur indiziert ist: *MüKo/Wagner*, § 823 BGB, Rn. 6, 21, 26; *Kötz/Wagner*, Deliktsrecht, 13. Aufl. 2016, S. 51, 57 ff.; *Looschelders*, Schuldrecht BT, 14. Aufl. 2019, S. 478; *Zech*, in: Jahrbuch SGHVR 2016, S. 17, 25.

¹¹ Zur Steuerungswirkung der Verschuldenshaftung *Wagner*, VersR 1999, 1441 ff.; *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, 5. Aufl., 2012, S. 181 ff.; *Zech*, JZ 2013, 21, 23 ff.; *ders.*, in: Jahrbuch SGHVR 2016, S. 17 ff.

Aus der Sicht der Technikanwender bedeutet die richterrechtliche Bestimmung schließlich geringere Rechtssicherheit. Dies wird noch dadurch verschärft, dass bei der Entscheidung konkreter Fälle auch Rückschaufehler drohen. Wenn der Wissensstand zum Zeitpunkt der richterlichen Beurteilung diese beeinflusst, wird das zum Zeitpunkt der Schädigung für den Schädiger erkennbare Risiko höher eingeschätzt als es tatsächlich war, was wiederum zu einer schärferen Haftung führt.

b) Steuerung des Verhaltens Betroffener, Technikförderung in historischer Perspektive

In der Verschuldenshaftung verbirgt sich noch ein weiterer Anreizeffekt, nämlich ein Anreiz für die Betroffenen (Dritte oder auch Nutzer), sich selbst um Schadensvermeidung zu kümmern. Soweit Hersteller und Betreiber sich pflichtgemäß verhalten und daher nicht haften, tragen die Betroffenen potentielle Schäden (allgemeines Lebensrisiko). Daraus ergibt sich für sie ein Anreiz, das ihnen Mögliche zu tun, um Schäden zu vermeiden. Dies setzt allerdings voraus, dass überhaupt für die Betroffenen eine Vermeidungsmöglichkeit besteht. Daher schlägt bei neuartigen Systemen, über die durchschnittliche Betroffene gerade keine Risikokenntnis haben, diese Anreizwirkung der Verschuldenshaftung fehl. Dies gilt umso mehr, wenn Vermeidungsmöglichkeiten faktisch fehlen, wie es bei ubiquitären Systemen, denen man im Alltag nicht entkommen kann, der Fall ist.¹²

Im Ergebnis führt daher eine Verschuldenshaftung bei vielen neuen Technologien zu einer Technikförderung auf Kosten der Betroffenen. So mag es etwa in der Frühphase der Industrialisierung gewesen sein.¹³ Heute erscheint er jedoch nicht akzeptabel, nicht nur aufgrund von Gerechtigkeitsbedenken, sondern auch, weil er in mangelnde Technikakzeptanz umschlagen kann. Bei Risiken neuer Technologien handelt sich gerade nicht um das allgemeine Lebensrisiko.

¹² Allerdings könnte man solche Vermeidungsmöglichkeiten auch regulatorisch sicherstellen, etwa indem man eine Pflicht zum alternativen Angebot lernfähiger und entsprechender nicht-lernfähiger Produkte normiert oder vorschreibt, dass die Lernfähigkeit für den Anwender abschaltbar sein muss.

¹³ *Ott*, in: ders./Schäfer (Hrsg.), *Allokationseffizienz in der Rechtsordnung*, 1989, S. 25, 29; *Kötz/Wagner*, *Deliktsrecht*, 13. Aufl. 2016, S. 13; vgl. *Ogorek*, *Untersuchungen zur Entwicklung der Gefährdungshaftung*, 1975, 135.

c) Herausbildung von Verkehrspflichten, Vorteile und Probleme

Der Benachteiligung Betroffener kann begegnet werden, indem die Verkehrspflichten, die sich richterrechtlich herausbilden, einfach sehr hoch eingestellt werden. Werden entsprechend hohe Verkehrspflichten angenommen, nähert sich die bestehende Verschuldenshaftung für bestimmte Bereiche, etwa Technikrisiken, der Gefährdungshaftung an. Dies begegnet jedoch kompetenziellen Bedenken (Verschiebung von Regelungskompetenz von der Legislative auf die Judikative) und führt – gerade bei neuen Technologien, für die noch keine feste Rechtsprechung existiert – zu Rechtsunsicherheit. Unter Anreizgesichtspunkten ist dies verheerend. Möchte man die Entwicklung und Einführung nützlicher neuer Technologien nicht unnötig behindern, sollte für größtmögliche Rechtssicherheit gesorgt werden.

Ein zweites Problem besteht darin, dass gerade bei solchen Technologien, die in rascher Entwicklung begriffen sind, zahlreiche Entwicklungsrisiken auftreten. Bei solchen Risiken, von denen zum Zeitpunkt der Schädigungshandlung niemand Kenntnis haben kann (vgl. § 1 Abs. 2 Nr. 5 ProdHaftG), versagt die Verschuldenshaftung. Es wäre aber sinnvoll, diejenigen haften zu lassen, die am nächsten an der Entwicklung dran sind und die Entwicklungsrisiken am ehesten überblicken können, damit hier ein Anreiz gesetzt wird, sich das erforderliche Risikowissen zu erarbeiten. Dieser Gedanke liegt – unter anderen – der Gefährdungshaftung zugrunde.

2. Gefährdungshaftung

Zur rechtlichen Bewältigung der mit künstlicher Intelligenz bzw. aktuellen digitalen Systemen verbundenen Risiken wird die Einführung von Gefährdungshaftungsregelungen diskutiert.¹⁴ Diese würden gegenüber der bestehenden Verschuldenshaftung (und auch gegenüber der Produkthaftung¹⁵) eine weitergehende Steuerungswirkung entfalten.

¹⁴ Vorschläge in der Literatur für eine Gefährdungshaftung der Hersteller: *Eidenmüller*, The Rise of the Robots and the Law of the Humans, SSRN Nr. 2941001, S. 8 ff. (mit besonderem Fokus auf das Verhältnis von Fahrzeughersteller und Herstellern von Fahrzeugteilen); *Lohmann*, EJRR 2/2016, 335, 337; vgl. *Hanisch*, in: Hilgendorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, S. 27, 54 ff.; *Zech*, in: Gless/Seelmann (Hrsg.), Intelligente Agenten und das Recht, 2016, S. 163, 200; *Hey*, Die außervertragliche Haftung des Herstellers autonomer Fahrzeuge bei Unfällen im Straßenverkehr, 2019, S. 231 ff.; *Spindler*, in: Lohsse/Schulze/Staudenmayer (Hrsg.), Liability for Artificial Intelligence and the Internet of Things, 2019, S. 125, 136 ff. Für eine Gefährdungshaftung der Anwender/Nutzer/Betreiber autonomer Systeme *Chrastaller u. a.*, Robotik, 2001, S. 154 ff.; *Schulz*, Verantwortlichkeit bei autonom agierenden Systemen, 2015, S. 364 ff. (für eine Haftung der „Halter“); *Gless/Janal*, JR 2016, 561, 574; *Horner/Kaulartz*, CR 2016, 7, 13 f.; *Schirmer*, JZ 2016, 660, 665 f.; *ders.* RW 2018, 453, 473; *Riehm/Meier*, in: DGRI Jahrbuch 2018 (erscheint demnächst), S. 11. Vgl. *Günther*, Roboter und rechtliche Verantwortung, 2016, S. 237 ff.; *Zech*, in: Gless/Seelmann (Hrsg.), Intelligente Agenten und das Recht, 2016, S. 163, 201 f.

¹⁵ Die Produkthaftung verzichtet zwar auf das Verschuldenserfordernis, dennoch handelt es sich um eine Haftung für pflichtwidriges Verhalten. Eine ausführliche Analyse der Haftung für digitale Systeme im Produktpfad findet sich bei *Wagner*, AcP 217 (2017), 707 ff.; *von Westphalen*, ZIP 2019, 889, 890; *Mazzini*, A System of Governance for Artificial Intelligence through the Lens of Emerging Intersections between AI and EU Law, SSRN Nr. 3369266, S. 16 ff.; *Riehm/Meier*, EuCML 2019, 161 ff.

a) Zusätzliche Steuerung des Aktivitätsniveaus

Gefährdungshaftungstatbestände steuern nicht nur das Sorgfalts-, sondern auch das Aktivitätsniveau.¹⁶ Sie weisen dem Schädiger das ökonomische Risiko unabhängig davon zu, ob er sich pflichtgemäß oder pflichtwidrig verhält, das heißt, sie internalisieren das Risiko vollständig (zumindest soweit es bzw. der Schaden ersatzfähig ist). Der Risikobeherrscher muss demnach abwägen, ob der zu erwartende Nutzen einer Aktivität deren Risiko übersteigt. Lohnt es sich nicht, so wird die riskante Tätigkeit unterbleiben.

Durch die Delegation dieser Abschätzung auf die Technikentwickler und -anwender wird privates Risikowissen nutzbar gemacht. Gefährdungshaftungstatbestände dienen damit der „Erschließung privater Risikoinformationen“¹⁷, sie haben sich insoweit als „bewährtes Instrument“¹⁸ erwiesen. Dies ist gerade bei neuartigen Technologien sinnvoll, bezüglich derer der Staat – sei es als Gesetzgeber, vollziehende Fachbehörde oder Gericht – über deutlich weniger Risikowissen verfügt als die Technikentwickler und -anwender.

Allerdings wird ein Wissensgefälle für Entwicklungsrisiken bei neuartigen Technologien generell bestritten. Wenn auch den beteiligten privaten Akteuren das Risiko nicht erkennbar ist, kann, so die Argumentation, die Gefährdungshaftung keine Steuerungswirkung entfalten.¹⁹ Dem ist jedoch entgegenzuhalten, dass bei neuartigen Technologien zumindest erkennbar ist, dass *möglicherweise* neue Risiken entstehen – das Vorliegen von Entwicklungsrisiken ist eher eine graduelle als eine einer Ja-Nein-Entscheidung zugängliche Frage.²⁰ Es ist daher durchaus sinnvoll, demjenigen, der erkennen kann, dass er mit einer bestimmten Wahrscheinlichkeit Risiken setzt, die Abwägung aufzubürden, ob es sinnvoll ist, dies zu tun.

Ergänzend sei noch erwähnt, dass auch sichergestellt werden müsste, dass der gewünschte Internalisierungseffekt nicht etwa dadurch fehlschlägt, dass als Technikanwender Kapitalgesellschaften mit zu geringer Kapitalausstattung auftreten (Plünderungsstrategien). Dies kann durch Pflichthaftpflichtversicherungen geschehen (was eher für eine Absicherung von massenhaft auftretenden kleinen Risiken sinnvoll wäre, vgl. § 1 PflVG) oder durch eine Deckungsvorsorge entsprechend den Regelungen bei der Umwelt- oder Gentechnikhaftung (§ 19 UmweltHG, § 36 GenTG).

b) Anreiz zur Weiterentwicklung

Hinzu kommt eine weitere wichtige Anreizwirkung der Gefährdungshaftung: Sie setzt einen Anreiz zur Weiterentwicklung von Technologien, um sie sicherer zu machen.²¹ Solange dem Technikanwender die Abschätzung des Risikos nicht mit hinreichender Zuverlässigkeit möglich ist, wird er die Anwendung unterlassen. Er hat aber zugleich einen Anreiz, die vorhandene Technologie weiterzuentwickeln bis sie sicher genug für die Anwendung ist. Dies setzt zwar voraus, dass der Haftende auch in der Lage ist, die Technologie weiterzuentwickeln. Durch vertragliche Beziehungen, etwa zwischen Hersteller und Anwender, werden jedoch entsprechende Internalisierungseffekte weitergereicht.

¹⁶ Zur Steuerungswirkung der Gefährdungshaftung *Wagner*, *VersR* 1999, 1441, 1443 f.; *Schäfer/Ott*, *Lehrbuch der ökonomischen Analyse des Zivilrechts*, 5. Aufl., 2012, S. 228 ff.; *Zech*, in: *Jahrbuch SGHVR* 2016, S. 17, 29 ff.

¹⁷ *Wagner*, *VersR* 1999, 1441, 1444.

¹⁸ *Lepsius*, *VVDStRL* 63 (2004), 264, 296.

¹⁹ *Schäfer/Ott*, *Lehrbuch der ökonomischen Analyse des Zivilrechts*, 5. Aufl., 2012, S. 385.

²⁰ *Zech*, in: *Jahrbuch SGHVR* 2016, S. 17, 30.

²¹ *Wagner*, *VersR* 1999, 1441, 1450; *Zech*, in: *Jahrbuch SGHVR* 2016, S. 17, 30 f.

c) Technikakzeptanz auf Betroffenenseite

Ein dritter Effekt von Gefährdungshaftungsregelungen liegt darin, dass bei einer vollständigen Risikointernalisierung die Technikakzeptanz auf Seiten der Betroffenen erhöht wird. Haftungsregelungen erleichtern damit die Einführung neuer Technologien und tragen so zur Technikförderung bei.

d) Problem: Abschreckungseffekt?

Das wichtigste Argument gegen die Einführung einer Gefährdungshaftung liegt darin, dass sie einen Abschreckungseffekt und in der Folge auch innovationshemmende Effekte zeitigen könnte. Allerdings zeigen die eben angestellten Überlegungen, dass eine Gefährdungshaftung dem Haftpflichtigen nur die Abschätzung aufbürdet, ob er die Technik verwenden möchte oder nicht; die Anwendung der Technik verbietet sie gerade nicht. Sie stellt damit im Vergleich zu einem unmittelbaren Eingriff durch Regulierungsrecht die bessere Alternative dar, da ein solcher Eingriff möglicherweise – wenn der Staat mangels eigenen Risikowissens nicht anders entscheiden kann – sogar ein Verbot sein müsste. Natürlich wären auch geringere Eingriffe denkbar wie ein Verbot mit Erlaubnisvorbehalt, was aber immer noch ein gewisses Risikowissen auf Seiten des Staates voraussetzt. Die Gefährdungshaftung ist damit ein Regelungsinstrument, das es erlaubt, mit einer Situation unsicherer Risikobewertung umzugehen und dabei von der grundsätzlichen Freiheit der Technikanwendung auszugehen („freiheitsbetonende Unsicherheitenregel“).²² Gefährdungshaftung ist die Haftung für gesellschaftlich grundsätzlich erwünschtes Tun, für das aber die passenden Anreize gesetzt werden sollen.

Wichtig ist dabei, dass die Gefährdungshaftung so ausgestaltet wird, dass sie rechtssicher anwendbar ist und in der Folge über die bestehenden Haftungsrisiken Klarheit besteht. Insbesondere müssen klare Voraussetzungen und klare Rechtsfolgen definiert werden, und auch Haftungshöchstgrenzen, mit denen die Versicherbarkeit des Haftungsrisikos gewährleistet wird, tragen dazu bei, dass das internalisierte Risiko für die Technikanwender kalkulierbar ist. Eine klare, kalkulierbare Haftung verhindert unnötige Abschreckungseffekte.

e) Problem: Nachweis kausaler Verursachung erforderlich

Ein Problem von Gefährdungshaftungsregelungen aus Geschädigtensicht besteht darin, dass sie – wie jede Haftungsregelung – die kausale Verursachung des Schadens durch den Haftpflichtigen voraussetzen. Hier kann dem Geschädigten zwar durch Vermutungsregelungen geholfen werden. Es bleibt aber dabei, dass materiell ein konkreter Verursacher feststehen muss. Dies setzt notwendig voraus, dass auch ein verursachendes System bestimmbar sein muss, was möglicherweise nicht immer der Fall ist. Gerade durch die zunehmende Vernetzung wird die Bestimmung einzelner Systeme und kausaler Verursacher immer schwieriger. Läuft eine Gefährdungshaftung wegen der Vernetzung ins Leere und können auch besondere Beweislastregelungen nicht helfen, drohen Fehlanreize.²³ So kann eine Gefährdungshaftung für denjenigen, der als Betreiber einer schädigenden Hardware-Komponente bestimmbar ist, während ein entfernterer pflichtwidriger Verursacher (etwa durch Lieferung fehlerhafter Daten) nicht zur Rechenschaft gezogen wird, bedingen, dass entferntere Schädiger sich unsorgfältig verhalten. Es kann auch Situationen geben, in denen niemand als zurechenbarer Verursacher identifiziert werden kann, mit dem Ergebnis, dass die Betroffenen das Risiko tragen, was ebenfalls Anreize für

²² *Spiecker gen. Döhmman*, in: Hill/Schliesky (Hrsg.), *Die Vermessung des virtuellen Raums*, 2012, S. 137, 152. Ebd., S. 153: „Daher sind Haftungsregeln Ausdruck einer Unsicherheitsregel, die zwar einen grundsätzlichen Ausgleich der Positionen von Risiko und Chance wahrt, aber in der Tendenz der positiven Bewertung den Vorrang einräumt.“

²³ *Spindler*, in: Lohsse/Schulze/Staudenmayer (Hrsg.), *Liability for Artificial Intelligence and the Internet of Things*, 2019, S. 125, 140.

unsorgfältige Technikanwendung schüfe. In solchen Szenarien muss man darüber nachdenken, über die Gefährdungshaftung hinauszugehen.

3. Haftungsersetzende Versicherung

Eine Möglichkeit, die Risiken vernetzter digitaler Systeme zu bewältigen, wäre eine gesetzliche Versicherung von „KI-Unfällen“ nach dem Vorbild der bestehenden gesetzlichen Unfallversicherung für Arbeitsunfälle.²⁴

a) Vergemeinschaftung von „KI-Risiken“ durch Versicherung

Eine haftungsersetzende Unfallversicherung würde zu einer Vergemeinschaftung von „KI-Risiken“ führen. Sie ähnelt darin einer Pflichthaftpflichtversicherung, die im Zusammenspiel mit einer Gefährdungshaftung zu ähnlichen Ergebnissen führen würde. Gefährliches, aber gesellschaftlich erwünschtes Verhalten wird in diesen Szenarien dadurch ermöglicht, dass das Risiko auf das Kollektiv der Haftpflichtigen bzw. Versicherten umgelegt wird.

b) Nachweis eines „KI-Unfalls“ genügt (haftungsersetzende Versicherung)

Eine haftungsersetzende Versicherung verzichtet auf eine Haftungsregelung und führt zu einem direkten Schadensausgleich durch das Kollektiv. Aus Geschädigtensicht besteht der entscheidende Unterschied darin, dass nicht mehr die Verursachung durch einen bestimmten Haftpflichtigen nachgewiesen werden muss, sondern dass die Verursachung durch eine ihrer Art nach von der Versicherung umfasste Schadensquelle genügt. Dies hätte den Vorteil, eine hohe Technikakzeptanz zu gewährleisten. Entsprechend den Regelungen bei der gesetzlichen Unfallversicherung könnten alle Schäden erfasst werden, für deren Entstehung der Einsatz bestimmter digitaler Systeme zumindest wesentliche Bedingung gewesen ist. Plakativ könnte man von „KI-Schäden“ sprechen, präziser von Schäden besonders riskanter digitaler Systeme (die dann im Einzelnen zu definieren wären).

Die Finanzierung einer solchen Versicherung würde über ein Beitragssystem erfolgen, wobei die Bestimmung des Kreises der Beitragspflichtigen ähnlich wie bei einer Gefährdungshaftung erfolgen kann. Als Versicherungsträger würde eine Organisation fungieren, die ähnlich wie Berufsgenossenschaften auch regulierend und mit Hoheitsgewalt tätig werden kann (alternativ könnte man auch daran denken, die Berufsgenossenschaften selbst heranzuziehen). Ein großer Vorteil solcher zentralen hoheitlichen Organisationen besteht darin, dass sie als Instrument staatlicher Informationsgewinnung dienen können. Sie können Risikowissen generieren, das dann der Allgemeinheit zur Verfügung steht.

²⁴ Dazu *Armbrüster*, in: Gless/Seelmann (Hrsg.), *Intelligente Agenten und das Recht*, 2016, S. 205, 221 ff.; *Janal*, in: Gless/Seelmann (Hrsg.), *Intelligente Agenten und das Recht*, 2016, S. 141, 157; *Schaub*, JZ 2017, 342, 345; *Borges*, in: Lohsse/Schulze/Staudenmayer (Hrsg.), *Liability for Artificial Intelligence and the Internet of Things*, 2019, S. 145, 159 ff.; *Hey*, Die außervertragliche Haftung des Herstellers autonomer Fahrzeuge bei Unfällen im Straßenverkehr, 2019, S. 209 ff.; *Zech*, in: Gless/Seelmann (Hrsg.), *Intelligente Agenten und das Recht*, 2016, S. 163, 202; *ders.*, ZfPW 2019, 198, 208 f., 216 f. Zu Innovationsrisiken *Ann*, in: Seehafer u. a. (Hrsg.), *Risikoregulierung und Privatrecht*, Stuttgart 1994, S. 37, 55; *Wagner*, ZEuP 2007, 1122, 1132 ff. Zur Haftungsersetzung durch Versicherungsschutz *Kötz/Wagner*, *Deliktsrecht*, 13. Aufl. 2016, S. 3 f., 230 f. Vgl. *Hanisch*, in: Hilgendorf (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, S. 27, 43; *Spiecker gen. Döhmman* CR 2016, 698, 704.

c) Problem: Fehlanreize?

Die Idee einer haftungsersetzenden Versicherung sieht sich aber auch Einwänden ausgesetzt. Entfällt die Steuerungswirkung von Haftungsregelungen ersatzlos, können Fehlanreize drohen.²⁵ Durch entsprechende Regelungen zu Regress (vgl. § 110 SGB VII) und Beitragsanpassung (vgl. §§ 157 ff. SGB VII) kann aber auch hier eine sinnvolle Anreizstruktur geschaffen werden.²⁶ Zwar besteht dann wieder die Herausforderung, die pflichtwidrige Verursachung eines KI-Schadens durch einen bestimmten Akteur nachzuweisen, der Versicherer als Prozesspartei steht aber mit Technikentwicklern und -anwendern eher auf Augenhöhe als ein Betroffener.

Als Beitragspflichtige könnte man, auch wenn jedermann versichert ist, nur diejenigen heranziehen, die kraft ihres Risikowissens Einfluss auf das Risiko haben. Dies wären typischerweise Hersteller und professionelle Betreiber. Ähnlich wie eine Gefährdungshaftung dieser Akteure könnte auch eine nach Risikoklassen gestufte Beitragspflicht mit einem Malussystem für bestimmtes Fehlverhalten die nötige Anreizwirkung für einen sorgfältigen und sinnvollen Einsatz digitaler Systeme sowie für deren Weiterentwicklung (insbesondere für mehr Sicherheit) entfalten.

²⁵ *Borges*, in: Lohsse/Schulze/Staudenmayer (Hrsg.), *Liability for Artificial Intelligence and the Internet of Things*, 2019, S. 145, 162.

²⁶ Zur Verhinderung des moralischen Risikos bei der gesetzlichen Unfallversicherung vgl. *Kötz/Wagner*, *Deliktsrecht*, 13. Aufl. 2016, S. 239.

IV. Vorschläge

Zusammenfassend ergeben sich folgende grobe Handlungsempfehlungen: Kurzfristig kann die Judikative durch Anpassen von Verkehrspflichten und Beweislastregeln auf die Risiken neuartiger digitaler Systeme reagieren. Hier spielt auch die Produkthaftung eine große Rolle. Mittelfristig wäre zu überlegen, eine Gefährdungshaftung für bestimmte, klar zu definierende digitale Systeme als sinnvolles Instrument der Risikosteuerung einzuführen. Die dafür erforderliche Klassifikation digitaler Systeme bildet auch einen gedanklichen Grundstein, der von der Datenethikkommission so bezeichneten „Kritikalitätspyramide“²⁷. Langfristig wäre zu diskutieren, ob nicht in einer Welt unentrinnbarer digitaler Systeme eine „KI-Unfallversicherung“ eingeführt werden sollte.

²⁷ Gutachten der Datenethikkommission, 2019, https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_DE.pdf?__blob=publicationFile&v=2 (zuletzt aufgerufen am 10.12.2019), S. 177 ff. Zu einer Kategorisierung digitaler Systeme entsprechend den mit ihnen verbundenen Risiken *Lohmann*, ZRP 2017, 168, 169; *Borges*, in: Lohsse/Schulze/Staudenmayer (Hrsg.), Liability for Artificial Intelligence and the Internet of Things, 2019, S. 145, 153.