Professor Dr. Bijan Fateh-Moghadam*

Innovationsverantwortung im Strafrecht: Zwischen *strict liability*, Fahrlässigkeit und erlaubtem Risiko – Zugleich ein Beitrag zur Digitalisierung des Strafrechts

https://doi.org/10.1515/zstw-2019-0031

I. Strafrecht und Risiko *revisited*

„Strafrecht und Kriminalpolitik sind in dieser ‚Risikogesellschaft' in eine Krise geraten [...]. "[1] So lautete der Ausgangsbefund der klassischen Untersuchung zu „Strafrecht und Risiko" von *Prittwitz* aus dem Jahr 1993[2]. Im selben Jahr diskutierte die Basler Strafrechtslehrertagung über die „Zukunftssicherung mit den Mitteln des Strafrechts"[3]. Den gesellschaftlichen Hintergrund für das Interesse an strafrechtlicher Risikobewältigung bildeten vor 26 Jahren die modernen Großrisiken, die von der Atom- und Gentechnologie für die Umwelt und für die Menschheit ausgingen[4]. Heute sind es Entwicklungen in den Informations- und Biotechnologien, die soziologische Beobachter von einer „vierten Medienepoche", der „vierten industriellen Revolution" oder einem „zweiten Maschinenzeitalter" sprechen lassen[5]. Die damit

1 *Prittwitz*, Strafrecht und Risiko, 1993, S. 29.

2 In etwa zeitgleich erschienen die Dissertation von *Hilgendorf*, Strafrechtliche Produzentenhaftung, 1993, sowie bereits 1991 die Habilitationsschrift von *Herzog*, Gesellschaftliche Unsicherheit und strafrechtliche Daseinsvorsorge, 1991.

3 Vgl. die Beiträge in ZStW 105 (1993), S. 679, insbesondere das wirkmächtige Einleitungsreferat von *Stratenwerth*, a. a. O., S. 679, sowie den Tagungsbericht, a. a. O., S. 803.

4 *Ulrich Beck*, Risikogesellschaft, 1986.

5 *Baecker*, 4.0 oder Die Lücke die der Rechner lässt, 2018, S. 10; aus ökonomischer Perspektive *Schwab*, Die Vierte Industrielle Revolution, 2016; *Brynjolfsson/McAfee*, The Second Machine Age, 2016. Anders setzt *Nassehi*, Muster, 2019, S. 12, an, indem er in gesellschaftstheoretischer Perspek-

Hinweis: Überarbeitete und ergänzte Fassung des gleichnamigen Vortrags vom 31. Mai 2019, gehalten auf der 38. Strafrechtslehrertagung 2019: „Wie viel Fahrlässigkeit verträgt die Gesellschaft?", Leibniz Universität Hannover. Der Vortragsstil wurde weitgehend beibehalten.

***Kontaktperson: Bijan Fateh-Moghadam,** Inhaber des Lehrstuhls für Grundlagen des Rechts und Life Sciences-Recht an der Universität Basel.

einhergehenden Herausforderungen für die Strafrechtswissenschaft, und hier gerade für die Fahrlässigkeitsdogmatik, sind so grundlegend, dass es sich lohnt, die Frage der strafrechtlichen Verantwortung für technische Innovationen erneut aufzugreifen. Der vorliegende Beitrag nimmt dabei eine primär grundlagenorientierte Perspektive ein, die sich für den Zusammenhang von technologischem und strafrechtlichem Wandel interessiert. Damit ist die klassische rechtssoziologische Fragestellung nach der Wechselwirkung zwischen (nicht-rechtlichen) gesellschaftlichen Innovationen und Innovationen im Recht angesprochen, die heute den Gegenstand interdisziplinärer Rechtsforschung bildet[6].

Der strafrechtliche Fahrlässigkeitsbegriff verweist auch und gerade auf die Risiken technischer Innovationen zurück[7]. Er bildet gleichsam die Linse, durch die das Strafrecht Veränderungen in der Risikoqualität technischer Innovationen beobachtet. Die von *Prittwitz* so bezeichnete „Risikodogmatik"[8] wird zum strafrechtlichen Ausdruck einer modernen Gesellschaft, die sich, in der Lesart des Soziologen *Rosa*, nur noch dynamisch stabilisieren kann und „systematisch auf Wachstum, Innovationsverdichtung und Beschleunigung" angewiesen ist[9]. Unter den Bedingungen der „beschleunigten Risikogesellschaft" – ein Begriff, der die Zeitdiagnosen von *Ulrich Beck* und *Rosa* zusammenliest[10] – kann sich auch die Rechtsordnung nur stabilisieren, wenn sie in der Lage ist, sich auf Innovationsrisiken laufend neu einzustellen[11]. Der Innovationsdruck, der von neuen Technologien auf das Recht ausgeht, soll hier über den rechtssoziologischen Begriff der „transformativen Technologie" erfasst werden (II). Als exemplarischer Anwendungsfall einer transformativen Technologie dient nachfolgend die Digitalisierung und hier insbesondere der Bereich der sogenannten künstlichen Intelli-

---

tive danach fragt, für welches (alte) gesellschaftliche Problem die Digitalisierung eine Lösung darstellt. Die digitale Beobachtung sei, hier treffe sie sich mit der empirischen Sozialforschung, weniger an konkreten Individuen interessiert als an der Entdeckung von Mustern in den Daten (a. a. O., S. 58 f.). So gesehen teilen sich die Soziologie und die Digitalisierung das Bezugsproblem, nämlich die Reduzierung von Komplexität im Wege des Sichtbarmachens von Regelmäßigkeiten in der Gesellschaft (a. a. O., S. 28).

6 Grundlegend *Hoffmann-Riem*, Innovation und Recht – Recht und Innovation, 2016; zur Öffnung der Rechtssoziologie hin zur interdisziplinären Rechtsforschung vgl. *Baer*, Rechtssoziologie, 3. Aufl., 2017, S. 51 ff., *Rosenstock*/*Singelnstein*/*Boulanger*, in: *Boulanger*/*Rosenstock*/*Singelnstein* (Hrsg.), Interdisziplinäre Rechtsforschung, 2019, S. 3 ff., sowie *Raiser*, Grundlagen der Rechtssoziologie, 6. Aufl., 2013, S. 4.

7 Siehe dazu etwa auch den Beitrag von *Susanne Beck* in diesem Heft, S. 967.

8 *Prittwitz* (Anm. 1), S. 262 ff.

9 *Rosa*, Resonanz, 2016, S. 673.

10 *Ulrich Beck* (Anm. 4); *Rosa*, Beschleunigung, 2005.

11 *Rosa* (Anm. 9), S. 685; vgl. auch *Baecker* (Anm. 5), S. 186 ff.

genz (KI) (III). Im abschließenden vierten Teil werden schließlich die möglichen konkreten Konsequenzen der Digitalisierung für die strafrechtliche Risikodogmatik diskutiert.

II. Transformative Technologien

Die Risikodogmatik erlebt derzeit eine Renaissance, weil die Informationstechnologie beginnt, ihr transformatives Potenzial zu entfalten. Transformative Technologien sind gemäß der hier vorgeschlagenen rechtssoziologischen Begriffsverwendung dadurch gekennzeichnet, dass sie nicht nur *disruptiv* wirken, sondern zugleich in der Lage sind, die ethischen und rechtlichen Rahmenbedingungen des Technikeinsatzes zu verändern[12]. Der hiervon abzugrenzende verbreitete ökonomische Begriff der *Disruption* zielt auf das marktverdrängende Potenzial einer Technologie und betrifft damit das Verhältnis unterschiedlicher, konkurrierender Technologien zueinander. Ein einfaches Beispiel für die disruptiven Effekte digitaler Technologien bildet die weitgehende Verdrängung der Vinylschallplatte durch die Compact Disc als optisches Speichermedium für digitale Musik, welche seit den 2000er Jahren ihrerseits durch die MP3-Technologie und die Verbreitung von Musik über online Musikstreamingdienste abgelöst wurde. Die in der Literatur gut belegten Gründe für die disruptiven Effekte der modernen Informationstechnologie sind zusammengefasst:

1. Eine über Jahrzehnte exponentiell steigende Leistungsfähigkeit von Computern, für die das sogenannte *Moore*sche Gesetz steht[13];
2. die Verfüg- und Verarbeitbarkeit großer Datenmengen, die es ermöglicht, über den Einsatz von informationstechnologischen Ansätzen in diesen Daten Muster zu erkennen und daraus neue Einsichten zu gewinnen (*Big Data*)[14];

12 Vgl. *Nuffield Council on Bioethics*, Emerging Biotechnologies, London 2012, S. 40 ff., abrufbar unter: http://nuffieldbioethics.org/wp-content/uploads/2014/07/Emerging_biotechnologies_full_report_web_0.pdf (Stand: 13.09.2019), sowie bereits *Fateh-Moghadam*, BJM 2018, 205, 209 ff.

13 Danach verdoppelt sich, vereinfacht formuliert, alle 18 Monate die allgemeine Rechnerleistung bei gleichbleibenden oder sinkenden Komponentenkosten, vgl. dazu ausführlich *Brynjolfsson/McAfee* (Anm. 5), S. 53 ff.

14 Zum nicht einheitlich verwendeten Begriff „Big Data" instruktiv *Deutscher Ethikrat* (Hrsg.), Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung, 2018, S. 54, abrufbar unter: https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf (Stand: 13.09.2019). Die vollständige Arbeitsdefinition des Deutschen Ethikrats lautet: „Big Data ist der Umgang mit großen Datenmengen, der darauf abzielt, Muster zu erkennen und daraus neue Einsichten zu gewinnen, und der hierzu angesichts der Fülle und Vielfalt der Daten sowie der Geschwindigkeit, mit der sie erfasst, analysiert und neu

3. die zunehmende Vernetzung informationstechnischer Systeme im sogenannten *Internet of Things* (*IoT*)[15]; sowie
4. die Erfolge bei der Automatisierung und Autonomisierung informationstechnisch gesteuerter Systeme wie Roboter, selbstfahrende Fahrzeuge und Software-Agenten[16]. Die erstaunlichen Fortschritte der künstlichen Intelligenz etwa im Bereich der Bilderkennung oder bei der Bewältigung von komplexen Strategiespielen wie *Go* gehen auf die Umstellung von *modellbasierter* künstlicher Intelligenz zu *funktionsbasierten*, induktiv vorgehenden Ansätzen des Maschinenlernens etwa in Form sogenannter neuronaler Netze zurück[17].

In der Zusammenschau dieser Entwicklungen wird deutlich, dass sich die datengetriebene Informationstechnologie zu einem sämtliche gesellschaftliche Lebensbereiche durchdringenden „sozialen Tatbestand" (*Durkheim*) entwickelt hat[18]. Der Rückgriff auf den klassischen soziologischen Begriff des sozialen Tatbestandes soll den kollektiven und strukturellen Charakter der digitalen Transformation der Gesellschaft betonen. Wenn uns die Produkte der digitalen Technologie heute als zweite Natur gegenübertreten, so lässt sich diese emergente Realität nicht mehr auf die Summe individuell motivierter Handlungen von Personen reduzieren. Als eine allgemein auftretende Art des Handelns, Denkens und Fühlens, die „ein von ihren individuellen Äußerungen unabhängiges Eigenleben besitzt"[19], wirkt die Digitalisierung von außen und mit zwingendem Charakter auf die Individuen ein. Die mit neunstelligen Investitionen geförderte Trias von Automatisierung (Robotik), Autonomisierung (KI) und Vernetzung (IoT/Big Data) stellt unseren Weltzugang schrittweise auf einen „*Onlife*"[20]-Modus um, in dem digitale und analoge Realitäten untrennbar miteinander verflochten sind[21]. Dies bedeutet zugleich, dass der Einzelne nicht mehr selbst bestimmen kann, ob er on- oder offline Leben möchte. Es wird zunehmend schwierig, sich aktiv der digitalen Transformation zu

verknüpft werden, innovative, kontinuierlich weiterentwickelte informationstechnologische Ansätze nutzt."

15 Vgl. dazu *Zech*, ZfPW 2019, 198, 203.

16 *Gless*/*Janal*, JR 2016, 561; *Wohlers*, BJM 2016, 113; vgl. ferner die Beiträge in *Gless*/*Seelmann* (Hrsg.), Intelligente Agenten und das Recht, 2016, sowie in *Hilgendorf* (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, und in *Susanne Beck* (Hrsg.), Jenseits von Mensch und Maschine, 2012.

17 Zu modellbasierten und funktionsbasierten Ansätzen in der KI-Forschung instruktiv *Darwiche*, Communications of the ACM 61 (2018), S. 56. Der Verfasser dankt Malte Helmert, Basel, für den Hinweis auf diese wichtige Unterscheidung in den Computerwissenschaften.

18 *Durkheim*, Regeln der soziologischen Methode, 1980, S. 114.

19 Siehe Anm. 18.

20 *Hildebrandt*, Smart Technologies and the End(s) of Law, Cheltenham/Northampton 2015, S. 41.

21 *Hildebrandt* (Anm. 20), S. 41 f.

entziehen. Wer etwa, wie der Bundesvorsitzende der Grünen, *Habeck*, seine Social Media Accounts (teilweise) löscht, weil er diesen einen negativen Einfluss auf den eigenen politischen Kommunikationsstil attestiert[22], beweist zwar eine bemerkenswerte Fähigkeit zur Selbstreflexion, definiert sich damit aber umso mehr über sein Verhältnis zu digitalen Medien. Eine vollständige digitale Abstinenz ist einer gesellschaftlich auch nur minimal integrierten Person ohnehin nicht mehr möglich[23].

Die quantitative Leistungssteigerung und ubiquitäre gesellschaftliche Verbreitung der digitalen Informationstechnologie hat damit den Punkt erreicht, an dem sie qualitativ zu einer transformativen Technologie umschlägt. Der Transformationsbegriff zielt gemäß der hier vorgeschlagenen rechtssoziologischen Lesart nicht auf die Relation zwischen konkurrierenden Technologien, sondern auf die voraussetzungsreichere Annahme einer Einwirkung der Technologie auf die normative Verfasstheit einer Gesellschaft, namentlich auf Ethik und Recht. Der damit behauptete Zusammenhang zwischen faktischen, wissenschaftlichen und technologischen Entwicklungen mit bestimmten normativen Dynamiken steht in einem Spannungsverhältnis zu der in der Rechts- und Moralphilosophie verbreiteten Überzeugung, dass man nicht (unmittelbar) von einem Sein auf ein Sollen schließen dürfe[24]. Und in der Tat: Auch die „Digitalisierung" trägt ihre Normativität nicht in sich; sie zwingt nicht unmittelbar zu spezifischen ethischen und rechtlichen Konsequenzen. Die mit der Digitalisierung verbundenen Transformationen im Recht können daher nicht im Sinne einer direkten kausalen Steuerung oder Determinierung rechtlicher Inhalte durch Technik verstanden werden. Gleichwohl lässt sich sowohl aus rechtshistorischer Perspektive als auch mit Blick auf die gegenwärtige Entwicklung zeigen, dass es bestimmte Mechanismen gibt, die die Co-Evolution von Technik und Recht miteinander koppeln. Im Folgenden werden vier Konzeptionen vorgestellt, die in der Lage sind, die digitale Transformati-

22 Seinen Rückzug von Twitter und Facebook gab *Habeck* auf seinem Blog – und damit ebenfalls digital vermittelt – bekannt, abrufbar unter: https://www.robert-habeck.de/texte/blog/bye-bye-twitter-und-facebook/ (Stand: 04.10.2019); zu den „asozialen" Effekten sozialer Medien vgl. *Lanier*, Zehn Gründe, warum Du deine Social Media Accounts sofort löschen musst, 2018, der freilich seinerseits auf digitale Alternativstrategien verweist.

23 Daher wird der Ausschluss aus der digitalen Welt durch algorithmisch determinierte, undurchsichtige elektronische Zugangssperren zum Ausdruck einer radikalen Krisen- und Entfremdungserfahrung im digitalen Zeitalter, literarisch verarbeitet bei *Schönthaler*, Der Weg aller Wellen, 2019.

24 Diese Kritik geht zurück auf *Hume*, A Treatise of Human Nature, Oxford 1955, und *Moore*, Principia Ethica, 1996; der Sein-Sollen-Fehlschluss ist eng verwandt mit dem sogenannten „naturalistischen Fehlschluss", bei dem tatsächliche Eigenschaften der Natur automatisch als „gut" bewertet werden.

on des Strafrechts zu erklären, ohne dabei die Autonomie der Begriffsbildung im Strafrecht in Frage zu stellen.

III. Vier Konzeptionen der digitalen Transformation des Strafrechts

1. *Reframing* des Selbstbildes des Menschen

Zunächst kennen wir aus der Ideengeschichte wissenschaftliche und technologische Revolutionen, die so grundlegend sind, dass sie das Bild verändern, das sich der Mensch von sich selbst und seinem Verhältnis zu Natur und Technik macht[25]. Als transformative wissenschaftliche Erkenntnis *avant la lettre* gilt insoweit die kopernikanische Wende vom geo- zum heliozentrischen Weltbild. *Blumenberg* beschreibt in seiner „Geistesgeschichte der Technik", wie die damit verbundene „Selbsteinsetzung des Menschen in die Mitte seiner Welt" eine tatsächliche Voraussetzung für die moderne Konstruktion des selbstbestimmten Individuums bildet[26]. Technik wird in dieser Lesart als „pure Konstruktion der freien Geistestätigkeit"[27] zum „Phänomen der zur unbedingten Macht entschlossenen Subjektivität"[28] des Menschen. Bemerkenswerterweise ist es nun die vom Menschen hervorgebrachte „dunkle Macht der Algorithmen", die sich verselbständigt und dem Menschen als „zweite Natur" gegenübertritt[29].

Damit verändert sich der Bezugsrahmen, das *Framing*, in dem der Mensch sein Verhältnis zu Natur und Technik deutet. Rahmen oder *Frames* sind in der Lesart von *Goffman* Heuristiken für die Interpretation von Ereignissen, über die wir unsere Alltagserfahrungen organisieren[30]. Wenn wir etwas beobachten, ordnen wir es – in der Regel unbewusst – entweder als Naturerscheinung ein oder als das Ergebnis von zweckgerichteten menschlichen Handlungen. Ein Ereignis in

25 Vgl. hierzu bereits *Fateh-Moghadam*, BJM 2018, 205, 210 f.

26 *Blumenberg*, Schriften zur Technik, 2015, S. 69.

27 *Blumenberg* (Anm. 26), S. 78.

28 Siehe Anm. 27.

29 *Harari*, Homo Deus, 2017, S. 97 ff.; *Nordmann*, Technikphilosophie zur Einführung, 2. Aufl. 2015, S. 84 ff., am Beispiel der „Krebsmaus"; auch *Blumenberg* (Anm. 26), S. 78, hatte bereits aufgezeigt, dass sich im weiteren Verlauf der Geschichte die Produkte der Technik verselbständigen und dem Menschen als eine „zweite Natur" gegenübertreten, die wiederum nicht für den Menschen gemacht zu sein scheint; *Bridle*, New Dark Age, 2019, S. 58, „Je obsessiver wir versuchen, die Welt zu berechnen, desto komplexer und unbegreiflicher erscheint sie."

30 *Goffman*, Rahmen-Analyse, 2018, S. 19 und 31 ff.

einem natürlichen Rahmen zu deuten bedeutet in den Worten von *Goffman*, und hier wird die strafrechtliche Relevanz sichtbar, „man führt sie vollständig, von Anfang bis Ende, auf ‚natürliche' Ursachen zurück. Man sieht keinen Willen, keine Absicht als Ursache am Werke, keinen Handelnden, der ständig auf das Ergebnis Einfluß nimmt."[31]

Die negativen Folgen eines solchen Naturereignisses erscheinen dann als „Unglück", die Qualifizierung als „Unrecht" macht dagegen nur in einem sozialen Deutungsrahmen Sinn[32]. Genau diese Unterscheidung wird durch die zunehmende Bevölkerung unserer Lebenswelt mit selbstlernenden Robotern und Software-Agenten unterlaufen, deren sozialer, moralischer und rechtlicher Status umstritten ist. Paradigmatisch für die Verflüssigung der Grenzziehung zwischen Person, Tier und Maschine steht der jüngst von 23 international führenden Computerwissenschaftlern veröffentlichte Aufruf zur Gründung einer neuen akademischen Disziplin der „maschinellen Verhaltensforschung" (*Machine Behaviour*)[33]. Die bisher auf Menschen und Tiere beschränkte Verhaltensforschung müsse auf Maschinen ausgeweitet werden, weil komplexe (lernende) KI-Agenten nicht länger allein aus ihrem internen Bauplan heraus verstanden werden könnten, sondern nur im Zusammenspiel von Maschine und Umwelt[34]. Die Diskussion über *Machine Behaviour* belegt eindrucksvoll, dass die moderne KI-Technologie das Potenzial besitzt, scheinbar verfestigte gesellschaftliche Konzeptionen zu verflüssigen und zu transformieren[35].

2. Technik als Medium des Rechts (*Legal Technology*)

Eine zweite Möglichkeit, wie technologische Entwicklungen das Strafrecht beeinflussen können, ist die Veränderung der Medien des Rechts. Die Medientheorie des Rechts hat gezeigt, dass die Erfindung der Lautsprache, der Schrift und des Buchdrucks jeweils auch rechtsbildende Konsequenzen hatte, indem etwa neue Möglichkeiten der Perpetuierung, Archivierung und Verbreitung von Recht ermöglicht wurden[36]. Dass sich auch die 4. Medienrevolution durch Computernetz-

31 *Goffman* (Anm. 30), S. 31.
32 Zur Dichotomie von Unglück und Unrecht vgl. bereits *Prittwitz* (Anm. 1), S. 378 ff.
33 *Rahwan et al.*, nature 568 (2019), S. 477.
34 *Rahwan et al.*, nature 568 (2019), S. 477, 480.
35 Vgl. ferner die Beiträge in *Brockman* (Hrsg.), Possible Minds, New York 2019.
36 *Vesting*, Die Medien des Rechts: Schrift, 2011; *ders.*, Die Medien des Rechts: Sprache, 2011; *ders.*, Die Medien des Rechts: Buchdruck, 2013; *ders.*, Die Medien des Rechts: Computernetzwerke, 2015; *Vismann*, Medien der Rechtsprechung, 2011; *Baecker* (Anm. 5), S. 10.

werke und Informationstechnologie bereits auf die Praxis des Straf- und Strafprozessrechts auswirkt, steht außer Frage. Damit ist der expandierende Bereich der *Legal Technology* angesprochen, der von relativ profaner Bürosoftware für Strafverteidigerinnen über die Europäische E-Evidence-Verordnung bis hin zu so problematischen Anwendungen wie algorithmengestützte Rückfallprognosen im Kontext der Strafjustiz durch Software-Lösungen wie COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) in den USA reicht[37]. Die hiermit verbundenen normativen Problemstellungen bedürfen empirisch informierter Einzelfallanalysen und können an dieser Stelle nicht weiter vertieft werden. Ein übergreifendes Risiko des Einsatzes von *Legal Technology* zur Erledigung klassischer juristischer Aufgaben besteht darin, dass nicht die Technologien an die juristischen Methoden, sondern die juristischen Methoden an die Funktionsweise der Technik angepasst werden. Es droht mit anderen Worten eine „Computerisierung“[38] rechtswissenschaftlicher Methoden. Während klassische juristische Argumentation qualitativ um gute Gründe für eine Entscheidung ringt, sind computergestützte Entscheidungsverfahren auf die Quantifizierbarkeit von Entscheidungsparametern angewiesen. Der gerechtigkeitstheoretische Gesichtspunkt der Billigkeit etwa, der gerade eine Abweichung von der strikten Regelanwendung im Einzelfall verlangen kann, dürfte sich über Computerprogramme kaum abbilden lassen[39].

3. Technik als Recht: „Normative Technosteuerung“?

Ein dritter Modus der digitalen Transformation besteht darin, dass Technik die verhaltenssteuernde Funktion des Rechts selbst übernimmt: Technologie ersetzt und verdrängt Recht (*code is law*)[40]. Sofern es nicht gelingt, die „Verhaltenssteue-

37 Einen Überblick zum Stand der *Legal Technology* gibt *Fries*, RW 9 (2018), S. 414; zur Automatisierung der öffentlichen Verwaltung vgl. *Braun Binder*, SJZ 115 (2019), S. 567; zur Kritik der E-Evidence-Verordnung vgl. *Burchard*, ZRP 2019, 164; zu COMPAS vgl. *ders.*, Künstliche Intelligenz als Ende des Strafrechts?, 15, abrufbar unter: https://www.normativeorders.net/de/publikationen/working-paper (Stand: 13.09.2019), sowie *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, S. 55 ff.

38 *Bridle* (Anm. 29), S. 27 ff.

39 Der Begriff geht auf *Aristoteles*, Nikomachische Ethik, 1956, Buch V Kap. 10 Rdn. 1137b, und seine Ausführungen zur *epieikeia* zurück; vgl. dazu *Bien*, Billigkeit, in: *Ritter* (Hrsg.), Historisches Wörterbuch der Philosophie, Bd. 1, 1971, S. 940.

40 Grundlegend *Lessig*, Code: Version 2.0, New York 2006, S. 1 ff., „code is law“; *ders.*, a. a. O., S. 81 ff., „regulation by code“; vgl. dazu *Hildebrandt* (Anm. 20), S. 165, sowie *Lobe*, Speichern und Strafen, 2019, S. 41 ff.

rung durch Algorithmen“[41] rechtlich einzuhegen, könnte dies auf einen fundamentalen Wandel der Struktur sozialer Kontrolle hinauslaufen. Rechtsnormen gelten gerade, weil sie übertreten werden können[42], und halten insofern stets alternative Möglichkeiten des Weltverlaufs offen[43]. Normative „Technosteuerung“[44] schließt bestimmte Verhaltensweisen dagegen faktisch aus: Das smarte Auto fährt nicht los, wenn die Insassen nicht angeschnallt sind, und es fährt nicht schneller als die jeweils zulässige Höchstgeschwindigkeit. Solche technisch vermittelten *impossibility structures*, die Rechtsverstöße schon rein physisch unmöglich machen sollen[45], führen dazu, dass Delinquenz allenfalls noch in Verbindung mit Computersabotage denkbar ist. Die praktische Einschränkung der Freiheit durch technisches Design reicht wegen ihres zwingenden Charakters weiter als die Formulierung von Verhaltenserwartungen durch Rechtsnormen[46]. Mit der faktischen Unterbindung abweichenden Verhaltens verhindern technische *impossibility structures* zugleich, dass Normverstöße unter bestimmten Voraussetzungen evolutionäre Anpassungsleistungen des Rechts ermöglichen – *Luhmann* hatte in diesem Zusammenhang von „brauchbarer Illegalität“[47] gesprochen. Der totalitäre Charakter einer drohenden umfassenden technischen Verhaltenssteuerung produziert die verfassungsrechtliche Frage nach der Existenz eines „Rechts zum Rechtsverstoß“[48]. Ein solches Recht kann widerspruchsfrei nicht als Freiheit verstanden werden, die Strafrechtsordnung sanktionslos zu übertreten. Vielmehr kann es allenfalls um ein Recht gehen, weiterhin die faktische Möglichkeit zu haben, Rechtsnormen zu befolgen oder zu übertreten. Freiheitsphilosophisch gewendet bedeutete ein solches Recht auf die Möglichkeit zum Normbruch zugleich, dass dem Einzelnen die Freiheit belassen wird, „sich *für* das Recht zu entscheiden“[49]. Richtigerweise geht es auch nicht um ein Recht auf Vollzugsdefizi-

41 *Hoffmann-Riem*, AöR 142 (2017), S. 1.

42 *Luhmann*, Rechtssoziologie, 4. Aufl. 2008, S. 43.

43 *Möllers*, Die Möglichkeit der Normen, 2018, S. 159.

44 *Hoffmann-Riem*, AöR 142 (2017), S. 1, 11 ff.

45 Dazu *Rademacher*, JZ 2019, 702 mit weiteren Nachweisen.

46 *Brownsword*, in: *Brownsword/Scotford/Yeung* (Hrsg.), The Oxford Handbook of Law, Regulation and Technology, Oxford 2017, S. 42.

47 *Luhmann*, Funktionen und Folgen formaler Organisation, 1964, S. 304 ff.; vgl. dazu bereits *Hoffmann-Riem*, AöR 142 (2017), S. 1, 34.

48 *Rademacher*, JZ 2019, 702, 707.

49 *Rostalski*, GA 2019, 481, 485. Wenig überzeugend ist allerdings die orthodox-hegelianische Annahme, aus der Vernunftbegabung des Menschen folge notwendig die Einsicht in die Richtigkeit der rechtlichen Ge- und Verbote, weshalb der Normbruch schon begrifflich kein Ausdruck wirklicher Freiheit vernunftbegabter Personen sein könne. Die mit dem Deutschen Idealismus nicht vereinbare „Freiheit zur Unvernunft“ und damit auch zum Rechtsverstoß ist eine Errungenschaft

te[50], da tatsächliche Normbefolgung und Normvollzug nicht identisch sind. Der Vollzug von Strafgesetzen, die immer schon mit delinquentem Verhalten rechnen, besteht in der Verfolgung und Ahndung von Normverstößen, nicht aber in deren Beachtung durch die Normadressaten. Selbst dann, wenn die Übertretung eines strafrechtlichen Verbots nachträglich lückenlos sanktioniert würde, bliebe abweichendes Verhalten möglich[51]. Die Radikalität von technischen *impossibility structures* besteht darin, dass sie diese Möglichkeit faktisch beseitigt – und nicht in einem lückenlosen Normvollzug. Aus strafrechtswissenschaftlicher Perspektive handelt es sich bei normativer Technosteuerung, die nicht auf Rechtsnormen, sondern auf – in der Regel privat kontrollierten – Computercodes beruht, streng genommen nicht mehr um eine Erscheinungsform strafrechtlicher Risikodogmatik, sondern vielmehr um deren Überwindung. Wenn man so will: Die ironische Antwort der digitalen Präventionsgesellschaft auf die liberale Forderung nach einer Abschaffung des Strafrechts[52].

4. Veränderung des „Realbereichs der Norm“

Die vierte Konzeption der digitalen Transformation des Rechts, schließlich, betrifft die Veränderung des „Realbereichs der Norm(en)“[53]. Der von *Hoffmann-Riem* geprägte Begriff des Realbereichs der Norm geht davon aus, dass Rechtsnormen nicht nur einen Sprach- oder Textbereich aufweisen, sondern zugleich auf einen bestimmten Ausschnitt von Realität bezogen sind[54]: „Dieser Ausschnitt betrifft die von der Norm in Bezug genommene technologische, naturwissenschaftliche, soziale, politische, ökonomische, kulturelle, ökologische u. a. ‚Wirklichkeit‘ in ihren Grundstrukturen.“[55] Der Realbereich gilt als ein konstitutiver Bestandteil der Norm, da er ihre sinnvolle Auslegung erst möglich macht[56]. Eine Rechtsnorm auszulegen bedeutet, unter anderem danach zu fragen, welches lebensweltliche Problem mit ihr bewältigt werden soll. Sie lässt sich nur dann sinnvoll verstehen,

eines modernen, grundrechtlichen Verständnisses der Selbstbestimmung des Menschen, die es zu verteidigen gilt, dazu *Fateh-Moghadam*, BJM 2018, 205, 215 ff.

50 So aber *Rademacher*, JZ 2019, 702, 708.

51 Zur Unterscheidung von lückenloser Straftatprävention und lückenloser Straftatenahndung vgl. auch *Rostalski*, GA 2019, 481, 483 ff. und 485 ff.

52 *Lüderssen*, Abschaffen des Strafrechts?, 1995.

53 *Hoffmann-Riem* (Anm. 6), S. 113.

54 *Hoffmann-Riem* (Anm. 6), S. 113 ff.

55 Siehe Anm. 53.

56 *Hoffmann-Riem* (Anm. 6), S. 114.

wenn man berücksichtigt, welchen Lebenssachverhalt der Gesetzgeber als unter die Norm zu subsumierenden „Normalfall“[57] vor Augen hatte. Die Auslegung der Norm muss mithin, wie man aus einer hermeneutischen Perspektive formulieren würde, die „Vorverständnisse“ des Gesetzgebers berücksichtigen[58], zumindest soweit sie sich in der Formulierung der Norm niedergeschlagen haben. Diese Vorverständnisse speisen sich nicht nur aus normativen Wertungen, sondern auch und gerade aus einer Vorstellung über die gesellschaftliche Wirklichkeit, die den Gegenstand der Regulierung bildet. Technische Innovationen sind daher in der Lage, mit den realen technologischen Voraussetzungen, auf die die Rechtsnorm bezogen ist, zugleich den durch Auslegung zu gewinnenden Normgehalt selbst zu verändern. Weiterhin kann die technologische Innovation unter Geltung des strafrechtlichen Analogieverbots dazu führen, dass bestimmte unerwünschte Erfolge oder Handlungen deshalb nicht mehr vom Anwendungsbereich einer Verbotsnorm erfasst werden, weil sich das neue technologische Verfahren nicht mehr unter den Wortlaut des gesetzlichen Tatbestandes subsumieren lässt; ein Vorgang, der aus dem Strafrecht der Biomedizin gut bekannt ist[59]. Die technologische Innovation erzeugt in diesen Fällen bei normativer Betrachtung eine Regelungs- bzw. Strafbarkeitslücke, die den Anlass für gesetzgeberische Reformen bilden kann.

In diesem Sinne lassen sich auch die ersten transformativen Effekte der digitalen Informations- und Computertechnologie im Strafrecht deuten. Aufgrund der strengen Gesetzesbindung sah sich der deutsche Strafgesetzgeber bereits frühzeitig gezwungen, auf Phänomene der computervermittelten Kriminalität mit neuen Spezialstraftatbeständen zu reagieren. So wurden schon 1986 mit dem 2. Gesetz zur Bekämpfung der Wirtschaftskriminalität (2. WiKG)[60] der Computerbetrug (§ 263a StGB) und das Ausspähen von Daten (§ 202a StGB) in das Strafgesetzbuch aufgenommen[61]. Bereits 1977 war *Siebers* Pionierarbeit zu „Computerkriminalität und Strafrecht“ erschienen[62]. Die neue Technologie, so schien es,

57 Dazu *Haft*, Einführung in das juristische Lernen, 5. Aufl. 1991, S. 113 ff.

58 Zur Rehabilitierung des „Vorurteils“ als universelle Bedingung des Verstehens vgl. *Gadamer*, Wahrheit und Methode, 1990, S. 281 ff.

59 Zur Bedeutung des strafrechtlichen Analogieverbots für die Auslegung des Embryonenschutzgesetzes gerade im Zusammenhang mit naturwissenschaftlichen Innovationen vgl. *Günther*, in: Embryonenschutzgesetz, 2014, Vor. § 1 Rdn. 10 ff.

60 BGBl. I (1986), S. 721.

61 Das 2. WiKG gilt daher als Geburtsstunde des Strafrechts der Informations- und Kommunikationstechnik, *Kochheim*, Cybercrime und Strafrecht in der Informations- und Kommunikationstechnik, 2. Aufl. 2018, S. 133 (dort auch zur weiteren Entwicklung der Gesetzgebung auf diesem Gebiet).

62 *Sieber*, Computerkriminalität und Strafrecht, 1977.

würde man durch die Schaffung eines neuen strafrechtlichen Spezialgebiets, dem Computer- und Internetstrafrecht[63], in den Griff bekommen. Die reizvolle Vorstellung, die strafrechtlichen Folgeprobleme der Digitaltechnologie könnten den informatikinteressierten „Computernerds" der Strafrechtswissenschaft überlassen werden, während die klassischen Gebiete des Strafrechts davon unberührt bleiben würden, erscheint aus heutiger Sicht naiv. Die Digitalisierung der Gesellschaft stellt die gesamten Strafrechtswissenschaften vor neue Herausforderungen. Unabhängig davon, ob man sich für den Allgemeinen Teil des Strafrechts, das Wirtschaft- oder Medizinstrafrecht, das Internationale Strafrecht oder für das Strafverfahrensrecht interessiert, ist man mit Folgeproblemen der Digitalisierung konfrontiert. Die Kriminologie, die Rechtssoziologie, die Rechtsphilosophie und die Rechtsgeschichte, um das Bild komplett zu machen, zählen ohnehin zu den Pionieren der rechtswissenschaftlichen Reflexion der Digitalisierung[64]. Eine Strafrechtswissenschaft auf der Höhe der Zeit muss die Digitalisierung *des* Strafrechts daher im Sinne beider Bedeutungen des Genitivs verstehen[65]: Einerseits als aktive strafrechtliche Beobachtung der Folgeprobleme der Digitalisierung gesellschaftlicher Lebenszusammenhänge, die nicht auf klassische Computer- und Internetkriminalität beschränkt sind (Strafrecht als aktive Regulierung der Digitalisierung sämtlicher Lebensbereiche), andererseits im Sinne einer technologischen Entwicklung, durch die das Strafrecht selbst transformiert wird (Strafrecht als passiver Gegenstand der Digitalisierung). Die damit verbundenen Aufgaben lassen sich nicht allein an das freilich expandierende und immer bedeutsamer werdende strafrechtliche Spezialgebiet des Computer- und Internetstrafrechts delegieren. Umso notwendiger erscheint eine grundlagenorientierte, interdisziplinär informierte Reflexion des Prozesses der Digitalisierung des Strafrechts, die zwar von je konkreten Problemstellungen ausgeht, diese aber zugleich abstrahiert und übergreifende theoretische Werkzeuge zur rechtswissenschaftlichen Deutung der Co-Evolution von Technik und Strafrecht bereitstellt[66].

63 Einen Gesamtüberblick über das Strafrechtsgebiet gibt *Kochheim* (Anm. 61).

64 Vgl. nur die Beiträge in *Hilgendorf* (Anm. 16), sowie in *Gruber*/*Bung*/*Ziemann* (Hrsg.), Autonome Automaten, 2015, und in *Gless*/*Seelmann* (Anm. 16).

65 Zu dieser Doppelperspektive auf die Digitalisierung *der* Gesellschaft vgl. allgemein *Baecker* (Anm. 5), S. 9.

66 Mit ähnlicher Stoßrichtung *Burchard*, Künstliche Intelligenz als Ende des Strafrechts?, insb. 24 ff., abrufbar unter: https://www.normativeorders.net/de/publikationen/working-paper (Stand: 13.09.2019).

Der erweiterte Einwirkungsradius der Digitalisierung lässt sich nicht zuletzt am Beispiel klassischer strafrechtlicher Zurechnungskategorien wie derjenigen der Fahrlässigkeit demonstrieren. Gerade das zuletzt beschriebene rechtssoziologische Konzept der Veränderung des „Realbereichs der Norm" ist geeignet, den Einfluss der Digitalisierung auf die strafrechtliche Fahrlässigkeitsdogmatik angemessen zu deuten. Die lebenswirklichen Effekte der Digitaltechnologie bestehen zumindest auch darin, dass neuartige Risiken für Rechtsgüter hervorgebracht werden, welche die Frage nach damit korrespondierenden strafrechtlichen Sorgfaltspflichten provozieren. Auf die strafrechtliche Risikodogmatik angewendet bedeutet dies, dass etwa die neue Risikoqualität von KI-Systemen, namentlich das *Autonomierisiko* bei maschinellem Lernen, einen Wandel der Auslegung und Anwendung strafrechtlicher Fahrlässigkeitsdelikte, wenn nicht gar die Schaffung innovativer Formen strafrechtlicher *strict liability*-Haftung bedingen könnte. Die damit angesprochene Dynamik der Risikodogmatik der künstlichen Intelligenz bildet den Gegenstand des letzten Teils dieses Beitrags.

IV. Vier Grundmodelle strafrechtlicher Innovationsverantwortung (Risikodogmatik der KI)

Die spezifische Risikoqualität von KI-Systemen ergibt sich aus dem Autonomierisiko, dem Verbundrisiko und dem Vernetzungsrisiko[67]. Das Verbundrisiko betrifft das Zusammenwirken von Mensch und Maschine in sogenannten sozio-technischen Systemen[68]. Das Vernetzungsrisiko entsteht durch die gleichzeitige Beteiligung mehrerer Computersysteme an einem Schadensereignis[69]. Das Autonomierisiko, auf das ich mich im Folgenden beschränken werde, bezieht sich auf die Fähigkeit selbstlernender Systeme, in Reaktion auf „Eindrücke aus der Außenwelt selbsttätig, das heißt ohne Einwirkung des Anwenders, ihr Verhalten zu ändern"[70] und insofern autonom zu agieren. Indem sie sich „auf eigene Wahrneh-

67 *Teubner*, AcP 218 (2018), S. 155, 164.

68 Zum Begriff „sozio-technisches System" vgl. *Ropohl*, Allgemeine Technologie, 3. Aufl. 2009, S. 58 f.; zu den strafrechtlichen Folgeproblemen vgl. *Simmler*, in: *Bendel* (Hrsg.), Handbuch Maschinenethik, 2019, S. 8 ff.; sowie am Beispiel des Medizinrechts *Steil et al.*, Methods Inf Med 58 (2019), S. 14, 19 ff.

69 *Zech*, ZfPW 2019, 198, 205.

70 *Zech*, ZfPW 2019, 198, 200; zu den technischen Grundlagen des Maschinenlernens vgl. *Norvig*, Artificial Intelligence, 3. Aufl., Boston 2016, S. 693 ff.

mungen“ verlassen, statt auf Eingaben des Anwenders angewiesen zu sein[71], gelingt es ihnen, selbständig „Entscheidungen unter Ungewissheit“[72] zu treffen. Die zunehmende Autonomie von selbstlernenden KI-Systemen führt zu Einschränkungen bei der Vorhersehbarkeit ihres Verhaltens und wirft die Frage auf, welche spezifischen Sorgfaltspflichten von Herstellerinnen, Programmiererinnen und Betreiberinnen gerade aus der „vorhersehbaren Unvorhersehbarkeit“[73] folgen. Die eingeschränkte wissenschaftliche Aufklärbarkeit der Entscheidungsgenese von KI-Systemen führt zudem zu Beweisproblemen in Bezug auf den Pflichtwidrigkeitszusammenhang. Vor diesem Hintergrund droht nach Ansicht zahlreicher Autoren eine strafrechtliche Verantwortungslücke (*responsibility gap*), die den strafrechtlichen Schuldgrundsatz unter Druck setzt[74]. So werden einerseits Durchbrechungen des Schuldgrundsatzes im Hinblick auf die Verantwortung von Herstellern und Betreibern diskutiert (*strict liability*)[75]. Andererseits geht es um die Ausweitung der Schuldfähigkeit von Menschen auf Roboter und Software-Agenten mit der Folge einer Verantwortungsverlagerung vom Menschen auf die Technik[76].

1. Verantwortungsverlagerung auf die Technik (Roboterstrafrecht)?

In die falsche Richtung weist m. E. die Diskussion über eine mögliche strafrechtliche Verantwortlichkeit von „intelligenten“ Software-Agenten[77]. Der maßgebliche Grund hierfür besteht indes nicht darin, dass (straf-)rechtliche Verantwortungsfähigkeit an spezifisch menschliche Qualitäten, innere psychologische Zu-

71 *Zech*, ZfPW 2019, 198, 200.

72 *Teubner*, AcP 218 (2018), S. 155, 174.

73 Vgl. hierzu *Gless*/*Weigend*, ZStW 126 (2014), S. 561, 581 f.

74 *Matthias*, Ethics and Information Technology 6 (2004), S. 175; *Gless*/*Weigend*, ZStW 126 (2014), S. 561, 582 („Verantwortungsdiffusion“); *Gless*, in: *Gless*/*Seelmann* (Anm. 16), S. 231; *Hilgendorf*, in: *Hornung* (Hrsg.), Rechtsfragen der Industrie 4.0, 2018, S. 120 ff.; *ders.*, in: Festschrift für Fischer, 2018, S. 110; *Simmler*, in: *Bendel* (Anm. 68), S. 4.

75 *Simmler*, Normstabilisierung und Schuldvorwurf, 2018; *Simmler*/*Markwalder*, ZStW 129 (2017), S. 20, 43.

76 Vgl. dazu bereits frühzeitig *Susanne Beck*, in: *Japanisch-Deutsches Zentrum* (Hrsg.), Mensch-Roboter-Interaktionen aus interkultureller Perspektive, 2012, S. 133 ff.; *Hilgendorf*, in: *Susanne Beck* (Anm. 16), S. 119; *Gless*/*Weigend*, ZStW 126 (2014), S. 561, 570 ff.; *Simmler*/*Markwalder*, ZStW 129 (2017), S. 20, 43; *Seher*, in: *Gless*/*Seelmann* (Anm. 16), S. 45; *Erhardt*/*Mona*, a. a. O., S. 61; *Gaede*, Künstliche Intelligenz – Rechte und Strafen für Roboter?, 2019.

77 Siehe Anm. 76.

stände oder ein philosophisch anspruchsvolles Verständnis von Willensfreiheit und Autonomie gebunden wäre[78]. Die verantwortliche Rechtsperson ist das Ergebnis eines sozialen, spezifisch rechtlichen Zuschreibungsprozesses, welcher einer funktionalen Logik folgt[79]. Sofern also „Personalität" als ein konstitutives Element des Schuldvorwurfs angesehen wird, so impliziert dies nicht notwendig eine Beschränkung auf menschliche Akteure[80]. Vielmehr ist es die rechtliche Zuschreibung, die die Person als mögliche Adressatin strafrechtlicher Schuldzuschreibung erst konstituiert[81]. Dies bedeutet andererseits nicht, dass die Zuschreibung des Status als verantwortliche Rechtsperson willkürlich erfolgen könnte; vielmehr müssen die Argumente für oder gegen die Anerkennung rechtlicher Verantwortlichkeit nichtmenschlicher Akteure nach Maßgabe der jeweiligen Funktionslogik des Rechtsgebiets beurteilt werden. Vor diesem Hintergrund mag es gute Gründe dafür geben, im Zivilrecht eine Teilrechtsfähigkeit von Software-Agenten einzuführen, wie dies *Teubner* vorgeschlagen hat[82]. Die zivilrechtliche Pointe besteht dabei darin, dass der menschlichen Betreiberin eines KI-Systems das schuldhafte Verhalten des Software-Agenten zugerechnet werden soll – so, wie dies für die Verrichtungsgehilfen (§ 831 BGB) anerkannt ist[83]. Auf das Strafrecht lässt sich diese Konstruktion indes nicht übertragen. Die Anerkennung der Verantwortungsfähigkeit des Roboters eignet sich im Strafrecht gerade nicht, um an die menschliche Akteurin hinter dem Roboter heranzukommen, weil das Strafrecht eine Haftung für allein fremdes Verschulden nicht zulässt. Vielmehr wäre das „eigenverantwortliche Dazwischentreten" des Roboters dazu geeignet, den menschlichen Täter hinter der Maschine zu entlasten[84].

Aber auch die unmittelbare strafrechtliche Verantwortung von Maschinen erscheint aus der Sicht des Strafrechts dysfunktional, da die Bestrafung eines KI-Systems nicht sinnvoll als normstabilisierende Reaktion auf den Bruch einer dem Rechtsgüterschutz dienenden Verhaltensnorm verstanden werden kann[85]. Denn

78 Zu diesen Einwänden ausführlich *Gless/Weigend*, ZStW 126 (2014), S. 561, 573 ff.

79 Grundlegend *Luhmann*, SozW 42 (1991), S. 166; für die Strafrechtswissenschaft *Jakobs*, Strafrecht Allgemeiner Teil, 2. Aufl. 1991, Abschn. 17 Rdn. 22; *Simmler/Markwalder,* ZStW 129 (2017), S. 20, 41 f.; *Simmler* (Anm. 75), S. 258 ff.

80 *Simmler* (Anm. 75), S. 258.

81 In diesem Sinne auch die am frühen *Marx* geschulte Einsicht von *Menke*, Kritik der Rechte, 2015, S. 17, wonach es „autonome Subjekte (gibt), weil es die moderne Form der Rechte gibt", und nicht etwa umgekehrt.

82 *Teubner*, AcP 218 (2018), S. 155.

83 Siehe Anm. 82.

84 *Gless/Weigend*, ZStW 126 (2014), S. 561, 588.

85 So überzeugend *Simmler/Markwalder*, ZStW 129 (2017), S. 20, 41, die freilich eine Roboterstrafbarkeit für die Zukunft nicht ausschließen wollen, sowie ausführlich *Simmler* (Anm. 75), S. 267 ff.;

dies würde die Fähigkeit des KI-Systems voraussetzen, normative Erwartungen entweder zu erfüllen oder zu enttäuschen. Selbst die modernsten selbstlernenden KI-Systeme sind in ihrem „Verhalten" nicht normativ motiviert. Sie treffen keine Entscheidungen als Folge der wertenden Abwägung von „guten Gründen", vielmehr basieren ihre Aktionen in der Außenwelt notwendig auf einer computervermittelten, auf Mustererkennung zielenden Auswertung derjenigen Daten, die ihnen entweder von Dritten zur Verfügung gestellt wurden oder die sie durch eigene Beobachtungen gewonnen haben. Ein selbstfahrendes Fahrzeug kann zwar darauf programmiert werden, eine Kollision mit menschlichen Verkehrsteilnehmern zu vermeiden – den normativen Sinn des Tötungs- oder Körperverletzungsverbots „versteht" es deshalb nicht. Kommt es aufgrund eines Fehlers des autonomen Fahrzeugs dennoch zu einem tödlichen Unfall[86], so wird dadurch möglicherweise unsere Vorstellung von der Leistungsfähigkeit der Technik enttäuscht. Dabei handelt es sich aber um eine *kognitive* Erwartung, die wir für die Zukunft korrigieren werden. Soweit wir zugleich an der Erwartung festhalten, dass selbstfahrende Fahrzeuge, die am Straßenverkehr teilnehmen, menschliche Hindernisse erkennen und rechtzeitig bremsen können *sollten*, so richtet sich diese – nunmehr *normative* – Erwartung an die Adresse der Menschen, die für die Konstruktion, Programmierung und Zulassung des Fahrzeugs zuständig sind[87].

Darüber hinaus wäre ein Roboterstrafrecht auch kriminalpolitisch dysfunktional, weil es die Menschen hinter der Maschine vorschnell von ihrer Verantwortung entlasten könnte. Zwar bliebe die fahrlässige Nebentäterschaft der Täterin hinter der Maschine weiterhin möglich. Entsprechend der Fahrlässigkeitsdogmatik für arbeitsteiliges Zusammenwirken müsste man dann aber von der Teilbarkeit der Verantwortungsbereiche und der Geltung des Vertrauensgrundsatzes ausgehen[88]. Damit droht eine Verantwortungsverlagerung vom Menschen auf die Maschine, welche die zentrale Rolle menschlicher Entscheidungen für das Funktio-

i.E. auch *Gless*/*Weigend*, ZStW 126 (2014), S. 561, 588 f., mit dem Hinweis auf die fehlende Eigenschaft des Roboters, ein „moralischer Akteur" zu sein, der zu eigener „Willensbildung" fähig sei; *Wohlers*, BJM 2016, 113, 123, für Roboterautos. Soll es nach *Wohlers* primär auf den „Grad an Autonomie" (von Roboterautos) ankommen, wenn es um die Frage geht, ob es (*de lege ferenda*) gerechtfertigt wäre, ihnen gegenüber einen individuellen Schuldvorwurf zu erheben, so könnte dies für zukünftige Formen selbstlernender KI-Systeme ebenfalls anders zu beurteilen sein.

86 Vgl. hierzu etwa den tödlichen Unfall, der 2018 durch ein selbstfahrendes Uber-Auto verursacht wurde, *NZZ* v. 19.3.2018, abrufbar unter: https://www.nzz.ch/panorama/selbstfahrendes-uber-auto-faehrt-frau-an-tot-ld.1367523 (Stand: 13.09.2019).

87 Zur Unterscheidung von kognitiven und normativen Erwartungen vgl. *Luhmann*, Rechtssoziologie (Anm. 42), S. 40 ff.

88 Zum Vertrauensgrundsatz allgemein *Sternberg-Lieben*/*Schuster*, in: *Schönke*/*Schröder*, 30. Aufl. 2019, § 15 Rdn. 151.

nieren auch der modernsten KI-Systeme verschleiert[89]. Das Roboterstrafrecht liefe daher dem Interesse eines effektiven strafrechtlichen Schutzes vor unerlaubten KI-Risiken entgegen.

Das vorstehende Ergebnis, wonach die Einführung eines „Roboterstrafrechts" keine sinnvolle Option der Strafgesetzgebung darstellt, gilt m. E. auch für den Fall, dass künftig eine sogenannte „starke künstliche Intelligenz" entwickelt werden sollte[90]. Der Begriff der „starken künstlichen Intelligenz" wird nicht einheitlich verwendet[91]. Er konkurriert mit Begriffen wie „Superintelligenz"[92] und „Singularität"[93]. In der Sache zielt er auf denkbare künftige Formen der künstlichen Intelligenz, deren Problemlösungskapazitäten nicht auf einen bestimmten Anwendungsbereich beschränkt sind (allgemeine Intelligenz), die zudem in der Lage sind, sich selbständig, ohne Eingaben eines Anwenders, weiterzuentwickeln (autonome Lernfähigkeit), und die auf diese Weise eine „Intelligenzstufe" erreichen, die dem Menschen gleichsteht und diese langfristig überschreiten werden. Das entscheidende Merkmal einer starken KI bzw. „Superintelligenz" besteht gemäß *Bostrom* darin, dass diese gegenüber dem Menschen einen „strategischen Vorteil" erlangt und dadurch nicht länger vom Menschen kontrolliert werden kann[94].

Folgt man dieser Analyse, so verweist „starke KI" aus normativer Perspektive vor allem auf ein höheres und nicht kontrollierbares Sicherheitsrisiko. Ein Roboterstrafrecht wäre nicht geeignet, um den mit einer möglichen maschinellen Superintelligenz angeblich verbundenen existenziellen Risiken zu begegnen, da dieses immer schon zu spät kommen würde[95]. Gerade aus der Perspektive derjenigen, die das Szenario einer dem Menschen überlegenen „Superintelligenz" für eine realistische Zukunftsprognose halten, erscheint die Vorstellung einer Kontrolle dieser Superintelligenz durch Strafandrohung gegenüber der Maschine als naiv. Warum sollte sich eine dem Menschen überlegene, mit einem strategischen Vorteil ausgestattete künstliche Intelligenz durch menschengemachte Strafgeset-

89 Den Vorrang menschlichen Handelns und menschlicher Aufsicht betonen nunmehr auch die Ethik-Leitlinien für eine vertrauenswürdige KI, *High-Level Expert Group on Artificial Intelligence*, Ethics Guidelines for Trustworthy AI, 2019, S. 17, abrufbar unter: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (Stand: 13.09.2019).

90 Für den Hinweis auf die möglichen Besonderheiten einer „starken Intelligenz" in der Diskussion bedanke ich mich bei Karsten Gaede. Vgl. dazu ausführlich *Gaede* (Anm. 76).

91 Zum Begriff vgl. auch *Gaede* (Anm. 76), S. 19 ff.

92 *Bostrom*, Superintelligenz, 2018.

93 *Kurzweil*, The singularity is near, London 2018.

94 *Bostrom* (Anm. 92), S. 115 ff.

95 Dies sieht auch *Gaede* (Anm. 76), S. 67 ff.

ze, Strafverfolgungsbehörden und die Strafjustiz beeindrucken lassen? Vielmehr müssten rechtliche Kontrollmechanismen darauf zielen, die Entstehung und „Freisetzung" einer Superintelligenz von vornherein zu verhindern[96]. Dies ist eine klassische Aufgabe des auf dem Vorsorgeprinzip[97] beruhenden Technikverwaltungsrechts, welches dann gegebenenfalls durch strafrechtliche Gefährdungsdelikte flankiert werden könnte[98]. Adressaten einer technik- und strafrechtlichen Kontrolle „starker KI" wären somit jedenfalls erneut die hinter der Technologie operierenden Menschen und nicht die Technik selbst.

Abschließend sei noch kurz auf die These eingegangen, ein Roboterstrafrecht sei dann sinnvoll und notwendig, wenn sich selbstlernende KI-Systeme so weit entwickelt hätten, dass wir sie als dem Menschen vergleichbare „moralische Akteure" verstehen könnten[99]. Nehmen wir an, dass eine so verstandene „starke künstliche Intelligenz" die Fähigkeit zur Ausbildung eigener normativer und rechtlicher Wertvorstellungen entwickelt, an denen sie ihr „Verhalten" orientiert. Tatsächlich müssten wir einer solchen Entität die Fähigkeit zubilligen, normative Erwartungen zu enttäuschen, sodass die normstabilisierende Funktion des Strafrechts grundsätzlich eröffnet wäre. Die Einwände gegen ein Roboterstrafrecht liegen hier auf der praktischen Ebene: Warum sollte eine mit der Fähigkeit vernünftiger Selbstbestimmung ausgestattete künstliche „Superintelligenz" sich ausgerechnet dem von menschlichen „Mängelwesen" produzierten jeweils geltenden Strafrecht unterwerfen? Die – für die Menschheit – entscheidende Frage dürfte dann vielmehr lauten, welchen moralischen und rechtlichen Status die aufgeklärt-autonomen Maschinen den Menschen zuschreiben werden[100].

2. *Strict liability*: Strafrechtliche Haftung für erlaubte Risiken?

Auf die menschliche Täterin hinter der Maschine würde demgegenüber die Einführung einer verschuldensunabhängigen strafrechtlichen Haftung für von KI-Systemen verursachte Rechtsgutsverletzungen im Sinne einer *strict liability* zielen. In der zivilrechtlichen Diskussion gilt die Einführung einer verschuldens-

96 So daher nachdrücklich *Bostrom* (Anm. 92), S. 184 ff.; vgl. auch *Gaede* (Anm. 76), S. 70 und 78.
97 Zum Vorsorgeprinzip vgl. *Calliess*, in: *Grunwald/Simonidis-Puschmann* (Hrsg.), Handbuch Technikethik, S. 390.
98 Siehe unten IV. 2.
99 Siehe dazu bereits Anm. 85.
100 Den Konflikt zwischen menschlicher und maschineller Moralität thematisiert *McEwan*, Maschinen wie ich, 2019, auf unterhaltsame Weise.

unabhängigen Gefährdungshaftung von Herstellern und/oder Betreibern von autonomen Software-Agenten als ein vielversprechender Lösungsansatz[101]. Die zivilrechtliche Gefährdungshaftung zielt darauf ab, dem gesellschaftlichen Bedürfnis nach technischer Innovation Rechnung zu tragen, das Schadensrisiko aber nicht den zufälligen Opfern unvorhersehbarer Folgen von KI-Systemen aufzubürden, sondern denjenigen, die wirtschaftlich von der Innovation profitieren. Damit wird zugleich die Risikoabschätzung an die Herstellerseite delegiert und dadurch nicht nur das Sorgfaltsniveau, sondern auch das Aktivitätsniveau gesteuert[102]. Für das Strafrecht wird eine verschuldensunabhängige Erfolgshaftung regelmäßig mit dem Hinweis auf das strafrechtliche Schuldprinzip abgelehnt[103]. Dies ist im Ergebnis überzeugend, bezeichnet aber nicht hinreichend konkret, worin das Problem einer Kausalhaftung für die Verletzungsfolgen des Einsatzes von KI-Systemen im Strafrecht besteht. Der entscheidende Gesichtspunkt ist, dass es sich um eine strafrechtliche Haftung für die Folgen einer erlaubten, nicht pflichtwidrigen Risikoschaffung handeln würde[104]. Dies erinnert an die Ende der 1950er Jahre geführte Debatte über die Frage, ob neben Vorsatz und Fahrlässigkeit das „riskante Verhalten" als schuldhafte Verhaltensform und Zurechnungsgrund im Strafrecht anzuerkennen sei[105]. Als Beispiele für eine Risikohaftung im Strafrecht wurden u. a. der Vollrauschtatbestand (§ 323a StGB) und die Beteiligung an einer Schlägerei (§ 231 StGB) diskutiert. Darüber hinaus wird auf das Zurechnungsprinzip des *versari in re illicita* verwiesen, welches den Täter wegen seines „Verweilens in einer unerlaubten Sache" für alle Folgen auch ohne sein Verschulden haften lässt[106]. Auch in den genannten Beispielen knüpft die Strafbarkeit indes an eine unerlaubte bzw. im Fall des – problematisch bleibenden – Vollrauschs an eine zumindest *nicht sozial adäquate, unerwünschte* Risikoschaffung durch den Täter an[107]. Erst vor diesem Hintergrund kann es hier – wenn überhaupt – als legitim erscheinen, den Eintritt des Erfolges bei den genannten Straftatbeständen als objektive Bedin-

101 Vgl. dazu *Hacker*, RW 2018, 243, 258 f.; *Schirmer*, RW 2018, 453, 473 ff.; *Wagner*, Robot Liability (June 19, 2018), S. 13 f., abrufbar unter: https://ssrn.com/abstract=3198764 (Stand: 16.09.19); *Zech*, in: *Gless/Seelmann* (Anm. 16), S. 197 ff.; *ders.*, ZfPW 2019, 198, 214 f.

102 *Zech*, ZfPW 2019, 198, 214; *Kötz/Wagner*, Deliktsrecht, 13. Aufl. 2016, Rdn. 503 ff.

103 *Gless*, recht 2013, 54, 57; *Gless/Janal*, JR 2016, 561, 564.

104 Vgl. *Kötz/Wagner*, Deliktsrecht (Anm. 102), Rdn. 491; *Zech*, in: *Gless/Seelmann* (Anm. 16), S. 197.

105 *Kaufmann*, Das Schuldprinzip, 1976, S. 145 ff.; *Schweikert*, ZStW 70 (1958), S. 394.

106 *Kaufmann* (Anm. 105), S. 146 f.; zum Begriff vgl. auch *Roxin*, Strafrecht Allgemeiner Teil, Bd. 1, 4. Aufl. 2006, § 10 Rdn. 122.

107 Zur Kritik an der herrschenden Konzeption der §§ 323a und 231 StGB vgl. *Roxin*, Allg. Teil I (Anm. 106), 23/7 ff.

gung der Strafbarkeit auszugestalten[108]. Daher stehen aktuelle Ansätze, die erwägen, für das Inverkehrbringen von gefährlichen KI-Systemen den Eintritt des Erfolges als objektive Bedingung der Strafbarkeit zu formulieren[109] vor dem Problem, dass sie primär die Unerlaubtheit des Inverkehrbringens begründen müssten. Dies setzte mithin zunächst die Schaffung eines abstrakten Gefährdungsdelikts des „Inverkehrbringens gefährlicher Produkte ohne hinreichende Sicherung" voraus[110]. Das nicht pflichtwidrige Herstellen, Inverkehrbringen oder Betreiben von KI-Systemen stellt demgegenüber eine erlaubte und gesellschaftlich erwünschte Tätigkeit dar. Ein erlaubtes Verhalten kann schon begrifflich keine rechtliche Verhaltenserwartung enttäuschen, sodass es keiner strafrechtlichen Bestätigung der Normgeltung bedarf. Eine strafrechtliche Risikohaftung für erlaubtes Verhalten ist daher als systemfremde Entgrenzung des Strafrechts ebenfalls abzulehnen[111].

3. Gefährdungsstrafrecht: Technische Innovation als unerlaubtes Risiko

Von der strafrechtlichen Erfolgshaftung für erlaubte Risiken ist drittens das Modell eines Gefährdungsstrafrechts zu unterscheiden, das auf abstrakten Gefährdungsdelikten basiert. Hier haftet die Täterin dafür, dass sie durch das Inverkehrbringen oder Betreiben von intelligenten Maschinen eine *unerlaubte* Gefahr für ein strafrechtlich geschütztes Rechtsgut schafft. Genauer genommen ist es das abstrakte Gefährdungsdelikt, welches bereits die Risikoschaffung unabhängig von einem Erfolgseintritt als unerlaubt qualifiziert. Mit Blick auf die gegenläufigen Funktionen des Technikrechts zwischen Innovationsförderung und Risikominimierung[112] sind abstrakte Gefährdungsdelikte problematisch, weil sie einseitig innovationsfeindlich sind. Eine Regulierung der künstlichen Intelligenz nach Maßgabe des technikrechtlichen Vorsorgeprinzips über abstrakte Gefährdungs-

108 Zur Gegenauffassung vgl. *Roxin*, Allg. Teil I (Anm. 106), 23/9, der für die Strafbarkeit fordert, dass der Täter hinsichtlich der im Rausch begangenen Tat fahrlässig gehandelt hat, um einen Verstoß gegen das Schuldprinzip zu vermeiden.

109 *Hilgendorf*, in: Festschrift für Fischer, S. 111.

110 Siehe Anm. 109 und unten IV. 4.

111 Gegen ein „Risikodelikt" als dritte Unrechts- bzw. Schuldkategorie auch *Radtke*/*Duttge*, in: Münchener Kommentar zum StGB, Bd. 1, 3. Aufl. 2017, § 15 Rdn. 30.

112 *Zech*, BJM 2014, 3, 6.; *Müller*/*Zech*, Sicherheit & Recht, 2019, 72, 78 f.; *Führ*, in: *Grunwald*/*Simonidis-Puschmann* (Anm. 97), S. 384, 386.

delikte kommt daher allenfalls bereichsspezifisch in Betracht[113]. Gleichwohl öffnet sich hier eine Möglichkeit für den Strafgesetzgeber, die Verantwortung des Menschen für Technikfolgen dadurch zu verdeutlichen, dass es untersagt wird, bestimmte Entscheidungen, etwa solche über Leben und Tod, an KI-Systeme zu delegieren oder KI-Systeme mit hohem Schadenspotenzial[114] in den Verkehr zu bringen.

Vor dem Hintergrund, dass KI-Systeme hier als innovative Risikotechnologien betrachtet werden, überrascht es nicht, dass in der Literatur die Regulierungsstrategie klassischer Risikoverwaltung als „Blaupause" für die KI-Regulierung vorgeschlagen wird[115]. Beispiele hierfür wären ein Verbot autonomer Waffensysteme[116], aber auch die Implementierung einer Gefährdungsstrafbarkeit für das Inverkehrbringen von autonomen Pflegerobotern ohne spezielle Zulassung, etwa in Anlehnung an die bereits existierende Regelung für Kraftfahrzeuge mit hoch- oder vollautomatisierter Fahrfunktion (§ 1a StVG)[117]. Schließlich wäre hier auch der Ort für eine rechtliche Vorsorge für die Gefahren des Entstehens und der Freisetzung einer sogenannten „starken KI" bzw. „Superintelligenz" (etwa in Analogie zum Gentechnikrecht)[118]. Da abstrakte Gefährdungsdelikte eine allenfalls bereichsspezifisch akzeptable (künftige) Lösung darstellen, wird die strafrechtliche Innovationsverantwortung für KI-Systeme auch künftig ganz überwiegend nach Maßgabe der allgemeinen Fahrlässigkeitsdogmatik zu beurteilen sein.

4. Fahrlässigkeit: Strafrechtliche Haftung für unerlaubte Risiken

Die strafrechtliche Fahrlässigkeitshaftung im Zusammenhang mit dem Inverkehrbringen oder Betreiben von KI-Systemen setzt voraus, dass der damit in Gang

113 Zur Notwendigkeit, die Risiken von KI anwendungsbezogen zu regulieren, vgl. auch *Meyer*, ZRP 2018, 233, 234.

114 Siehe Anm. 109.

115 *Martini* (Anm. 37), S. 113 ff. mit den Beispielen Nanotechnologie, Humangenetik und Arzneimittelrecht; vgl. ferner *Gaede* (Anm. 76), S. 76 f.

116 Vgl. *Dederer*, RW 2018, S. 380; zum Stand der politischen Diskussion vgl. Draft Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, abrufbar unter: https://www.unog.ch/80256EDD006B8954/(httpAssets)/5497DF9B01E5D9CFC125845E00308E44/$file/CCW_GGE.1_2019_CRP.1_Rev2.pdf (Stand: 18.09.2019); aus technikphilosophischer Perspektive *Weber*, in: *Gruber/Bung/Ziemann* (Anm. 64), S. 267; ferner *Lenzen*, Künstliche Intelligenz, 2018, S. 211 ff.

117 Zum Einsatz von Robotern im Palliativ- und Hospizbereich vgl. *Susanne Beck*, MedR 2018, 772.

118 Dazu bereits oben IV. 1. am Ende.

gesetzte schadensträchtige Verlauf für den Täter individuell vorhersehbar und vermeidbar war. Orientierung im Sinne des von *Duttge* geforderten „Veranlassungsmoments“[119] könnten insoweit gesetzliche oder durch Industrieverbände definierte Verhaltens- und Qualitätsstandards für spezifische Automatisierungen durch KI-Systeme gewährleisten. Ein Beispiel hierfür bildet die Regulierung von Kraftfahrzeugen mit hoch- oder vollautomatisierter Fahrfunktion im Straßenverkehrsrecht (§§ 1a und 1b StVG)[120]. Darüber hinaus fehlt es aber weitgehend an verbindlicher Normierung von Standards für die KI-Entwicklung. Wie *Yuan* in einem computerwissenschaftlich gut informierten Beitrag gezeigt hat, lassen sich indes auch für diesen ungeregelten Bereich eine Reihe von typischen Sorgfaltspflichtverstößen bei der Entwicklung von KI-Systemen identifizieren[121]. Soweit ein solcher Sorgfaltspflichtverstoß vorliegt, erscheint die spätere Verwirklichung des Autonomierisikos eines KI-Systems durch die Verletzung oder Tötung eines Menschen nicht als „technisches Versagen“, sondern als Folge einer unerlaubten Risikoschaffung, die der Programmiererin oder Herstellerin zugerechnet werden kann, wenn und soweit der Schadensverlauf für sie individuell vorhersehbar und vermeidbar war. So lassen sich allgemein etwa *Modellierungsfehler*, Fehler bei der *Auswahl von Trainingsdaten* und Fehler bei der *Evaluation der Sicherheit* der erzielten Ergebnisse unterscheiden[122]. Bei diesen Fehlern handelt es sich um klassische Produktionsfehler, deren Schadenspotenzial in der Regel von vornherein erkennbar war und nicht erst – wie bei einem sogenannten *Entwicklungsfehler* – nachträglich beim Betrieb des KI-Systems. In der Gerichtspraxis wird es insoweit freilich auf die genaue Kenntnis der jeweiligen technischen Funktionsweise eines KI-Systems ankommen, für die der Rechtsanwender letztlich auf externen Sachverstand der Computerwissenschaft angewiesen ist.

Für die Phase nach dem Inverkehrbringen des KI-Systems, hierauf haben *Gless* und andere bereits ausführlich hingewiesen[123], greifen die Grundsätze strafrechtlicher Produkthaftung, namentlich die Garantenpflicht der Herstellerin, auf nachträglich erkannte oder erkennbare riskante Fehlentwicklungen eines KI-Produktes zu reagieren.

Mit Blick auf das Erfolgsunrecht der Fahrlässigkeit ist es vor allem der Nachweis des Pflichtwidrigkeitszusammenhangs, der beim Einsatz von *Machine Lear-*

119 *Radtke/Duttge*, in: MK (Anm. 111), § 15 Rdn. 121 ff.; sowie grundlegend *Duttge*, Zur Bestimmtheit des Handlungsunwerts von Fahrlässigkeitsdelikten, 2001, S. 279 ff.

120 Vgl. dazu *Kaler/Wieser*, NVwZ 2018, 369, 370 f.

121 *Yuan*, RW 2018, 477, 495 ff.

122 *Yuan*, RW 2018, 477, 496.

123 *Gless/Janal*, JR 2016, 561, 565; *Gless/Weigend*, ZStW 126 (2014), S. 561, 582; *Gless*, recht 2013, 54.

ning-Systemen problematisch ist[124]. Die hierfür erforderliche Feststellung der hypothetischen Entwicklung des „Verhaltens" eines KI-Systems bei pflichtgemäßer Programmierung fällt insbesondere beim funktionsbasierten Maschinenlernen schwer, denn insoweit gelten KI-Systeme ungeachtet der neueren Forschungsarbeiten zu *explainable AI* nach wie vor weitgehend als *Black Boxes*[125]. Zu optimistisch könnte daher die Einschätzung sein, wonach gerade das mathematische Design von KI den Nachweis des Pflichtwidrigkeitszusammenhangs zum Beispiel über den Einsatz von Computersimulationen erleichtern könnte[126]. Gleichwohl weist die Überlegung in die richtige Richtung: Eine seriöse interdisziplinäre Wissenschaft maschinellen Verhaltens (*Machine Behaviour*)[127] könnte dazu beitragen, die für die strafrechtliche Fahrlässigkeitszurechnung erforderlichen Kausalzusammenhänge zu klären.

De lege ferenda ließe sich das Problem des Nachweises zwischen Pflichtwidrigkeit und Erfolg auch dadurch beseitigen, dass man, wie von *Hilgendorf* vorgeschlagen, den Eintritt des Erfolges beim Inverkehrbringen gefährlicher KI-Systeme ohne hinreichende Sicherung als objektive Bedingung der Strafbarkeit ausgestaltet[128]. Indem *Hilgendorf* das Inverkehrbringen von gefährlichen KI-Systemen nicht generell, sondern nur dann strafrechtlich erfassen möchte, wenn dies „ohne hinreichende Sicherung" erfolgt, erübrigt sich in seinem Modell der Nachweis der Fahrlässigkeit gerade nicht. Im Kern scheint sein Vorschlag vielmehr auf eine Beweiserleichterung bei der Fahrlässigkeitszurechnung zu zielen, da sich der Nachweis des Pflichtwidrigkeitszusammenhanges erübrigt, wenn der Eintritt des Erfolgs lediglich objektive Bedingung der Strafbarkeit ist. Demgegenüber gilt es zu erinnern, dass die Feststellung des Pflichtwidrigkeitszusammenhangs beim Fahrlässigkeitsdelikt auch in anderen Anwendungsfeldern und namentlich im Arztstrafrecht häufig schwierig ist, ohne dass darin unerträgliche Strafbarkeitslücken erblickt werden würden. Die Einführung einer *strict liability* in Form des Verzichts auf den Nachweis des Pflichtwidrigkeitszusammenhangs droht den Fahrlässigkeitstatbestand zu entgrenzen und ist daher auch im Kontext von KI-Systemen problematisch. Sofern das Inverkehrbringen von KI-Systemen vom Gesetzgeber für bestimmte Technologien oder Anwendungsbereiche nicht generell über ein echtes abstraktes Gefährdungsdelikt untersagt ist (dazu oben IV. 3), sollten Hersteller für durch das Produkt verursachte Schäden nur dann haften, wenn ihnen

124 Vgl. dazu *Yuan*, RW 2018, 477, 498 ff.

125 *Martini* (Anm. 37), S. 28; *Lenzen* (Anm. 116), S. 75 ff.; *Rahwan et al.*, nature 568 (2019), S. 477, 478; *Wischmeyer*, AöR 143 (2018), S. 1, 43 und 61 ff.

126 *Yuan*, RW 2018, 477, 499.

127 *Rahwan et al.*, nature 568 (2019), S. 477.

128 Siehe Anm. 109.

diese nach den allgemeinen Grundsätzen der Fahrlässigkeit zugerechnet werden können.

Blickt man abschließend auf die Seite des Betreibers von KI-Systemen wie etwa selbstlernenden Pflegerobotern, so lassen sich auch hier durchaus *konkrete Sorgfaltspflichten* definieren, bei deren Verletzung der Eintritt eines hinreichend bestimmten Verletzungserfolges in der Regel auch objektiv und subjektiv vorhersehbar ist. Im Vordergrund steht dabei die Frage, ob und inwieweit die Betreiberin eine Aufgabe, etwa im Bereich der Pflege, überhaupt auf ein im Einsatz nicht überwachtes KI-System delegieren durfte. Jedenfalls trifft auch den Betreiber die Pflicht, die sorgfältige Erfüllung der Aufgaben durch das KI-System laufend zu kontrollieren und auf Fehlentwicklungen zu reagieren. Ein lernendes KI-System ist eben gerade keine verantwortlich handelnde Person, auf die man sich nach dem Vertrauensgrundsatz verlassen darf; vielmehr stellt es, wie man in Anlehnung an die strafrechtliche Rechtsprechung zu Tiergefahren formulieren könnte, eine Gefahrenquelle dar, da es „in seinem Verhalten nicht vernunftgesteuert und im Allgemeinen unberechenbar ist“[129]. Für eine derartige „Gefahrenquelle“ ist „im Interesse eines gefahrlosen Gemeinschaftslebens derjenige verantwortlich, in dessen sozialem Herrschaftsbereich sie liegt (…)“[130]. Eine am Menschen orientierte strafrechtlichen Haftung für KI-Systeme erweist sich damit auch für die Fahrlässigkeit als Grundmodell strafrechtlicher Innovationsverantwortung als angemessen.

V. Fazit

Die vorstehenden Überlegungen zur Risikodogmatik der KI laufen darauf hinaus, dass sich die transformativen Effekte hochmoderner KI-Systeme auf die strafrechtliche Zurechnung von Innovationsrisiken in Grenzen halten bzw. halten sollten. Weder die Verantwortungsverlagerung auf Maschinen noch die Einführung einer verschuldensunabhängigen *strict liability* für KI-induzierte Verletzungen erscheinen als adäquate Reaktionen des Strafrechts. Es wird vielmehr darauf ankommen, einerseits absolute rechtliche Grenzen für die Delegation von Entscheidungen unter Ungewissheit auf selbstlernende Automaten im öffentlichen Technik- und Sicherheitsrecht zu definieren. Andererseits gilt es, für jeden einzelnen Anwendungsbereich von KI-Systemen im Lichte der allgemeinen Fahrlässigkeitsdogma-

129 BayObLG, 30.12.1987 – RReg. 3 St 226/87.

130 Siehe Anm. 129. Zu dieser Analogie zur Tiergefahr vgl. bereits *Zech*, in: *Gless*/*Seelmann* (Anm. 16), S. 196.

tik nach der individuellen Verantwortung der menschlichen Akteurinnen hinter der Maschine zu fragen. Darin wird zugleich sichtbar, dass das Konzept transformativer Technologien nicht im Sinne eines notwendigen Kausalzusammenhangs zwischen technologischen und strafrechtlichen Innovationen verstanden werden darf. Vielmehr ist es die Einsicht in das transformative Potenzial einer Technologie, die eine kritische Reflexion über die „Risiken des Risikostrafrechts“[131] erst ermöglicht. Es gilt, mit anderen Worten, zu verstehen, was dem Strafrecht passiv widerfährt, indem es sich in Reaktion auf die Digitalisierung aktiv verändert[132].

131 *Prittwitz*, in: *Frehsee/Löschper/Smaus* (Hrsg.), Konstruktion der Wirklichkeit durch Kriminalität und Strafe, 1997, S. 47 ff.

132 In Anlehnung an eine Formulierung von *Baecker* (Anm. 5), S. 9.