

Maschinenethik und strafrechtliche Verantwortlichkeit

Monika Simmler

Inhalt

1	Einleitung	2
2	Autonomie, Agency und Verantwortung	3
3	Strafbarkeit von Maschinen	4
4	Strafbarkeit von moralischen Maschinen	6
5	Verteiltes Handeln und geteilte Verantwortung	8
6	Fahrlässigkeit bei der Delegation an Technik	10
7	Sorgfaltsmaßstab und technische Autonomie	11
8	Maschinelle Moral als Sorgfaltspflicht	13
9	Zusammenspiel von Recht, Technik und Maschinenethik	14
10	Resümee	14
	Literatur	15

Zusammenfassung

Handlungen sind zunehmend das Resultat eines Zusammenwirkens von Mensch und (teil-)autonomer Maschine. Das Verhalten dieser Maschinen lässt sich weder dem Menschen noch der Technik selbst ohne Weiteres als Handlung zurechnen. Dieser Herausforderung für die strafrechtliche Verantwortungszuschreibung nimmt sich der vorliegende Beitrag an. Es wird festgestellt, dass Maschinen keine Strafrechtssubjekte darstellen, was allerdings nicht für alle Zeiten gelten muss. Insbesondere die Emergenz moralischer Handlungsfähigkeit von Technik könnte zu einem Wandel beitragen. Bis dahin steht jedoch die Verantwortlichkeit der Menschen im Zentrum. Sie sind gefordert, adäquate Sorgfalt walten zu lassen. Dabei kann gerade das Anliegen der Maschinenethik – die Moralimplementation – zur eigentlichen Sorgfaltspflicht werden. Die Disziplin entfaltet damit Relevanz für das Strafrecht, hängen doch maschinelle Moral, Handlungsträgerschaft und Verantwortung eng zusammen.

M. Simmler (✉)

Law School, Universität St. Gallen, St. Gallen, Schweiz

E-Mail: monika.simmler@unisg.ch

Schlüsselwörter

Strafrecht · Rechtstheorie · Maschinenethik · Schuldvorwurf · Handlungsfähigkeit

1 Einleitung

Das digitale Zeitalter ist durch zwei wesentliche Transformationen gekennzeichnet: Durch das Aufkommen der künstlichen Intelligenz lässt sich einerseits ein Übergang von der bloßen Automatisierung zur *Autonomisierung* beobachten (Müller-Hengstenberg und Kirn 2016). Maschinen werden komplexer, undurchsichtiger (Rammert 2003) und zeichnen sich durch Interaktivität und Anpassungsfähigkeit aus (Floridi und Sanders 2004). Das resultiert in neuen Freiheitsgraden der Technik. Diese technische Autonomie kann mehr oder weniger ausgeprägt sein (Simmler und Frischknecht 2019) und beschreibt relational den Grad der Unabhängigkeit der Maschine von etwas, wie dem Einfluss einer anderen Entität, der Umwelt oder interner Zwänge (Verhagen 2014; Castelfranchi und Falcone 2004; Müller-Hengstenberg und Kirn 2016). Andererseits lässt sich eine *Hybridisierung* der Gesellschaft feststellen, da Handlungen zunehmend das Resultat soziotechnischer Handlungszusammenhänge sind (Weyer und Reineke 2005). Das Verhältnis von Mensch und Maschine verändert sich: Aus Instrumentalität wird Interaktivität (Rammert 2003).

Das moderne Strafrecht sieht sich aufgrund dieses Wandels mit einer Welt konfrontiert, in der neue Akteure hervortreten, die nicht menschlich sind und nicht ohne Weiteres unter herkömmliche Verantwortlichkeitskategorien subsumiert werden können. Dennoch führen sie aufgrund ihrer Komplexität Aktionen aus, die bis anhin nur dem Menschen zuteilwurden (Seher 2016). Je mehr Maschinen ein derartiges Ausmaß an Eigenaktivität aufweisen, desto schwieriger wird es, sie als simple und passive Instrumente zu beschreiben (Rammert 2003). Es stellt sich entsprechend die Frage, ob Technik handlungsfähig ist – sowohl moralisch als auch strafrechtlich. Ebenso ist zu klären, inwiefern die erhöhte Autonomie der Technik die Verantwortlichkeit der Menschen berührt.

Im Spannungsfeld jahrhundertlang verhandelter Handlungs- und Verantwortungsbegriffe erkämpft sich nun eine neue Disziplin ihren Platz: die Maschinenethik. Sie hat die Entwicklung „moralischer Maschinen“ zum Ziel, d. h. Maschinen sollen befähigt werden, eigens moralisch zu handeln (Misselhorn 2018a). In Analogie zur künstlichen Intelligenz wird bei der „maschinellen Moral“ also versucht, menschliche Moral zu simulieren (Bendel 2018a, b). Dabei solle die Maschinenethik diese *Artificial Morality* (Misselhorn 2018b) nicht nur untersuchen, sondern im besten Falle auch schaffen, was aufgrund der zunehmenden Autonomie technischer Systeme notwendig werde (Bendel 2018a, b). Ein kurzer Blick auf den Gegenstand der Maschinenethik genügt, um festzustellen, dass ihr Unterfangen das Potenzial hat, auch die strafrechtliche Verantwortungszuschreibung zu prägen. Der Zusammenhang zwischen Strafrecht und Maschinenethik verdient deshalb eine nähere Untersuchung.

Im vorliegenden Beitrag ist zunächst grundlegend zu beleuchten, inwiefern der technische Fortschritt die strafrechtliche Verantwortungszuschreibung berührt (Abschn. 2). In der Folge wird diskutiert, ob avancierte Maschinen Subjekte strafrechtlicher Verantwortlichkeit darstellen (Abschn. 3) und ob insbesondere das Aufkommen *moralischer* Maschinen zu einer derartigen Maschinenverantwortlichkeit führen kann (Abschn. 4). In einem zweiten Teil ist zu untersuchen, wie soziotechnisch verteiltes Handeln die Verantwortlichkeit von Menschen berührt (Abschn. 5). Beim (ethischen) Design von Technik erlangt vor allem der durch Sorgfaltspflichten konstituierte Fahrlässigkeitsvorwurf Bedeutung (Abschn. 6). Der ihm zugrunde liegende Sorgfaltsmaßstab hängt vom Grad technischer Autonomie ab (Abschn. 7), wobei auch die Maschinenethik Einfluss auf dessen Konstituierung gewinnen kann. So könnte gerade die Möglichkeit einer „Moralimplementation“ zur eigentlichen Sorgfaltspflicht werden (Abschn. 8). Daran anknüpfend schließt der Beitrag mit einer Betrachtung des allgemeinen Zusammenspiels der technischen Entwicklung, des (Straf-)Rechts und der (Maschinen-)Ethik (Abschn. 9), bevor ein Resümee den Artikel abrundet (Abschn. 10).

Zusammengefasst bietet dieser Beitrag also einen Überblick, wie Maschinenethik und Strafrecht zusammenhängen. Angetrieben ist dieses Unterfangen durch die Hoffnung, die verwandten Diskussionen dieser Disziplinen mögen sich gegenseitig bereichern, und durch die Überzeugung, dass Fragen der *Verantwortung* im digitalen Zeitalter längst nur noch interdisziplinär zu beantworten sind.

2 Autonomie, Agency und Verantwortung

Das Forschungsgebiet der künstlichen Intelligenz hat Maschinen zu Tage getragen, die aufgrund ihres Grades an Autonomie den traditionellen Mensch-Technik-Dualismus mindestens herausfordern, wahrscheinlich aber gar überwinden. Überwindung soll hierbei nicht als vollständige Auflösung der Differenz verstanden werden. Selbstverständlich sind Mensch und Maschine grundverschieden. Daran soll diese Auffassung nicht rütteln. Trotzdem erfährt das Verständnis der determinierten Technik auf der einen und der indeterminierten Menschen auf der anderen Seite – dieses klare Gegenüber von Notwendigkeit und Freiheit – eine Veränderung, welche an ideengeschichtlichen Grundfesten nagt und eine theoretische Neuauslegung verlangt.

Die „Annäherung von Mensch und Maschine auf der Autonomieskala“ fordert insbesondere das Fundament des strafrechtlichen Verantwortungskonzeptes heraus (Beck 2015, S. 180). Dieses baut auf einem *individuellen Schuldvorwurf*. Das Schuldprinzip wird damit zum Kern strafrechtlicher Verantwortlichkeit. Es verlangt, dass Verurteilung und Strafe den Einzelnen nur dann treffen dürfen, wenn ihm ein persönlicher Vorwurf gemacht werden kann. Persönlich vorwerfbar und damit schuldhaft ist eine Handlung nur, wenn sie von einem adäquaten Strafrechtssubjekt (*Person*) ausgeführt wurde, das zum Handlungszeitpunkt die *Kompetenz* hatte, sich auch anders zu verhalten, und wenn dieses Subjekt aufgrund eines vorsätzlichen oder fahrlässigen Verhaltens die Geltung einer durch das Strafrecht geschützten

Norm *infrage stellt* (Simmler 2018). Betrachtet man diesen Grundstein der Strafbarkeit, welcher sich in der Dogmatik durchgängig niederschlägt, wird schnell deutlich, weshalb die zunehmende Autonomie von Maschinen das strafrechtliche Verantwortungskonzept im Kern betrifft. Das Strafrecht rechnet Personen Handlungen (*Unrecht*) aufgrund ihrer allgemeinen und situativen *Autonomie* zu. Ihnen wird die Handlung vorgeworfen (*Schuld*), weil sie aufgrund dieser Autonomie auch hätten anders handeln können. Autonomie, Handlungsträgerschaft (*Agency*) und Schuldvorwurf hängen also direkt zusammen. Werden Maschinen nun immer autonomer, stellen sich grundsätzlich zwei Fragen: Führt die zunehmende Autonomie von Maschinen dazu, dass diese eigens strafrechtlich handlungs- und schuldfähig werden? Und: Inwiefern berührt diese Autonomie die Verantwortlichkeit ihrer menschlichen Schöpferinnen und Interaktionspartner?

Die Klärung dieser Fragen ist herausfordernd und wichtig zugleich, steht doch der Verdacht im Raum, dass die zunehmende Autonomie von Technik es erschwert oder gar unmöglich macht, ihre Handlungen einem Menschen zuzurechnen. Die hohen Ansprüche an die *individuelle* Zuschreibung eines Schuldvorwurfs sind nur schwierig mit diesem vermehrten Zusammenspiel menschlicher und nichtmenschlicher Akteure in Einklang zu bringen (Beck 2017). Komplexen soziotechnischen Handlungszusammenhängen ist regelmäßig das Potenzial inhärent, zu einer Diffusion von Verantwortung zu führen (Maring 2008). Solange das Verhalten einer Maschine einfach einem Menschen zugeordnet werden kann, können Handlungen in Anwendung bestehender rechtsdogmatischer Konzepte ohne Weiteres zugerechnet werden. Das ist der Fall, wenn Maschinen aufgrund ihrer relativen Primitivität als Werkzeug zu gelten haben oder bei fortschrittlicherer Technologie, die im expliziten Auftrag eines Menschen agiert (Chinen 2016). Avancierte Maschinen schaffen aber neue Konstellationen. Ihr Output ist nicht mehr klar vorhersehbar und kontrollierbar. Es wäre einerseits ungerecht, Menschen für solche nicht absehbaren Konsequenzen zur Rechenschaft zu ziehen. Andererseits darf der Tendenz, sich hinter der Maschine zu „verstecken“, auch keinen Vorschub geleistet und der Mensch nicht vorschnell aus seiner Verantwortung entlassen werden (vgl. Mittelstadt et al. 2016). Ansonsten droht eine „Verantwortlichkeitslücke“ zu entstehen (Matthias 2004). Wer für den neuartigen Freiheitsgrad der Maschinen einzustehen hat, bleibt für den Moment also rechtlich unklares Terrain. Als mächtiges Instrument sozialer Kontrolle (Hallevy 2010) steht das Strafrecht jedoch auch im digitalen Zeitalter vor der Aufgabe, sich fortzuentwickeln und die Verantwortlichkeit der beteiligten Akteure möglichst klar festzulegen. Dabei kann die Maschinenethik eine neue Perspektive bieten – sowohl hinsichtlich sich neu eröffnender Möglichkeiten maschineller Verantwortung als auch hinsichtlich der Pflichten der menschlichen Erschaffer.

3 Strafbarkeit von Maschinen

Aufgrund des Gleichschritts von Autonomie und Verantwortung stellt sich der Strafrechtswissenschaft die Frage, ob es sich bei intelligenten Maschinen potenziell um eigenständige Subjekte des Strafrechts handelt, ob Formen künstlicher Intelli-

genz also selbst Adressaten eines Schuldvorwurfs sein können (z. B. Hallevy 2010). Es gilt entsprechend zu klären, ob das Strafrecht gegenüber einer künstlich erschaffenen Entität jemals den es so sehr in seiner Identität als Schuldstrafrecht bestimmenden persönlichen Vorwurf erheben könnte, Technologie also schuldfähig sein kann (Simmler und Markwalder 2017). Den heute vorherrschenden Vorstellungen folgend ist das nicht möglich, da Maschinen nicht als verantwortliche Agenten begriffen werden und daher nicht Adressaten einer Strafe sein können (vgl. Gless et al. 2016). Wie lange sich diese traditionelle Annahme noch behaupten kann, ist allerdings fraglich, ist doch der Schuldvorwurf – wie die zunehmende Verbreitung der Unternehmensstrafbarkeit in vielen Rechtsordnungen (wenn auch nicht in Deutschland) zeigt – nicht ausschließlich Menschen vorbehalten (Simmler und Markwalder 2017).

Die Ansprüche an den Subjektbegriff sind im Strafrecht allerdings besonders hoch. Die traditionelle Schuldlehre basiert auf der Annahme der menschlichen Willensfreiheit bzw. eines *Andershandelnskönnens*. Gemäß der Definition des deutschen Bundesgerichtshofs bedeutet Schuld nichts weniger, als dass sich ein selbstbestimmter und mit einem freien Willen ausgestatteter Mensch gegen das Recht und für das Unrecht entscheidet (BGHSt. 2, 194, 200). Auf den ersten Blick scheint dieses Verständnis eine Verantwortlichkeit von Maschinen auszuschließen. Diese „klassische“ Schuldauflassung sieht sich allerdings seit jeher Kritik und Relativierungen ausgesetzt. Auch wenn diese Debatte hier nicht im Detail nachgezeichnet werden kann (siehe dafür Simmler 2018), lässt sich summarisch konstatieren, dass Willensfreiheit und Andershandelnskönnen sozial konstruiert sind und vergleichbare Zuschreibungen Maschinen nicht per se verwehrt bleiben müssen (Simmler und Markwalder 2017). Die Freiheit, welche dem Schuldvorwurf zugrunde liegt, ist eine „Freiheit zur Selbstverwaltung“ (Jakobs 1993, S. 34). Eine solche könnte teilweise oder vollständig auch technischen Systemen eingeräumt werden, würde die soziale Praxis der Zukunft eine solche Zuschreibung erlauben.

Die Beantwortung der Frage nach der Möglichkeit strafrechtlicher Personenqualität künstlicher Gebilde hängt ebenfalls davon ab, welcher Auffassung man im Rahmen der grundlegenden theoretischen Auseinandersetzung folgt. Ist man beim Personenbegriff der idealistisch-philosophischen Annahme verpflichtet, dass nur als Person gelten könne, wer sich seiner Freiheit bewusst und reflexionsfähig sei, und sich ferner selbst als Entität mit Vergangenheit und Zukunft verstehen könne, muss schnell zum Schluss kommen, dass eine persönliche Verantwortung von Maschinen auszuschließen ist (vgl. Gless und Weigend 2014; Gless et al. 2016). Diese Auffassung vernachlässigt allerdings den Umstand, dass sich auch der Personenbegriff alleine in der sozialen Wirklichkeit konstituiert. Dabei knüpft er nicht notwendigerweise an naturwissenschaftliche Kategorien an. Auch bei Personalität handelt es sich um das Resultat eines Zuschreibungsprozesses und nicht um eine ontologische Gegebenheit (Simmler und Markwalder 2017). Als normative Konstruktion ist der Personenbegriff folglich gesellschaftsrelativ und sozial wandelbar, was wiederum die Möglichkeit einer Maschinenstrafbarkeit nicht von vornherein ausschließt.

Dass der Personenbegriff keinen naturwissenschaftlichen Kategorien folgt, darf nun nicht dahingehend missverstanden werden, dass die Gesellschaft Personalität

willkürlich zuschreibt. Während Unternehmen aufgrund der Emergenz kollektiver Handlungsträgerschaft als Adressaten normativer Erwartungen aufgefasst werden können, ist dies bei Maschinen auf dem heutigen technischen Stand fraglich (Simmler und Markwalder 2017). Wenn in dieser Debatte vermehrt die Konstruktion sogenannter „e-Personen“ als Haftungssubjekte eigener Art vorgeschlagen wird, mittels welcher Maschinen ein Rechtsstatus eingeräumt werden könnte (Beck 2013; Gruber 2012; Gless und Weigend 2014), mag das für das Zivilrecht eine spannende Idee sein. Für das Strafrecht kann das aber kaum überzeugen. Der Schuldvorwurf folgt der Funktion des Strafrechts, d. h. der spezifischen Stabilisierung von Normen mittels Sanktion (Simmler 2018). Die Strafe entfaltet ihre Wirkung jedoch nur, wenn wir Subjekten überhaupt die Kompetenz zuerkennen, Erwartungen zu enttäuschen und Normen infrage zu stellen (Jakobs 1993). Maschinen als „ePersonen“ zu konstruieren könnte eine Folge der Zuerkennung dieser Kompetenz sein, nicht aber die Grundlage maschineller Personalität (Simmler und Markwalder 2017).

Erst die Funktion des Strafrechtssystems füllt den strafrechtlichen Subjektbegriff und den strafrechtlichen Schuldvorwurf mit Sinn (Simmler 2018). Nur mit Blick auf die Funktion der Normstabilisierung lässt sich deshalb auch die Frage nach einer Maschinenverantwortlichkeit beantworten. Es ist deshalb zu klären, ob Maschinen das Potenzial mit sich bringen, Normen zu destabilisieren. Täten sie dies, müsste die Gesellschaft einen Weg finden, derartige Angriffe auf die Normgeltung zu erwidern, um sie weiterhin garantieren zu können (Simmler und Markwalder 2017). Dafür müsste die Personalisierung von Maschinen allerdings weit fortgeschritten sein. Wir müssten an sie nicht nur kognitive, sondern auch normative Erwartungen richten. Nur dann wären ihre Enttäuschungen nicht bloßes „Pech“, sondern Ausdruck eines Mangels an Rechtstreue (vgl. Jakobs 1991, 1993). Als Adressat normativer Erwartungen sähe sich eine derart avancierte Maschine mit dem Anspruch konfrontiert, ihr Verhalten eigenständig normgemäß anpassen zu können (Simmler und Markwalder 2017). Eine Maschinenstrafbarkeit würde also voraussetzen, dass Maschinen aufgrund der ihnen zugeschriebenen Kompetenzen Normen destabilisieren können.

4 Strafbarkeit von moralischen Maschinen

Wie soeben skizziert, würde eine Strafbarkeit von Maschinen verlangen, dass diese als gleichgestellte „Teilnehmer“ an der normativen Kommunikation der Gesellschaft erlebt würden. Sie müssten in der Lage sein, die Geltung geschützter Normen durch verübtes Unrecht zu berühren. Strafe würde in diesem Fall erforderlich, um die Normgeltung zu bekräftigen. Dies setzt allerdings voraus, dass wir Maschinen zuerkennen, sich schuldhaft über Normen hinwegsetzen und damit einen Mangel an Rechtstreue zum Ausdruck bringen zu können. Diese Fähigkeit zu strafbarem Handeln ist nun eng mit derjenigen zu moralischem Handeln verknüpft bzw. zeigen sich hier diverse Analogien. Eine Strafbarkeit von Maschinen kommt kaum in Betracht, solange diese nicht auch zu „moralischen Agenten“ werden (Gless und

Weigend 2014), womit das Vorhaben der Maschinenethik für die Frage der Strafbarkeit von Technik an Bedeutung gewinnt.

Bei der Diskussion des Verhältnisses von strafrechtlicher Handlungsfähigkeit und dem Gegenstand der Maschinenethik, der maschinellen Moral, gilt es vorweg zu betonen, dass das Rechtssystem nicht mit der Moral gleichgesetzt werden darf (vgl. Misselhorn 2018b). Das Strafrecht ist operativ geschlossen (Simmler 2018). Der daraus resultierende Rechtspositivismus führt dazu, dass das Strafrecht die Geltung seiner Normen voraussetzt. Sie werden im Einzelfall nicht auf ihre moralische „Richtigkeit“ hin überprüft – sollte eine solche der Erkenntnis überhaupt zugänglich sein. Ebenso kann es nicht die gesellschaftliche Funktion des Strafrechtssystems sein, die Moralität einer einzelnen Handlung zu beurteilen. Dennoch schützt das Strafrecht bis zu einem gewissen Grad das „moralische Band der Nation“ (Durkheim 1893). Es gilt damit als Indikator für zu einem Zeitpunkt vorherrschende Moralvorstellungen. Wenig überraschend weisen deshalb die strafrechtliche und die moralische Verantwortungszuschreibung bzw. deren Fundamente eine Verwandtschaft auf, die es erlaubt oder geradewegs dazu verpflichtet, bei der Verfeinerung Ersterer (maschinen-)ethische Diskussionen miteinzubeziehen.

Die Meinungen, ob Maschinen moralisch handlungsfähig sein können, gehen allerdings auseinander: So meint z. B. Seng (2018), Maschinen hätten per se keine Moral, da diese eine „menschliche Kategorie“ sei. Auch Weber und Zoglauer (2018) halten für existierende Maschinen fest, dass diese keine moralischen Agenten seien. Künstliche Intelligenz könne allerdings durchaus moralische oder unmoralische „Produkte“ erzeugen (Seng 2018). In die Strafrechtsterminologie übersetzt bedeutete diese Annahme, Maschinen könnten Unrecht verwirklichen, dieses Unrecht könnte ihnen aber nicht vorgeworfen werden. Sie wären trotz der moralischen Wirksamkeit ihres Verhaltens nicht schuldfähig.

Es gibt aber auch andere Ansätze: So erblickt Bendel (2018a, S. 34) in Maschinen „neue, fremde, merkwürdige Subjekte der Moral“. Ihm zufolge stellt „maschinelle Moral“ – in Analogie zu „künstlicher Intelligenz“ – ein Terminus technicus dar. Es sei möglich, bestimmte Aspekte menschlicher Moral maschinell abzubilden, bestimmte andere Aspekte dagegen nicht (Bendel 2019). Auch Misselhorn (2018b, S. 13 f.) anerkennt die Möglichkeit, Maschinen moralisches Handeln zuzugestehen, gradualisiert jedoch sogleich: Zwar seien Maschinen moralische Akteure, sie seien aber nicht zu „vollumfänglichem moralischen Handeln wie Menschen in der Lage“. Dazu würden ihnen verschiedene für die Übernahme moralischer Verantwortung wesentliche Eigenschaften fehlen. Auch Wallach und Allen (2009) differenzieren verschiedene Stufen von *Moral Agency*, um sogleich festzuhalten, dass Technik die höchste noch nicht erreicht habe. Spricht man von Maschinen als moralische Agenten, könne das auch gemäß Weber und Zoglauer (2018) vieles heißen: Deren Verhalten könne nach den von Menschen implementierten Vorgaben moralisch relevant sein. Maschinen könnten aber auch selbst die Fähigkeit besitzen, intentional moralische Normen zu beachten. Als Steigerungsformen könne es ferner bedeuten, dass Maschinen als ethische Agenten selbst moralische Normen begründen oder sich gar gegen die von ihnen gewählten Normen stellen. Dies würde sie zu einem „Spiegelbild des Menschen mit all seinen moralischen Defekten“ machen. Auch

diese Autoren gehen also davon aus, dass *Moral Agency* in ihrem Grad variieren kann und sie (ganz oder bis zu einem gewissen Grad) nicht nur Menschen vorbehalten ist.

Die Frage nach der *moralischen* Handlungsfähigkeit ist diesem Ansatz folgend keine binäre. Eine solche *Gradualisierung* ist für das Strafrecht interessant, berührt sie doch die herkömmlich binäre „Entscheidung“ zwischen Person/Nicht-Person und den Dualismus von Handlung/Nicht-Handlung. Auch eine Gradualisierung der *strafrechtlichen* Handlungsfähigkeit ist nicht von vornherein ausgeschlossen. Dies verlangte allerdings eine vertiefte Auseinandersetzung mit dem strafrechtlichen Handlungsbegriff und dem Verhältnis von Handlungs- und Schuldfähigkeit. Für den Moment muss die Feststellung genügen, dass der Grad an Verantwortlichkeit zwar variieren kann, die Fragen nach der strafrechtlichen Subjektqualität und nach der Erhebung eines Schuldvorwurfs grundsätzlich aber nicht graduell beantwortet werden können: Jemand ist Subjekt strafrechtlicher Verantwortlichkeit oder jemand ist es nicht. Jemand wird bestraft oder jemand wird es nicht. Dennoch ist dieser Diskurs keineswegs nur ein „akademisches Glasperlenspiel“ (Rammert und Schulz-Schaeffer 2002, S. 32), kann sich eine mögliche (gradueller) moralische Handlungsfähigkeit von Technik doch maßgeblich auf die Verantwortungszuschreibung auswirken.

Für den heutigen Stand der Technik und unserer aktuellen sozialen Interaktion folgend muss die Frage nach einer strafrechtlichen Verantwortung von Maschinen verneint werden (Simmler und Markwalder 2017; Markwalder und Simmler 2017). Sie sind keine Adressaten der spezifisch strafrechtlichen normativen Erwartungen. Sie können diese nicht enttäuschen, da ihnen die nötige Personenqualität und die vorausgesetzte Kompetenz nicht zugesprochen werden. Dieses Zwischenfazit ist allerdings nur ein Zwischenfazit. Sollten Maschinen vermehrt zu moralischen Agenten höherer Stufen werden, könnte sich der Befund ändern. Das Strafrecht folgt, wie bereits betont, seiner Funktion innerhalb der Gesellschaft. Es folgt der sozialen Wirklichkeit, dem sozialen Spiel. Dabei ist es nicht an naturalistische Kategorien gebunden. Genau wie die Maschinenethik ist das Strafrecht nicht „speziesistisch“ (vgl. Nida-Rümelin und Battaglia 2018). Es wird die Entwicklung – insbesondere diejenige der Schaffung maschineller Moral – deshalb weiterverfolgen müssen. Weisen Maschinen irgendwann eine dem Menschen vergleichbare Stufe moralischer Handlungsfähigkeit auf, dreht sich auch das Rad des strafrechtlichen Personen- und Handlungsbegriffs neu.

5 Verteiltes Handeln und geteilte Verantwortung

Wenn Maschinen (zumindest zu einem gewissen Grad) moralisch handeln können, nicht jedoch moralisch oder strafrechtlich zur Verantwortung gezogen werden (vgl. Misselhorn 2018a), stellt sich vordringlich die Frage, wer für ihr Handeln verantwortlich ist, wenn sie es nicht (oder noch nicht) selbst sind. Obwohl moralfähige Maschinen keine Subjekte des Strafrechts sind, ist die Frage nach der maschinellen Moral für das Strafrecht also bedeutend. Dabei steht die Frage im Zentrum, inwie-

fern das „Einpflanzen“ von Moral bzw. die Repräsentation menschlicher Moral in Maschinen Programmiererinnen zugeschrieben werden kann und diese Erweiterung der maschinellen Kompetenzen auch deren Verantwortlichkeit mitkonstituiert. Die Maschinenethik wird deshalb zum jetzigen Zeitpunkt vorwiegend für die (menschliche) Verantwortung bei der Entwicklung und im Umgang mit Maschinen relevant.

Das digitale Zeitalter ist durch soziotechnisch verteiltes Handeln gekennzeichnet (Rammert 2003). Die Frage, wie in derartigen hybriden Systemen Verantwortung adäquat zugeschrieben werden kann, muss aktuell als unbeantwortet gelten. Es ist zunächst plausibel, dass verteilte Handlungsträgerschaft auch in einer verteilten Verantwortlichkeit bzw. in einem eigentlichen „Verantwortlichkeitsnetzwerk“ resultiert (Loh und Loh 2017). Künstliche Systeme berühren den menschlichen „Verantwortungsanteil“. Auch wenn man eine originäre Verantwortlichkeit technischer Entitäten ablehnt, befreit das also nicht von der Aufgabe, das Verantwortungskonzept zu verfeinern und anzureichern, sodass die technische „Beteiligung“ einbezogen werden kann (vgl. Nida-Rümelin 2017). Auch Maschinen in soziotechnischen Handlungszusammenhängen stellen jedoch gewissermaßen eine „Repräsentation der ihre Regeln definierenden natürlichen Personen“ dar (Rath 2019, S. 228). Sie unterliegen einem bestimmten „Verhaltensregelwerk“ und die sie leitenden Algorithmen repräsentieren die Wertvorstellungen des Programmierers (Rath 2019, S. 230). Auch wenn dieses „lineare Modell“ für komplexere Maschinen schnell zu simplistisch wird, da sich die Programmierung unter Mitwirkung autonomer Technik „mediatisiert realisiert“ (Rath 2019, S. 230 f.), führt diese „Repräsentation“ und „Delegation“ dazu, dass der Mensch auch für den „Handlungsanteil“ der Maschinen grundsätzlich nicht aus der Verantwortung zu entlassen ist – zumindest solange nicht, wie Maschinen keine eigenständigen Verantwortlichkeitssubjekte darstellen.

Es gibt verschiedene Szenarien, wie Maschinen in eine Straftat „verwickelt“ sein können. Einerseits gibt es Fälle, in denen eine Maschine ein Unrecht verübt, weil sie vorsätzlich so programmiert wurde. Solche Vorsatztaten dürften wenig problematisch sein: Der Mensch dahinter bleibt strafrechtlich verantwortlich. Es handelt sich hierbei schlicht um einen bewussten Missbrauch intelligenter Agenten für kriminelle Zwecke (Günther 2015; Seher 2016; Gless et al. 2016; Markwalder und Simmler 2017). Schwieriger sind diejenigen Fälle, bei denen eine Maschine die Straftat „begeht“, weil sie fehlerhaft programmiert wurde. Die Fahrlässigkeitsstrafbarkeit steht im Zusammenhang mit avancierter Technik deshalb im Vordergrund. Entwicklerinnen stehen vor der Herausforderung, einem Fahrlässigkeitsvorwurf zu entgehen und nicht für fehlerhaftes Verhalten der Maschine einstehen zu müssen. Gleichzeitig bestehen noch keine eindeutigen Regeln, die ihnen helfen, das geforderte Maß an Sorgfalt klar zu bestimmen (Beck 2017). Damit wird die Frage, welche Maßstäbe im Bereich der Fahrlässigkeit angelegt werden und wie sich die entsprechenden Sorgfaltspflichten konstituieren, zum Dreh- und Angelpunkt der Strafbarkeit für Technik (Günther 2015).

6 Fahrlässigkeit bei der Delegation an Technik

Die Zuschreibung von Verantwortung für die fehlerhafte Programmierung, Implementierung oder Nutzung von technischen Systemen ist überall dort problematisch, wo Arbeitsabläufe mit komplexen technologischen Hilfsmitteln unterstützt werden, die aufgrund eines Mindestmaßes an Autonomie mehr als ein bloßes Werkzeug darstellen. Es stellt sich hier die Frage, welchem Akteur die Folgen technischer Autonomie als Handlung zugerechnet werden können. Dies betrifft primär den strafrechtlichen Handlungsbegriff, der enger oder weiter gefasst sein kann. Während Vorsatzdelikte durch die Finalität der Handlung gekennzeichnet sind, ist es beim Fahrlässigkeitsvorwurf die *Sorgfaltspflichtverletzung*, welche die Zurechnung eines Geschehens zu einer Person erlaubt.

Als „Moment der Rechtsuntreue“ (Kindhäuser 1994, S. 211) legitimiert die Nichterfüllung der Sorgfaltspflicht die Erhebung eines Vorwurfes in gewissen Fällen also auch in Abwesenheit eines Vorsatzes. Dem hierbei angesetzten Sorgfaltsmaßstab liegen die an einen rechtstreuen Normadressaten gesetzten Erwartungen zugrunde, der im Rahmen seiner Möglichkeiten Normbrüche vermeiden soll. Diese „Kategorie der Pflicht“ enthält also die Information darüber, in welchem Maß der Rechtsadressat „seine Kapazitäten in den Dienst der Normrealisierung zu stellen hat“ (Kindhäuser 1994, S. 198). Als nachgelagertes System beurteilt das Strafrecht den anzuwendenden Sorgfaltsmaßstab *ex post*. Die Herausforderung im Zusammenhang mit intelligenten Agenten liegt nun darin, dass es in diesem Bereich an einer bisherigen rechtlichen Praxis mangelt, auf die man ohne Weiteres verweisen könnte. Es ist deshalb die dringliche Aufgabe der Rechtswissenschaft, dennoch möglichst klare Leitplanken vorzugeben. Ansonsten würde strafrechtliche Verantwortlichkeit zur Lotterie, was weder für die Glaubwürdigkeit des Rechtssystems noch für den technischen Fortschritt eine wünschenswerte Ausgangslage wäre.

Die Delegation von Aufgaben an Technik ist eine bewusste Entscheidung. Bei der Implementierung soziotechnischer Systeme ist eine Balance der Aufgabenteilung zwischen Mensch und Maschine herzustellen. Dabei sind Kompetenzen, Verantwortlichkeiten und Kontrollmöglichkeiten adäquat zu verteilen (Flemisch et al. 2012). Für die Konkretisierung der Sorgfaltspflichten wird es entscheidend sein, inwieweit die Gesellschaft bereit ist, Risiken in Kauf zu nehmen. Ein Null-Risiko-Ansatz würde dazu führen, dass Entwickler ihre Projekte aufgeben müssten, wollen sie sich keiner Globalverantwortlichkeit ausgesetzt sehen (Simmler und Markwalder 2017). Eine solche Globalverantwortlichkeit wäre jedoch kaum mit dem Erfordernis der persönlichen Vorwerfbarkeit vereinbar. Andererseits stellen drohende „Verantwortlichkeitslücken“ ebenso eine Gefahr für die Normstabilität dar (Matthias 2004). In diesem Spannungsfeld gilt es nun eine strafrechtliche Fahrlässigkeitsdogmatik zu entwickeln, deren zugrunde liegender Sorgfaltsmaßstab eine funktionale Handlungs-zurechnung ermöglicht.

7 Sorgfaltsmaßstab und technische Autonomie

In allen deutschsprachigen Rechtsordnungen wird Fahrlässigkeit gleichermaßen definiert: So begeht ein Delikt fahrlässig, wer die Folge seines Verhaltens aus pflichtwidriger Unvorsichtigkeit nicht bedenkt oder darauf nicht Rücksicht nimmt (Art. 12 Abs. 3 des CH-StGB), wer also die Sorgfalt außer Acht lässt, zu der er verpflichtet und befähigt gewesen wäre (§ 6 Abs. 1 des Ö-StGB). Der Sorgfaltsmaßstab im Strafrecht ist ein *individueller*, d. h. eine Unvorsichtigkeit ist nur vorwerfbar, wenn der Täter die Vorsicht nicht beachtet, zu der er nach den konkreten Umständen und seinen persönlichen Verhältnissen verpflichtet ist. Das Strafrecht verlangt also von niemandem mehr als dieser persönlich zu leisten in der Lage ist (Seelmann und Geth 2016). Ansonsten wäre eine Bestrafung dysfunktional.

Soll der Fahrlässigkeitsvorwurf funktional sein, d. h. dem Schuldprinzip folgen (Simmler 2018), können Handlungen nicht willkürlich zugerechnet werden. Auch technische Fehlleistungen werden dem Entwickler dementsprechend nur zugerechnet, wenn ihm gegenüber ein persönlicher Vorwurf erhoben werden kann. Verantwortlichkeit wird einerseits dadurch eingeschränkt, was eine Person weiß und andererseits dadurch, was sie kontrollieren kann (vgl. Nida-Rümelin und Battaglia 2018); oder in der Sprache einer Strafrechtlerin: durch die *Voraussehbarkeit* und die *Vermeidbarkeit*. So kann jemandem eine Pflichtwidrigkeit nur vorgehalten werden, wenn die Gefährdung des Rechtsguts und die Möglichkeit der Tatbestandsverwirklichung vorausgesehen werden konnte. Ebenso muss es dem Täter überhaupt möglich gewesen sein, die Tat zu verhindern. Es sind genau diese Voraussetzungen, die in Bezug auf intelligente Agenten problematisch werden. Während sich die Zurechnung zum fahrlässig Handelnden darauf fokussiert, ihm das Nichtbefolgen der Sorgfaltspflicht *trotz* ausreichender Kontrolle über den Handlungsablauf vorzuwerfen, lässt gerade die zunehmende Autonomie avancierter Technik diese Kontrolle erodieren. Zusammengefasst: Umso autonomer die Technik wird, desto schwieriger wird es, ihr Verhalten dem Menschen zuzurechnen.

Autonomie ist eine graduelle Eigenschaft (vgl. Loh 2018a). Es können verschiedene Level technischer Autonomie unterschieden werden (Simmler und Frischknecht 2019), die sich in soziotechnischen Handlungskontexten wiederum direkt auf die Handlungsträgerschaft der Technik auswirken. Hybride Systeme sind allem voran durch zwei Faktoren charakterisiert: Durch die Aufgabenteilung zwischen Mensch und Maschine (*Automatisierungslevel*) und durch den Freiheitsgrad der Technik bei der Erfüllung dieser Aufgabe (*Autonomielevel*) (Simmler und Frischknecht 2019). Die Ausgestaltung dieser Faktoren determiniert die Verteilung der Handlungsträgerschaft, die Kontrollmöglichkeiten und damit die Verantwortlichkeit der Akteure. Es stellt sich allerdings die Frage, ob dieser Umstand die Verantwortung des Menschen schlichtweg *vorverlagert*. Ist die autonome Maschine einmal implementiert, kann ihr Verhalten nur noch beschränkt gesteuert werden. Dies erhöht die Verantwortung im Vorfeld der Implementierung. Das Sorgfaltsmaß steigt und es findet eine Verschiebung von der unmittelbaren Handlungsverantwortung zu einer vorgelagerten „Designverantwortung“ (Maring 2008) statt. Dieser positive Zusammenhang zwischen technischer Autonomie und Sorgfaltsmaßstab bedeutet folglich,

dass ein hohes Level an Autonomie den Menschen nicht aus seiner Verantwortung entlässt. Nicht jedes unvorhersehbare Verhalten der Maschine kann dem menschlichen (Co-)Akteur zugerechnet werden. Der Verweis auf die Autonomie der Technik als neue Form der Exkulpation funktioniert allerdings nur, wenn die Schaffung dieses Grades an Autonomie sorgfältig vonstattenging.

Was die *sorgfältige* Schaffung maschineller Autonomie bedeutet, gilt es zu präzisieren. Es bedeutet nicht, dass keinerlei Risiken eingegangen werden dürfen. Das Strafrecht kennt – speziell im Bereich technischer Neuerungen – auch die Formel des „erlaubten Risikos“ (Killias et al. 2017; Hoyer 2009). Bei der Bestimmung des Bereichs erlaubter Risiken wird u. a. auf „allgemein anerkannte Regeln der Technik“, auf den „Stand der Technik“ oder auf den „Stand der Wissenschaft“ verwiesen (Hoyer 2009). Wie diese Klauseln bereits indizieren, ist das Ziehen der Grenze einer allzu großen Generalisierbarkeit kaum zugänglich. Den Inhalt der Sorgfaltspflicht auszumachen, erfordert die schwierige Abwägung zwischen dem erwarteten Nutzen einer bestimmten Tätigkeit und den mit ihr einhergehenden Risiken (Güngerich und Walpen 2011), wobei der Grad der geforderten Vorsicht insbesondere von der Nähe der Gefahr und vom Wert des gefährdeten Rechtsguts abhängt (Seelmann und Geth 2016). Das Erkennen der Sorgfaltspflichten bedarf folglich „prognostischer Elemente“ (Kindhäuser 1994, S. 210).

Die Voraussehbarkeit muss im Moment der Tathandlung gegeben sein, wobei als Tathandlung nicht nur die letzte Handlung des Täters vor dem Erfolgseintritt infrage kommt (Jakobs 1991). Vielmehr ist es oftmals so, dass „der Täter sich selbst fahrlässig als unvermeidbar handelndes Werkzeug entlässt“ (Jakobs 1991, 9/14). Damit verlagert sich der Handlungszeitpunkt vor. Den Akteur trifft eine Übernahmeverantwortlichkeit für die Konsequenzen seiner fahrlässigen Kontrollaufgabe. Sie werden ihm als Handlung zugerechnet. Es stellt sich nun die Frage, ob sich auch im Falle potenziell „fahrlässiger Delegationen“ an technische Agenten eine solche Übernahmeverantwortlichkeit ergibt. Eine solche Zurechnung der Folgen maschineller Autonomie ist denkbar, wenn bereits die Schaffung dieses Freiheitsraums unsorgfältig und damit vorwerfbar geschehen ist. Als Anhaltspunkt für diese Beurteilung könnten z. B. die im Haftpflichtrecht gebräuchlichen drei *curae* dienen. So wäre eine Delegation an eine Maschine dann fahrlässig, wenn die *cura in eligendo* (Sorgfalt in der Auswahl), die *cura in instruendo* (Sorgfalt in der Unterweisung) und die *cura in custodiendo* (Sorgfalt in der Überwachung) nicht gewahrt blieben. Dazu wird teilweise noch eine vierte Pflicht konstatiert, die *cura in organisando* (Sorgfalt in der Organisation) (siehe z. B. Schweizer BGE 110 II 456).

Als Zwischenfazit lässt sich festhalten, dass die Emergenz avancierter Maschinen den Menschen nicht aus der Pflicht entlässt: Technische Autonomie entschuldigt nicht. Vielmehr verpflichtet sie zu erhöhter Wachsamkeit bei der *Schaffung* dieser Autonomie. In Befolgung dieses Ansatzes ist bei der Implementierung von Technik neben der regulären Risikoevaluation zunächst auszumachen, ob sich der fragliche Bereich überhaupt für Automatisierung eignet, diese Aufgaben also überhaupt delegiert werden dürfen. In der Folge muss das Mensch-Maschine-Zusammenwirken angemessen überwacht und organisiert werden, wobei Organisation insbesondere das Festlegen eines sinnhaften Ausmaßes an Automatisierung und technischer

Autonomie bedeutet. Die Technik muss ferner angemessen unterwiesen werden. Die Maschinenethik kann nun wertvolle Hinweise liefern, was eine ethische und normgemäße Erfüllung einer solchen *curae in instruendo* bedeuten könnte.

8 Maschinelle Moral als Sorgfaltspflicht

Technische Systeme sind wertgeladen. Sie schaffen moralische Konsequenzen und haben das Potenzial, ethische Prinzipien zu untergraben (Martin 2018). Die Delegation einer Handlung an eine Maschine stellt immer auch eine moralische Delegation dar, für die zunächst die Entwicklerinnen in der Verantwortung stehen. Die Maschinenethik adressiert deshalb nicht nur die Maschinen selbst, sondern auch „deren Schöpfer auf allen Ebenen der Gestaltung“ (Weber und Zoglauer 2018). Das Kernstück dieser Disziplin – die eigentliche „Moralimplementation“ (Misselhorn 2018b) – gewinnt deshalb für das Strafrecht nicht nur im Falle der bereits diskutierten Möglichkeit originärer Subjektqualität Relevanz. Sie berührt ferner die Verantwortlichkeit ebendieser menschlichen „Schöpfer“. So verlangt es die Verbreitung immer autonomerer Maschinen, dass diese zu moralischem Verhalten befähigt werden. Wer Maschinen Freiheitsräume einräumt, muss sichergehen, dass die damit entstehende *Agency* von Technik auch eine moralische ist. Die bewusste Auseinandersetzung mit maschineller Moral wird damit zur Sorgfaltspflicht, deren Vernachlässigung einen Fahrlässigkeitsvorwurf zu legitimieren vermag.

Eine solche „Pflicht zur maschinellen Moral“ ist wiederum zu konkretisieren. Sie ist je nach Einsatzbereich variabel und steigt bei höheren Freiheitsgraden der Technik. Die Pflicht zur Implementierung prosozialen, moralischen und normkonformen Verhaltens hängt insbesondere von der Ausgereiftheit der Maschine ab (Chinen 2016). Je autonomer eine Maschine ist, desto notwendiger ist die Berücksichtigung dieser Faktoren bereits in frühen Entwicklungsstadien. Vor allem in Bereichen von hoher moralischer Relevanz muss eine solche normative Konstruktion von Technik stets auch ein maschinenethisches Unterfangen sein, wobei die Berücksichtigung dieser Faktoren zum Verantwortungsbereich der Entwicklerinnen gehört. Sowohl bei der Moralität (vgl. Loh 2018b) als auch bei der strafrechtlichen Normtreue verbleibt die Kompetenz und die dazugehörige Kompetenzkompetenz folglich zunächst bei den menschlichen Designern und bei anderweitig am sozio-technischen System beteiligten Individuen.

Das Vorhaben der Maschinenethik – die Schaffung der Möglichkeit, moralische Maschinen zu designen – unterminiert also am Schluss nicht die Verantwortung der Menschen und verlagert sie auf die Technik. Ganz im Gegenteil entsteht gerade aufgrund der durch sie geschaffenen erweiterten Möglichkeiten eine zusätzliche Pflicht für Entwickler. Ihr Verantwortungsbereich erweitert sich auf eine Pflicht zur sorgfältigen und bedachten Moralimplementation, d. h. geradewegs zur Schaffung maschineller Moral.

9 Zusammenspiel von Recht, Technik und Maschinenethik

Es ist naheliegend, dass die Transformationen des digitalen Zeitalters und deren weitreichenden Konsequenzen für die Gesellschaft auch das Strafrecht und die soziale Praxis rund um dessen Verantwortungszuschreibung prägen. Dem Recht kommt eine entscheidende Rolle dabei zu, dem technischen Fortschritt Rahmen und Grenzen vorzugeben (Wein 1992). Es hat den „Siegeszug der Maschinen von Anfang an begleitet und mitgestaltet“ (Hilgendorf 2015, S. 13). Wie die Gesellschaft und mit ihr das Strafrecht die Verantwortlichkeit im soziotechnischen Zusammenspiel verteilen, wird für die zukünftige technologische Entwicklung also nicht wirkungslos bleiben (Markwalder und Simmler 2017).

Aus strafrechtlicher und kriminalpolitischer Sicht stellt sich beim Aufkommen neuer Technologien die Frage, was für Probleme sich für das gemeinschaftliche Zusammenleben und für die Stabilität der durch das Strafrecht zu schützenden Normen ergeben können. So betrifft technischer Fortschritt stets die soziale Interaktion der Gesellschaftsmitglieder und entfaltet so Wirkung auf die „normative Konstruktion der Gesellschaft“ (Popitz 1980). Das Strafrecht ist aufgrund seines spezifischen Instrumentariums besonders dazu geeignet, soziale Ordnung und Verhaltensabstimmung zu sichern und gewisse Normen als *ultima ratio* einem speziellen Schutz zu unterstellen (Simmler 2018). Dies gilt gerade im Kontext technologischer Entwicklung. Es ist dabei die Aufgabe der Strafrechtswissenschaft, diejenigen Bereiche auszumachen, in denen die kommunikative Kraft der Strafe nachhaltig zur Stabilität der Gesellschaft beitragen kann (Beck 2017).

Diese Ausgangslage verlangt es, dass Juristinnen, Informatikerinnen und Ethikerinnen näher zusammenrücken (vgl. Seidel 2017). Die Aufgabe der Vertreterinnen der Rechtswissenschaft beschränkt sich dabei keineswegs auf Fragen der Regulierung von Technik, der Entwicklung von Spezialgesetzen oder Standards. Vielmehr sind es gerade rechtliche Grundlagenfragen, welche die Disziplin im Zusammenhang mit avancierter Technologie herausfordern (Hilgendorf 2015). Diese elementaren Fragen rund um die strafrechtliche Handlungs- und Schuldlehre weisen dabei einen engen Zusammenhang zum philosophischen und ethischen Grundgerüst des Rechts auf. Gleichfalls verkörpern die Entscheidungen, die beim Design von Technik getroffen werden, Wertentscheidungen. Damit haben sie sowohl ethische als auch oftmals rechtliche Implikationen. Die Rechtstheorie wird im digitalen Zeitalter somit zu einem Grundpfeiler der Maschinenethik – und die Maschinenethik zu einem Grundpfeiler der Rechtstheorie.

10 Resümee

Dieser Beitrag hat sich dem Zusammenhang von Maschinenethik und Strafrecht angenommen und aufgezeigt, dass diese beiden Disziplinen eine unweigerliche Nähe aufweisen. Autonomisierung der Technik und Hybridisierung der Gesellschaft wirken sich sowohl auf die moralische als auch auf die strafrechtliche Handlungsfähigkeit und damit Verantwortung aus. Die Emergenz künstlicher moralischer

Akteure zieht einerseits die Frage nach sich, ob diese strafrechtliche Subjektqualität erlangen können. Andererseits führt sie – möge sich die technische Handlungsfähigkeit graduell auch von derjenigen der Menschen unterscheiden – unweigerlich zur Frage, wie technische Handlungsfähigkeit die Verantwortlichkeit der Menschen berührt.

Es konnte festgestellt werden, dass Maschinen aktuell keine strafrechtliche Subjektqualität zukommt. Auch wenn sie bis zu einem gewissen Grad Subjekte der Moral darstellen können, erfüllen sie die Ansprüche des strafrechtlichen Personenbegriffs nicht, setzt der strafrechtliche Schuldvorwurf doch einen sehr hohen Grad an Autonomie und Handlungsträgerschaft voraus. Es wurde gleichfalls festgehalten, dass dieser Befund sich wandeln könnte. Die strafrechtliche Verantwortungszuschreibung adressiert nicht Menschen, sondern Personen als Objekte der Zurechnung. In Anbetracht des rasanten technischen Fortschritts kann nicht ausgeschlossen werden, dass zukünftige Maschinen diesen Personenstatus irgendwann erreichen könnten. Solange dies nicht der Fall ist, steht aber zweifellos die Verantwortlichkeit des Menschen im Zentrum. Ihr soll deshalb das primäre wissenschaftliche Interesse gelten.

Wie in diesem Beitrag aufgezeigt wurde, wird die Verantwortung des Menschen folglich nicht per se unterminiert. Es ist sich deshalb der Frage anzunehmen, wann und inwieweit Menschen für (teil-)autonome Maschinen zur Rechenschaft zu ziehen sind. Insbesondere die Erhebung eines Fahrlässigkeitsvorwurfs bereitet bei Vorliegen fortgeschrittener Autonomiestufen dogmatische Schwierigkeiten. Dieser konstituiert sich durch eine Sorgfaltspflichtwidrigkeit, weshalb für eine funktionale Handlungszurechnung die Fixierung eines adäquaten Sorgfaltsmaßstabs notwendig wird. Beim Design von Technik hängt dieser primär von der technischen Autonomie ab. Umso autonomer und komplexer das technische System, desto höher ist der Anspruch an die Vorsicht im Vorfeld der Implementierung.

Die Sorgfaltspflichten gilt es in Zukunft weiter zu konkretisieren. Es hat sich aber gezeigt, dass insbesondere die Maschinenethik zur Präzisierung des Pflichtinhalts beitragen kann. Darüber hinaus können die Auseinandersetzung mit der maschinellen Moral und deren Implementierung selbst zur Sorgfaltspflicht werden, wäre es doch gerade bei avancierten Maschinen fahrlässig, diese Möglichkeiten zu vernachlässigen. Entsprechend bestätigte sich in dieser Abhandlung, dass die Maschinenethik für das Strafrecht fruchtbar gemacht werden kann. Das Zusammenspiel von Recht, Informatik und Ethik offenbarte sich als evident und dessen Intensivierung im Zeitalter autonomer Maschinen als wissenschaftliche Pflicht.

Literatur

- Beck, Susanne. 2013. Über Sinn und Unsinn von Statusfragen – zu Vor- und Nachteilen der Einführung einer elektronischen Person. In *Robotik und Gesetzgebung*, Hrsg. Eric Hilgendorf und Jan-Philipp Günther, 239–261. Baden-Baden: Nomos.
- Beck, Susanne. 2015. Technisierung des Menschen – Vermenschlichung der Technik. Neue Herausforderungen für das rechtliche Konzept „Verantwortung“. In *Autonome Automaten: Künst-*

- liche Körper und artifizielle Agenten in der technisierten Gesellschaft, Hrsg. Malte-Christian Gruber, Jochen Bung und Sascha Ziemann, 2. Aufl., 173–187. Berlin: BWV.
- Beck, Susanne. 2017. Google cars, software agents, autonomous weapons systems – New challenges for criminal law? In *Robotics, autonomies, and the law*, Hrsg. Eric Hilgendorf und Uwe Seidel, 227–252. Baden-Baden: Nomos.
- Bendel, Oliver. 2018a. Überlegungen zur Disziplin der Maschinenethik. *APuZ* 6–8:34–38.
- Bendel, Oliver. 2018b. Wozu brauchen wir die Maschinenethik? In *Handbuch Maschinenethik*, Oliver Bendel. Wiesbaden: Springer.
- Bendel, Oliver. 2019. Moralische Maschinen. In *Beitrag für das Gabler Wirtschaftslexikon*. Wiesbaden: Springer Gabler. <https://wirtschaftslexikon.gabler.de/definition/moralische-maschinen-119940/version-369160>. Zugriffen am 31.01.2019.
- Castelfranchi, Christiano, und Rino Falcone. 2004. The dialectics between (social) environment and agent's architecture and powers. In *Agents and computation autonomy*, Hrsg. Matthias Nickles, Michael Rovatsos und Gerhard Weiss, 40–54. Berlin/Heidelberg: Springer.
- Chinen, Mark. 2016. The co-evolution of autonomous machines and legal responsibility. *Virginia Journal of Law & Technology* 20(2): 338–393.
- Durkheim, Émile. 1893. *De la division du travail social*. Paris: Félix Alcan.
- Flemisch, Frank, Matthias Heesen, Tobias Hesse, Johann Kelsch, Anna Schieben, und Johannes Beller. 2012. Towards a dynamic balance between humans and automation: Authority, ability, responsibility and control in shared and cooperative control situations. *Cognition, Technology & Work* 14:4–18.
- Floridi, Luciano, und J. W. Sanders. 2004. On the morality of artificial agency. *Minds and Machine* 14:349–379.
- Gless, Sabine, und Thomas Weigend. 2014. Intelligente Agenten und das Strafrecht. *ZStW* 126: 561–591.
- Gless, Sabine, Emily Silverman, und Thomas Weigend. 2016. If Robots cause harm, Who ist to blame? Self-Driving Cars and Criminal Liability. *New Criminal Law Review* 19(3): 412–436.
- Gruber, Malte-Christian. 2012. Rechtssubjekte und Teilrechtssubjekte des elektronischen Geschäftsverkehrs. In *Jenseits von Mensch und Maschine*, Hrsg. Susanne Beck, 133–159. Baden-Baden: Nomos.
- Günther, Andreas, und Adrian Walpen. 2011. Rechtliche Aspekte eines risiko- und effizienz-basierten Sicherheitskonzeptes. *Sicherheit & Recht* 2:105–117.
- Günther, Jan-Philipp. 2015. Embodied Robots – Zeit für eine rechtliche Neubewertung? In *Autonome Automaten: Künstliche Körper und artifizielle Agenten in der technisierten Gesellschaft*, Hrsg. Malte-Christian Gruber, Jochen Bung und Sascha Ziemann, 2. Aufl., 155–172. Berlin: BWV.
- Hallevy, Gabriel. 2010. The criminal liability of artificial intelligence entities – From science fiction to legal social control. *Akron Intellectual Property Journal* 4:171–201.
- Hilgendorf, Eric. 2015. Recht und autonome Maschinen – ein Problemaufriß. In *Das Recht vor den Herausforderungen der modernen Technik*, Hrsg. Eric Hilgendorf und Sven Hötzsch, 11–40. Baden-Baden: Nomos.
- Hoyer, Andreas. 2009. Erlaubtes Risiko und technologische Entwicklung. *ZStW* 121:860–881.
- Jakobs, Günther. 1991. *Strafrecht Allgemeiner Teil, Die Grundlagen und die Zurechnungslehre*, 2. Aufl. Berlin: de Gruyter.
- Jakobs, Günther. 1993. *Das Schuldprinzip*. Opladen: Westdeutscher Verlag.
- Killias, Martin, Nora Markwalder, André Kuhn, und Nathalie Dongois. 2017. *Grundriss des Allgemeinen Teils des Schweizerischen Strafgesetzbuches*, 2. Aufl. Bern: Stämpfli.
- Kindhäuser, Urs. 1994. Erlaubtes Risiko und Sorgfaltswidrigkeit. Zur Struktur strafrechtlicher Fahrlässigkeitshaftung. *Goltdammer's Archiv für Strafrecht* 141(5): 197–223.
- Loh, Janina. 2018a. Maschinenethik und Roboterethik. In *Handbuch Maschinenethik*, Hrsg. Oliver Bendel. Wiesbaden: Springer.
- Loh, Janina. 2018b. Verantwortung und Roboterethik – ein kleiner Überblick – Teil 2. *InTeR* 1: 29–35.

- Loh, Wulf, und Janina Loh. 2017. Autonomy and responsibility in hybrid systems: The example of autonomous cars. In *Robot ethics 2.0: From autonomous cars to artificial intelligence*, Hrsg. Patrick Lin, Keith Abney und Ryan Jenkins, 35–50. Oxford: Oxford University Press.
- Maring, Matthias. 2008. Mensch-Maschine-Interaktion: Steuerbarkeit – Verantwortbarkeit. In *Maschinen, die unsere Brüder werden. Mensch-Maschine-Interaktion in hybriden Systemen*, Hrsg. Christoph Hubig und Peter Koslowski, 113–130. München: Wilhelm Fink.
- Markwalder, Nora, und Monika Simmler. 2017. Roboterstrafrecht – Zur strafrechtlichen Verantwortlichkeit von Robotern und künstlicher Intelligenz. *AJP* 2:171–182.
- Martin, Kirsten. 2018. Ethical implications and accountability of algorithms. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-018-3921-3>.
- Matthias, Andreas. 2004. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology* 6:175–183.
- Misselhorn, Catrin. 2018a. Maschinenethik und Philosophie. In *Handbuch Maschinenethik*, Hrsg. Oliver Bendel. Wiesbaden: Springer.
- Misselhorn, Catrin. 2018b. *Grundfragen der Maschinenethik*, 2. Aufl. Ditzingen: Reclam.
- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, und Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(1): 1–21.
- Müller-Hengstenberg, Claus D., und Stefan Kim. 2016. *Rechtliche Risiken autonomer und vernetzter Systeme. Eine Herausforderung*. Berlin: de Gruyter.
- Nida-Rümelin, Julian. 2017. Handlung, Technologie und Verantwortung. In *Berechenbarkeit der Welt?* Hrsg. Wolfgang Pietsch, Jörg Wernecke und Maximilian Ott, 497–513. Wiesbaden: Springer.
- Nida-Rümelin, Julian, und Fiorella Battaglia. 2018. Mensch, Maschine und Verantwortung. In *Handbuch Maschinenethik*, Hrsg. Oliver Bendel. Wiesbaden: Springer.
- Popitz, Heinrich. 1980. *Die normative Konstruktion von Gesellschaft*. Tübingen: Mohr.
- Rammert, Werner. 2003. Technik in Aktion: Verteiltes Handeln in sozio-technischen Konstellationen. In *Autonome Maschinen*, Hrsg. Thomas Christaller und Josef Wehner, 298–315. Wiesbaden: Westdeutscher Verlag.
- Rammert, Werner, und Ingo Schulz-Schaeffer. 2002. Technik und Handeln – wenn soziales Handeln sich auf menschliches Verhalten und technische Artefakte verteilt. *TUTS – Working Papers* 4/2002, Berlin.
- Rath, Matthias. 2019. Zur Verantwortungsfähigkeit künstlicher „moralischer Akteure“. Problem-anzeige oder Ablenkungsmanöver? In *Maschinenethik: Normative Grenzen autonomer Systeme*, Hrsg. Matthias Rath, Friedrich Kotz und Matthias Karmasin, 223–242. Wiesbaden: Springer.
- Seelmann, Kurt, und Christopher Geth. 2016. *Strafrecht Allgemeiner Teil*, 6. Aufl. Basel: Helbing Lichtenhahn.
- Seher, Gerhard. 2016. Intelligente Agenten als „Personen“ im Strafrecht. In *Intelligente Agenten und das Recht*, Hrsg. Sabine Gless und Kurt Seelman, 45–60. Baden-Baden: Nomos.
- Seidel, Uwe. 2017. Industry 4.0 and law – Experiences from AUTONOMICS. In *Robotics, autonomics, and the law*, Hrsg. Eric Hilgendorf und Uwe Seidel, 11–26. Baden-Baden: Nomos.
- Seng, Leonie. 2018. Maschinenethik und Künstliche Intelligenz. In *Handbuch Maschinenethik*, Hrsg. Oliver Bendel. Wiesbaden: Springer.
- Simmler, Monika. 2018. *Normstabilisierung und Schuldvorwurf*. Bern/Berlin: Carl Grossmann Verlag.
- Simmler, Monika, und Ruth Frischknecht. 2019. *Capturing distributed agency in socio-technical systems: An interdisciplinary taxonomy of automation and autonomy in human-machine interaction*. Unveröffentlichtes Manuskript: 2018 zur Publikation eingereicht.
- Simmler, Monika, und Nora Markwalder. 2017. Roboter in der Verantwortung? – Zur Neuauflage der Debatte um den funktionalen Schuldbegriff. *ZStW* 129:20–47.
- Verhagen, Harko. 2014. Autonomy and reasoning for natural and artificial agents. In *Agents and computation autonomy*, Hrsg. Matthias Nickles, Michael Rovatsos und Gerhard Weiss, 83–93. Berlin/Heidelberg: Springer.

- Wallach, Wendel, und Colin Allen. 2009. *Moral machines, teaching robots right from wrong*. Oxford: Oxford University Press.
- Weber, Karsten, und Thomas Zoglauer. 2018. Maschinenethik und Technikethik. In *Handbuch Maschinenethik*, Hrsg. Oliver Bendel. Wiesbaden: Springer.
- Wein, Leon E. 1992. The responsibility of intelligent artifacts: Toward an automation jurisprudence. *Harvard Journal of Law & Technology* 6(1): 103–154.
- Weyer, Johannes, und Simone Reineke. 2005. Creating order in hybrid systems: Reflexions on the interaction of man and smart machines. In *Soziologische Arbeitspapiere 7*, Hrsg. Technische Universität Dortmund, Wirtschafts- und Sozialwissenschaftliche Fakultät. Dortmund: Universität Dortmund.