

Prof. Dr. Sabine Gleß und Prof. Dr. Thomas Weigend
Intelligente Agenten und das Strafrecht

DOI 10.1515/zstw-2014-0024

I. Intelligente Agenten – Potential und Risiko

„*Schuld ist der Algorithmus*“, meinte die Presse¹. Die Ehefrau eines früheren Bundespräsidenten hatte sich rechtlich dagegen zur Wehr gesetzt, dass ihr Name, wenn man ihn bei „google-search“ eingab, automatisch mit Begriffen wie „Prostitution“ oder „Escort Service“ verbunden wurde. Diese Ergänzungen, die die Klägerin als beleidigend empfand, beruhten allerdings nicht auf individuellen Entscheidungen von Personen im Bereich der beklagten Firma, sondern auf den Funktionen „google-bot“ und „google-autocomplete“, die nach allgemein festgelegten Handlungsvorschriften Nutzeranfragen verarbeiten². Beide sind Beispiele für so genannte Intelligente Agenten³: Die Funktionen operieren nach bestimmten vorgegebenen Regeln, aber sie verarbeiten die Informationen (das Suchverhalten der Nutzer von Google) in jedem Einzelfall selbständig.

Intelligente Agenten werden – in ihren unterschiedlichen Formen⁴ – dort eingesetzt, wo die rasche Verarbeitung sehr vieler Informationen präzise Kombination und schnelle Reaktion erfordert, oder auch dort, wo der Einsatz physischer Kräfte gefordert ist, die die Möglichkeiten von Menschen übersteigen. Intelligente Agenten gibt es heute schon in vielen Lebensbereichen: Sie bestimmen – in der recht einfachen Form eines Softwareagenten – als Suchmaschinen im

1 „Die Welt“ v. 10.09.2012, http://www.welt.de/welt_kompakt/print_politik/article109114305 (besucht am 18.07.2014).

2 Zu der Frage der zivilrechtlichen Haftung siehe BGH NJW 2013, 2348; *Gounalakis*, NJW 2013, 2321; *Hager*, JA 2013, 630; *Mäsch*, JuS 2013, 814.

3 *Russell/Norvig*, Artificial Intelligence. A Modern Approach, 3. Aufl. 2014, S. 34 ff.

4 Als Synonym für den Oberbegriff Intelligente oder Autonome Agenten benutzt man vielfach auch den Begriff „Roboter“; siehe etwa *Beck*, JR 2009, 226; *dies.*, in: *Beck* (Hrsg.), *Jenseits von Mensch und Maschine*, 2012, S. 15. Gebräuchliche Unterbegriffe für verschiedene Formen der Intelligenzen Agenten sind z. B. „Softwareagenten“, „Android“ oder „Drohne“.

Sabine Gleß: Ordinaria für Strafrecht und Strafprozessrecht an der Juristischen Fakultät der Universität Basel

Thomas Weigend: Professor und Vorstand des Instituts für ausländisches und internationales Strafrecht der Rechtswissenschaftlichen Fakultät der Universität zu Köln

Internet unseren Zugang zu Informationen und damit in gewissem Umfang unsere Wahrnehmung der Realität⁵; „lesende Softwareagenten“ treffen im Online-Börsenhandel aufgrund bestimmter Meldungen autonome Kauf- oder Verkaufentscheidungen⁶. Komplexere Systeme wie „intelligente Skalpelle“ und „Fräsroboter“ führen medizinische Operationen durch⁷, und intelligente Flugkörper („Drohnen“) werden nicht nur für militärische Zwecke⁸, sondern auch zur Überwachung und Sicherung sensibler ziviler Bereiche und vielleicht in naher Zukunft auch zur Lieferung von Waren eingesetzt⁹. Am sichtbarsten in das Bewusstsein der Öffentlichkeit rücken Intelligente Agenten derzeit wohl durch Bemühungen der Industrie, fahrerlose Autos zum Einsatz auf öffentlichen Straßen zu fertigen.¹⁰

Die Voraussetzungen dafür, dass eine Maschine als Intelligenter Agent bezeichnet werden kann, sind in der Informatik ebenso umstritten wie der Inhalt des Begriffs Intelligenz¹¹. Für eine Bewertung aus rechtlicher Sicht erscheint es sinnvoll, Maschinen nach ihrer technischen Machart in drei Arten von Systemen zu unterteilen: (a) einfache Datenverarbeitungssysteme, wie etwa Taschenrechner, bei denen der Entwickler des Systems die Art der möglichen Datenaufnahme wie auch deren Verarbeitung auf der Grundlage hierarchisch definierter Algorithmen abschließend festgelegt hat; (b) geöffnete Systeme, wie etwa automatisierte

5 Siehe dazu etwa *Egermann*, Suchmaschinen, in: *Kilian/Heussen*, Computerrecht, Stand 2013, Rdn. 7–10.

6 *Bram/Stalknecht*, Handelsblatt Online v. 24.04.2013, <http://www.handelsblatt.com/finanzen/boerse-maerkte/boerse-inside/blitzabsturz-an-der-boerse-maschinen-an-der-macht/8113842.html> (besucht am 18.07.2014).

7 *Hilgendorf*, in: *Rotsch* (Hrsg.), Criminal Compliance vor den Aufgaben der Zukunft, 2013, S. 24 f. Zur Zulässigkeit des Einsatzes von „Fräsrobotern“ bei Gelenkoperationen siehe die zivilrechtlichen Urteile OLG Dresden NJOZ 2008, 247 und OLG Frankfurt a. M., Urt. v. 08.11.2013–25 U 79/12 (<http://openjur.de/u/668787.html>).

8 Die Zulässigkeit des Einsatzes von Drohnen zur militärischen Nutzung ist unter dem Gesichtspunkt des Humanitären Völkerrechts umstritten; siehe dazu etwa *Asaro*, International Review of the Red Cross 94 (2012), S. 687; *Heller*, Journal of International Criminal Justice 11 (2013), S. 89; *May*, Journal of International Criminal Justice 11 (2013), S. 47; *Pagallo*, The Laws of Robots, 2013, S. 51 f.; *Sparrow*, Journal of Applied Philosophy 24 (2007), S. 62.

9 Siehe etwa <http://www.amazon.com/b?node=8037720011>.

10 Siehe etwa <http://www.google.ch/about/careers/lifeatgoogle/self-driving-car-test-steve-mahan.html>.

11 *Alan Turing* hat dazu 1950 folgenden Test entwickelt: Ein menschlicher Fragesteller führt über eine Tastatur und einen Bildschirm ohne Sicht- und Hörkontakt mit zwei ihm unbekanntem Gegenübern eine Unterhaltung; der eine ist ein Mensch, der andere eine Maschine. Beide versuchen, den Fragesteller davon zu überzeugen, dass sie denkende Menschen sind. Wenn der Fragesteller am Ende nicht sagen kann, welcher von beiden die Maschine ist, hat die Maschine den Turing-Test bestanden. Siehe dazu *Alan Turing*, Computing Machinery and Intelligence, Mind 59 (1950), S. 433.

Rasenmäher, die – über Sensoren – unterschiedliche Daten aus der Außenwelt aufnehmen können, diese jedoch nach Regeln verarbeiten, die der Entwickler unveränderbar festgelegt hat; (c) intelligente Systeme, die große Mengen von Daten nicht nur aufnehmen, sondern auch auf bestimmte Muster hin interpretieren und selbständig auf die verarbeiteten Daten reagieren. Geöffnete und intelligente Systeme handeln insofern autonom, als sie sensorisch ihre Umgebung analysieren und verarbeiten. Aber nur intelligente Systeme sind darüber hinaus in der Lage, auf die aufgenommenen und von ihnen interpretierten Informationen unmittelbar zu reagieren und sie außerdem in einer Weise abzuspeichern, die es erlaubt, darauf für spätere Entscheidungsprozesse zurückzugreifen. Das kann auf sehr einfacher Stufe funktionieren, wie etwa bei der Internetsuchmaschine, die alle Nutzeranfragen speichert, die meistgefragten Kombinationen anbietet und sich merkt, was akzeptiert wird. Manche Intelligente Agenten müssen aber auch sehr komplexe Entscheidungen treffen, wie etwa ein selbstfahrendes Auto, das eine verkehrsreiche, durch Verkehrsschilder geregelte Kreuzung in Interaktion mit menschlichen Fahrern überqueren muss.

Auch ein sehr hoch entwickelter Intelligenter Agent ist allerdings immer nur so intelligent wie seine Programmierung. Die ihm aufgrund von umfangreichen Simulationen mitgegebenen Verhaltensregeln¹² können ihn nie auf alle denkbaren Wechselfälle des wirklichen Lebens vorbereiten. Hinzu kommt, dass Intelligente Agenten die ihnen einprogrammierten Regeln nicht – wie Menschen – ohne weiteres durch Analogie oder „Intuition“ auf neue Situationen übertragen können. Angenommen, in dem eingangs geschilderten Fall würde „google-autocomplete“ die Regel vorgegeben bekommen, dass der Name einer bestimmten Person nicht mit den Begriffen „Prostitution“ oder „Escort Service“ ergänzt werden darf; das System würde daraus nicht ohne Weiteres schließen, dass etwa auch die Ausdrücke „Callgirl“ oder „Rotlicht“ verboten sind¹³.

„Wirklich“ intelligente Systeme ordnen freilich das aufgenommene Datenaustausch nicht nur nach Mustern, um auf diese Weise ein vorgegebenes Ziel möglichst effizient zu erreichen¹⁴, sondern sie können darüber hinaus (im weitesten Sinne) *lernen*¹⁵, indem sie ihr Verhalten den gewonnenen Informationen anpassen. Auch das funktioniert auf sehr einfacher Stufe, etwa bei einem *recom-*

12 Russell/Norvig (Anm. 3), S. 57 ff.

13 Auch einfache Softwareagenten können jedoch „Übertragen“ lernen. Dies geschieht beispielsweise dadurch, dass ein Intelligenter Agent mit einem anderen verknüpft wird oder dass er auf der Grundlage eines Zufallsgenerators aus den Reaktionen der Umwelt Schlüsse zieht.

14 Siehe dazu Wooldridge/Jennings, Knowledge Engineering Review 10 (1995), S. 115; John, Haftung für künstliche Intelligenz, 2007, S. 15 ff.

15 Hofstadter, Fluid Concepts and Creative Analogies, 1995, S. 13 ff., 62 ff.

mender system, das dem Käufer im Onlineshop mitteilt, was ihn „auch noch interessieren könnte“ (und dessen Empfehlungen der Intuition eines realen Verkäufers recht nahe kommen). Hinter dem *recommender system*, das dem Nutzer automatisch weitere Angebote unterbreitet, steht ein Intelligenter Agent, der die Eingaben des Nutzers hinsichtlich seiner Präferenzen und Abneigungen kontinuierlich auswertet und mit den Daten anderer Nutzer vergleicht, wobei er ein spezifisches Statistiksystem mit einem Zufallsgenerator kombiniert¹⁶. Zu diesem Zweck speichert der Agent die Informationsspuren aller Nutzer, denn ein *recommender system* benötigt riesige Datenmengen, um Cluster von Kunden zu bilden, die – statistisch gesehen – gleiche Interessen haben¹⁷. Auf einem viel komplexeren Niveau „autonom“ funktioniert ein führerloses Auto: Da es die durch Sensoren aufgenommenen Daten selbständig auf Muster untersucht und unmittelbar darauf reagiert, kommt es nicht nur einem Außenstehenden so vor, als agiere das Auto; auch der Entwickler kann nicht mit Sicherheit vorhersagen, wie das Auto in einer bestimmten Verkehrssituation durch die Verknüpfung der unterschiedlichen Datenverarbeitungsvorgänge reagieren wird, und er kann einzelne „Entscheidungen“ des Autos auch retrospektiv nicht mehr vollständig nachvollziehen. Der Ingenieur, der ein fahrerloses Auto entwickelt, kann also unter Umständen erst nach einer Reihe von Unfällen erkennen, dass sein Auto beispielsweise bei bestimmten Licht- oder Witterungsverhältnissen Vorfahrtszeichen nicht erkennt.

Unvorhersehbar wird das Verhalten des Intelligenten Agenten nicht nur dann, wenn er in eine Lage gerät, für die ihm keine adäquate Reaktion einprogrammiert wurde, sondern auch dann, wenn er sich aufgrund seiner „Erfahrung“ autonom für eine bestimmte Aktion entscheidet¹⁸. Intelligente Agenten versuchen den vom Menschen vorgegebenen Auftrag zu erfüllen, auch wenn sie sich dazu auf neue Situationen einstellen müssen. Darin liegt gerade ihr großer Nutzen. Damit kann aber der Mensch das Verhalten des Agenten in solchen ungeplanten Situationen weder vollständig voraussehen noch kontrollieren; ein Mensch kann beispielsweise nicht vorhersehen, ob und wann ein selbstfahrendes Auto ein nicht eindeutig erkennbares Verkehrsschild aus seiner „Erfahrung“ mit ähnlichen Verkehrsschildern falsch interpretiert. Die fehlende Kontrollierbarkeit ist also

16 Dadurch kommt es bei großen Online-Händlern zu den „Empfehlungen aufgrund Ihrer Bestellung“.

17 Dadurch entsteht ein massives Problem des Datenschutzes. Siehe dazu etwa *Jeckmans/Beye/Erkin/Hartel/Lagendijk/Tang*, *Privacy in Recommender Systems*, 2013, abrufbar unter: http://eprints.eemcs.utwente.nl/22141/01/Privacy_in_Recommender_Systems.pdf (besucht am 18.07.2014).

18 *Russell/Norvig* (Anm. 3) S. 35ff., 57ff.; *Wooldridge/Jennings*, *Knowledge Engineering Review* 10 (1995), S. 118f.

gewissermaßen die Kehrseite des besonderen Werts, den lernfähige Roboter für den Menschen haben.

Der nächste Schritt könnte die Entwicklung „emotionaler“ Interaktionen zwischen Maschine und Mensch sein: Man arbeitet daran, Intelligente Agenten so zu programmieren, dass sie auch die emotionalen Bedürfnisse des Menschen erkennen und empathische Reaktionen simulieren. Auf einfachem Niveau geschieht dies schon heute, etwa wenn ein Onlineshop, der auf ein Angebot keine Buchung registriert, dem Nutzer automatisch mitteilt: „Es tut uns leid, für Sie nicht das Passende gefunden zu haben.“ Mancher Nutzer wird sich über das ihm persönlich entgegengebrachte Mitgefühl der Maschine freuen (und beim nächsten Mal vielleicht eher bereit sein, dort etwas zu bestellen), obwohl es sich in Wahrheit um einen ganz simplen Automatismus handelt. Auch in der Krankenpflege hat man festgestellt, dass Demenzkranke über einen Roboter in Tiergestalt eine „Gefühlsbrücke“ aufbauen können, indem sie der Maschine Empfindungen zuschreiben. In der Informatik spricht man vom *Eliza-Effekt*¹⁹, wenn ein Mensch der Aktion eines Intelligenzen Agenten unbewusst ein menschliches Handlungsschema und eine eigene Motivation unterlegt²⁰. Der Effekt birgt allerdings die Gefahr der Fehlinterpretation der (Re-)Aktionen einer Maschine und damit von „Missverständnissen“ bei der Interaktion zwischen Menschen und Intelligenzen Agenten.

Wie schon der kurze Überblick gezeigt hat, eröffnen lernfähige Intelligente Agenten dem Menschen eine neue Dimension des Fortschritts. Gleichzeitig wirft die rechtliche Beurteilung ihres Verhaltens neue und schwierige Fragen auf. Brisant ist vor allem die Frage der Haftung für die Aktionen Intelligenzen Agenten. Im Zivilrecht hat man sich damit bereits auseinandergesetzt²¹; für das Strafrecht steht die Diskussion jedoch noch am Anfang²². Wird durch einen Intelligenzen Agenten ein strafrechtlich relevanter Erfolg verursacht (z. B. ein Mensch getötet oder verletzt oder eine Sache zerstört), so stellt sich die Frage nach der Verantwortung: Wer hat gehandelt und wer trägt strafrechtliche Schuld? Der Mensch, der die Technologie in Funktion gesetzt hat? Oder der Intelligente Agent selbst?

19 ELIZA ist ein von Joseph Weizenbaum entwickeltes Computerprogramm, bei dem der Computer einen Psychotherapeuten simuliert, der (scheinbar) sprachlich mit einem Menschen interagiert; *Weizenbaum*, *Computer Power and Human Reason*, 1976; *Hildebrandt*, in: *Hildebrandt/Gaakeer* (Hrsg.), *Human Law and Computer Law: Comparative Perspectives*, 2013, S. 27f.

20 Siehe hierzu *Picard*, *Affective Computing*, 1995; *Tao/Tan*, in: *Tao/Tan/Picard* (Hrsg.), *Affective Computing and Intelligent Interaction*, 2005, S. 981.

21 Siehe etwa *Gruber*, in: *Beck* (Hrsg.), *Jenseits von Mensch und Maschine*, 2012, S. 150; *Hanisch*, in: *Hilgendorf* (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, S. 27; *Spindler*, in: *Hilgendorf* (Hrsg.), a. a. O., S. 63.

22 Siehe jedoch den Sammelband von *Beck* (Anm. 21).

Oder haftet vielleicht gar niemand strafrechtlich, weil die Allgemeinheit die Entwicklung und den Nutzen solcher Technologien grundsätzlich wünscht und dann auch die Konsequenzen ihrer (teilweisen) Unberechenbarkeit tragen muss? Mit diesen Fragen beschäftigt sich die folgende Skizze möglicher Begründung und Begrenzung von strafrechtlicher Haftung beim Einsatz von Intelligenten Agenten.

II. Der Intelligente Agent als strafbares Subjekt?

Unser Strafrecht ist für Menschen gemacht. Dem Menschen wird der vom Tatbestand beschriebene Erfolg zugerechnet, wenn er ihn durch sein Verhalten verursacht hat. Und der Mensch trägt strafrechtliche Schuld, wenn er vorsätzlich oder fahrlässig einen Straftatbestand erfüllt und dabei das Unrecht seines Verhaltens erkennen und in zumutbarer Weise vermeiden kann. Allerdings haben wir gesehen, dass auch Intelligente Agenten lernfähig sind und zumindest in übertragenem Sinne „autonom“ agieren können. Heißt das, dass sie auch wie Menschen bestraft werden können?

1. Vorläufer: Tiere und Roboter als Subjekte des Strafrechts

Bevor wir eine Antwort auf diese Frage versuchen, wollen wir den Blick kurz auf Erfahrungen und Vorstellungen über die strafrechtliche Sanktionierung von Nicht-Menschen richten.

In früheren Jahrhunderten hatte der Mensch noch kein Monopol auf Schuld und Strafe. Vergeltung wurde vielmehr auch an „unbeseelten Gegenständen“²³ geübt. Die Strafrechtsgeschichte berichtet über Strafprozesse und Strafen gegenüber Tieren in Europa bis in die frühe Neuzeit²⁴. Einer der jüngeren Tierprozesse ist aus Basel (1474) überliefert: Verhandelt wurde gegen einen Hahn; ihm wurde vorgeworfen, ein Ei gelegt zu haben, aus dem ein Basilisk schlüpfen sollte²⁵. Der Basler Hahn wurde – so wie auch andere verurteilte Tiere – mit dem Tode bestraft. Welche Motivation hinter der Bestrafung von Tieren stand, ist bis heute umstrit-

²³ Bauer, in: *Perels/Wojak* (Hrsg.), *Die Humanität der Rechtsordnung. Ausgewählte Schriften*, 1998, S. 251.

²⁴ Die Tierprozesse enthielten Elemente aus dem Strafprozess gegen Menschen, wie die Untersuchungshaft oder die Bestellung eines Verteidigers für das angeklagte Tier. Siehe *Dinzelbacher*, *Das fremde Mittelalter. Gottesurteil und Tierprozess*, 2006, S. 113.

²⁵ *Brednich*, *Enzyklopädie des Märchens*, Bd. 13, 2010, S. 651.

ten. Historiker verweisen zunächst auf ein ganz profanes Tötungsrecht der Bauern gegenüber schädigendem, streunendem Vieh²⁶. Eine andere Quelle der Tierprozesse dürfte das Streben nach magischem Schutz vor Teufelswerk gewesen sein. Auch der Vergeltungsgedanke hat möglicherweise eine Rolle gespielt: Im Mittelalter nahm man an, dass Gott den Tieren eine den Menschen untergeordnete, „dienende“ Funktion zugewiesen habe²⁷; bei der Verletzung dieser Funktion hatten Tiere dann – ebenso wie Menschen – Strafe verdient, durch die die hierarchischen Verhältnisse in der Welt wiederhergestellt wurden²⁸. Mit der Aufklärung und der Durchsetzung präventiver Straftheorien endete das Strafrecht gegen Tiere²⁹. Tiere werden danach nicht wie Personen, sondern wie Sachen behandelt (siehe § 90a Satz 2 BGB). Allerdings wird im Rahmen der „Animal Rights“-Bewegung heute durchaus überlegt, ob Tiere nicht nur Rechte, sondern auch Pflichten haben sollten³⁰.

Maschinen haben nach modernem Verständnis ebenso wenig wie Tiere den Status von Personen. Sie werden nicht als selbstbestimmte Wesen wahrgenommen, da sie in ihrer Funktionsweise festgelegt sind. *Mark Twain* hat dies schon 1906 so formuliert³¹: „[The machine] does no more and no less than the law of its make permits and compels it to do. There is nothing personal about it; it cannot choose.“ Auch die als „Sozialroboter“ für die Interaktion mit Menschen vorgesehenen Intelligenzen Agenten³² sind im Grunde genommen nicht mehr als eine Verbindung verschiedener technischer Komponenten, wie etwa Gehäuse, Rechner und Sensoren.

Im Genre der Science Fiction finden wir aber immer wieder Vorahnungen, dass eine als Helfer des Menschen erdachte Maschine vielleicht doch einmal mehr werden könnte. Der sich langsam zum Menschen wandelnde Roboter ist ein wiederkehrendes Motiv von Science-Fiction-Romanen. Schon 1950 hat *Isaac Asimov* die Frage nach der Rechtspersönlichkeit von Maschinen plastisch formuliert. In dem Roman mit dem bezeichnenden Titel „I, Robot“ präsentiert *Asimov* Ver-

26 *Brednich* (Anm. 25), S. 651.

27 *Gergen*, NuR 2012, 96, 102.

28 *Fischer*, in: *Rehberg* (Hrsg.), *Die Natur der Gesellschaft*, 2008, S. 5151, 5159.

29 *Dinzelbacher* (Anm. 24), S. 111.

30 Dazu *Ritvo*, *Daedalus* 136 (2007), S. 118; *De Mello*, *Animals and Society*, 2012; *Peters/Stucki/Boscardin*, *The Animal Turn-What is it and Why Now?*, *Verfassungsblog.de* 14. April 2014, abrufbar unter: <http://www.verfassungsblog.de/the-animal-turn-what-is-it-and-why-now> (besucht am 18.07.2014).

31 *Mark Twain*, *What is man?* 1906, abrufbar unter: <http://www.gutenberg.org/ebooks/70> (besucht am 18.07.2014).

32 Siehe dazu *Kobald*, *Ideengeschichtliche Aspekte zur Technikgenese der sozialen Robotik unter weltanschaulicher Perspektive*, 2013, S. 13f.

haltensregeln, die das (hierarchisch geordnete) Verhältnis zwischen Mensch und Roboter bestimmen sollen³³. Auch in zahlreichen Science-Fiction-Filmen spielen Maschinen mit wachsendem Selbstbewusstsein eine wichtige Rolle³⁴. Ein typisches Beispiel ist das Androiden-Kind David in *Steven Spielbergs* Film „AI“, das seiner Menschen-Mutter reuevoll verspricht, sich zu bessern, um dann – verwandelt wie der klassische Pinocchio – nach Hause zurückkehren zu dürfen³⁵.

2. Intelligente Agenten als Personen?

Auch wenn es sich bei Darstellungen von Intelligenzen Agenten, die gelernt haben, Selbstbewusstsein, Gefühle, Empathie und Moral zu entwickeln, und denen deshalb eine Art Willensfreiheit und vielleicht sogar ein Status als Moralischer Agent zugeschrieben wird, nach wie vor nur um fiktive Vorahnungen und nicht um Realität handelt, werfen diese Bücher und Filme wichtige und interessante Fragen auf – nicht zuletzt die Frage, ob und gegebenenfalls unter welchen Voraussetzungen eine Bestrafung von Intelligenzen Agenten in Betracht kommt³⁶. Möglicherweise erscheinen uns Roboter ja bald nicht mehr stärker „ferngesteuert“ als Menschen, die ihre vermeintlich freien Entscheidungen in einem undurchschaubaren Geflecht von Einflüssen aus Genetik, Erziehung und sozialer Umwelt treffen und die wir dennoch als „frei“ verantwortlich definieren.

In diesem Sinne stellt sich zunächst die Frage, wo die Grenze zwischen der Maschine und der menschlichen Person als traditionellem Subjekt des Strafrechts verläuft. Genügt es, dass ein Roboter die vorprogrammierte Fähigkeit besitzt, aus der Umgebung aufgenommene Daten mit bekannten Mustern zu vergleichen und

33 Asimov, I, *Robot*, 1950, S. 27. Danach sollen folgende drei Regeln gelten: „1. A robot may not injure a human being or, through inaction, allow a human being to come to harm; 2. A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law. 3. A robot must protect its own existence, as long as such protection does not conflict with the First or Second Laws.“ (abrufbar unter: http://cdn.preterhuman.net/texts/literature/books_by_author/A/Asimov,%20Isaac%20-%20I,%20Robot.pdf [besucht am 18.07.2014]).

34 Siehe etwa 2001: *A Space Odyssey* (1968); *Star Trek, The Ultimate Computer* (1968); *The Stepford Wives* (1975); *Blade Runner* (1982); *Artificial Intelligence* (2001).

35 David: „I’m sorry I broke myself. I’m so sorry I cut your hair off. I’m sorry I hurt Martin. [...] Mommy if Pinocchio became real and I become a real boy can I come home?“

36 Dazu Beck, in: *Battaglia/Mukerji/Nida-Rümelin/Bisol* (Hrsg.), *Rethinking Responsibility in Science and Technology*, Pisa 2014, 177 ff.; *Koops/Hildebrandt/Jaquet-Chiffelle*, *Minnesota Journal of Law, Science & Technology* 11 (2010), S. 522ff.; *Hildebrandt*, in: *Hildebrandt/Rouvroy* (Hrsg.) *Law, Human Agency and Autonomic Computing*, 2011, S. 141ff.

auf dieser Grundlage zu handeln, damit wir ihn als eine für ihr Verhalten verantwortliche „Person“ ansehen können?

Nach einer von *John Locke*³⁷ und *Immanuel Kant* geprägten Linie der Rechtsphilosophie gilt als Person, wer ein gedankliches Verhältnis zu sich selbst entwickeln kann. Die Fähigkeit zur Selbstreflexion ermöglicht der Person die Selbstbestimmung; oder, wie *Kant* es ausdrückt: Da der Mensch sich seiner selbst bewusst sei, betrachte er auch sich selbst „nur als bestimmbar durch Gesetze, die er sich durch Vernunft selbst gibt, und in diesem seinem Dasein ist ihm nichts vorhergehend vor seiner Willensbestimmung, sondern jede Handlung (...) ist im Bewußtsein seiner intelligibelen Existenz nichts als Folge, niemals aber als Bestimmungsgrund seiner Kausalität, als Noumens, anzusehen.“³⁸ Aus diesem Grund, fährt *Kant* fort, könne das vernünftige Wesen von jeder gesetzwidrigen Handlung, die es verübt, mit Recht sagen, dass es sie auch hätte unterlassen können³⁹. Wenn die Person aber im Selbstbewusstsein ihrer existentiellen Freiheit nicht umhin kann, sich als verantwortlich für ihre Taten zu verstehen⁴⁰, so sind auch die Anderen berechtigt, der Person diese Verantwortung zuzuschreiben. Eine andere, auf *Hegel* zurückreichende Denktradition trennt den Begriff der Person streng vom Begriff des Menschen und versteht erstere als die Bezeichnung des menschlichen Wesens als möglichen Träger von Rechten und Pflichten in der Gesellschaft⁴¹.

Nun sind diese in der idealistischen Philosophie gründenden Erwägungen gewiss nicht auf die modernen Roboter bezogen. Man kann aber sagen, dass es dem Intelligenzen Agenten an wesentlichen Elementen des Person-Seins fehlt: Mag er auch lernen und – von anderen nicht vorhersehbare – Entscheidungen

37 *John Locke*, *An Essay Concerning Human Understanding*, 1690, S. 331 (§ XVII Nr. 26), nimmt als Voraussetzung des Person-Seins an, dass jemand als „intelligent Agent“ (!) „capable of a Law, and Happiness and Misery“ ist. Das Bewusstsein der eigenen Vergangenheit ist der Schlüssel zur Person: „This personality extends itself beyond present existence to what is past, only by consciousness, – whereby it becomes concerned and accountable; owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present.“ Siehe hierzu auch *Seelmann*, *Rechtsphilosophie*, 5. Aufl. 2010, S. 132ff.; ausführlich *ders.*, in: *Heer* u. a. (Hrsg.), *Toujours agité, jamais abattu*, Festschrift für Wiprächtiger, 2011, S. 575.

38 *Kant*, *Kritik der praktischen Vernunft*, in: *Weischedel* (Hrsg.), *Immanuel Kant, Werke in zehn Bänden*, 1975, Bd. 6, S. 223.

39 *Kant* (Anm. 38), S. 223.

40 *Kant* (Anm. 38), S. 223f., verweist hier auf die Stimme des Gewissens, deren Selbsttadel und Verweis der Mensch – er „mag künsteln soviel als er will“ – nicht zum Verstummen bringen könne.

41 Siehe hierzu *Jakobs*, *Norm, Person, Gesellschaft*, 3. Aufl. 2008, S. 37ff.; *Pawlik*, *Das Unrecht des Bürgers*, 2012, S. 141 ff.

treffen können, so kann er sich doch nicht seiner eigenen Freiheit bewusst werden; und noch weniger kann er sich selbst als Träger von Rechten und Pflichten in der Gesellschaft verstehen⁴². Intelligente Agenten können zwar bestimmte Aufgaben „automatisch“ erledigen, aber sie folgen – selbst wenn sie lernfähig sind – letztlich den durch die Programmierung vorgegebenen Optionen und geben sich nicht selbst über ihr Verhalten Rechenschaft ab⁴³. Daher nehmen wir Intelligente Agenten als „unfrei“ wahr und machen sie nicht personal verantwortlich für Schäden, die sie anrichten.

Diese Feststellungen beziehen sich – darüber muss man sich im Klaren sein – auf den gegenwärtigen Stand der Entwicklung künstlicher Intelligenz einerseits und auf das traditionelle Verständnis des Personseins im Sinne der deutschen Philosophie andererseits, wobei letztere ja für das Strafrecht immer noch die maßgeblichen grundlegenden Parameter vorgibt. Falls man Intelligente Agenten auch im Hinblick auf Selbst- und Rechtsbewusstsein, ethische wie emotionale Ansprechbarkeit und andere „typisch menschliche“ Dimensionen fortentwickeln könnte, wäre natürlich auch die Frage neu zu betrachten, ob sie nicht doch als (dann womöglich schuldfähige) Rechtspersonen angesehen werden könnten⁴⁴.

3. Strafbarkeit Intelligenter Agenten?

Das Problem, dass Intelligente Agenten *rebus sic stantibus* für ihr Verhalten nicht in rechtlicher oder moralischer Hinsicht verantwortlich gemacht werden können, spiegelt sich auch in den Kategorien der Strafrechtsdogmatik wider. Ebenso wie bei der in manchen Punkten ähnlich gelagerten Fragestellung einer Strafbarkeit juristischer Personen⁴⁵ lassen sich die Probleme zunächst bei den Voraussetzungen der „Handlung“ und der „Schuld“ verorten.

⁴² Grundsätzliche Überlegungen zur Rechtspersönlichkeit im Bereich künstlicher Intelligenz etwa bei *Solum*, North Carolina Law Review 70 (1992), S. 1231; *Matthias*, Automaten als Träger von Rechten, 2008.

⁴³ Siehe auch *Koops/Hildebrandt/Jaquet-Chiffelle* (Anm. 36), S. 522ff.

⁴⁴ Siehe zur Statusfrage *Beck*, in: *Hilgendorf/Günther* (Hrsg.), Robotik und Gesetzgebung, 2013, S. 239ff.; *Hildebrandt*, in: *Hildebrandt/Gaakeer* (Anm. 19), S. 27f.

⁴⁵ Siehe hierzu grundlegend *Böse*, ZStW 126 (2014), S. 126 ff.; *Frisch*, Festschrift für Wolter, 2013, S. 349; *Heine*, Die strafrechtliche Verantwortlichkeit von Unternehmen, 1995; *Hettinger* (Hrsg.), Reform des Sanktionenrechts, Bd. 3: Verbandsstrafe, 2003; *Jakobs*, Festschrift für Lüderssen, 2002, S. 559; *Mitsch*, NZWiSt 2014, 1; *Schünemann*, Festschrift für Tiedemann, 2008, S. 429; *ders.*, in: Leipziger Kommentar, 12. Aufl. 2007, vor § 25 Rdn. 17 ff.; *Volk*, JZ 1993, 429. Aus neuester Zeit siehe den Lösungsversuch im Gesetzentwurf des Landes Nordrhein-Westfalen für ein Verbandsstrafgesetzbuch; abzurufen unter: <http://www.justiz.nrw.de/JM/justizpolitik/jumiko/beschlues>

a) Handlungsfähigkeit Intelligenter Agenten

Eine Internet-Suchmaschine stellt rufschädigende Wortkombinationen zusammen, eine Drohne verletzt harmlose Personen, ein fahrerloses Auto überfährt einen Fußgänger: Die so beschriebenen Vorgänge scheinen auf Handlungen des Intelligenten Agenten selbst zu verweisen, sofern dieser unabhängig von aktueller menschlicher Steuerung auf der Grundlage seines – im oben erläuterten Sinne lernfähigen⁴⁶ – Algorithmenprogramms agiert. Aber sind dies wirklich „Handlungen“ des Intelligenten Agenten im strafrechtlichen Sinne?

Mit der Frage, wann eine rechtlich relevante Handlung vorliegt, hat sich die Strafrechtswissenschaft vor allem um die Mitte des letzten Jahrhunderts beschäftigt. Für die überkommene „kausalistische“ Theorie genügt jede gewillkürte Körperbewegung als strafrechtlich relevante Handlung; ihre modernere „soziale“ Variante verlangt darüber hinaus einen irgendwie gearteten Sozialbezug der Körperbewegung⁴⁷. Wendet man diese Definitionen auf Intelligente Agenten an, so sieht man jedenfalls bei jenen Intelligenten Agenten, die eine reine Software-Umgebung (wie das Internet) verlassen haben und uns als mechanisch operierende Roboter gegenüber treten⁴⁸, ohne Weiteres, dass sich deren „Körper“ bewegen; man stößt dann allerdings auf die Frage, ob Intelligente Agenten „willkürlich“ handeln können. Nach dem kausalistischen Handlungsverständnis werden jedoch keine hohen Ansprüche an die Willenssteuerung des Verhaltens gestellt; verneint wird eine willkürliche Handlung nur dann, wenn der Körper des Menschen bloß „automatisch“ reagiert (wie bei bestimmten nicht zu beherrschenden Reflexen) oder von einem anderen Menschen als physisches Objekt benutzt wird (A wirft B auf C)⁴⁹. Letzteres ist bei einfachen menschengesteuerten Maschinen wie Staubsaugern oder herkömmlichen Autos der Fall, nicht aber bei Intelligenten Agenten, die sich durch ihre Lernfähigkeit der jederzeitigen Kontrolle seitens eines Menschen entziehen; sie können im Sinne der kausalen Handlungstheorie „willkürliche“ Entscheidungen treffen.

se/2013/herbstkonferenz13/zw3/TOP_II_5_Gesetzesentwurf.pdf (besucht am 18.07.2014); dazu *Hoven*, ZIS 2014, 19; *Löffelmann*, JR 2014, 185; *Zieschang*, GA 2014, 91.

46 Siehe oben II. 1.

47 Zu den Handlungstheorien siehe die Überblicke bei *Jescheck/Weigend*, Lehrbuch des Strafrechts, Allgemeiner Teil, 5. Aufl. 1996, S. 217 ff.; *Puppe*, in: *Nomos Kommentar StGB*, 4. Aufl. 2013, vor § 13 Rdn. 31 ff.; *Roxin*, Strafrecht Allgemeiner Teil, Band I, 4. Aufl. 2006, S. 236 ff.

48 Problematisch ist das Vorliegen einer „Handlung“ im herkömmlichen Sinne allerdings bei rein „intellektuell“ operierenden Agenten, wie etwa Suchmaschinen im Internet. Hier müsste man den Handlungsbegriff schon stark „entkörperlichen“, um ihr Verhalten noch als „Handlung“ bezeichnen zu können.

49 Siehe *Baumann/Weber/Mitsch*, Strafrecht Allgemeiner Teil, 11. Aufl. 2003, S. 207–211.

Im Unterschied zu der zuvor herrschenden kausalen Handlungslehre machten die Vertreter eines finalen Handlungsbegriffs die Zweckgerichtetheit zum zentralen Kriterium der menschlichen Handlung⁵⁰. Nach ihrer Auffassung unterscheidet sich die menschliche Handlung von sonstigen Geschehnissen, die für einen Schadenseintritt kausal werden können, dadurch, dass sie vom Willen des Menschen determiniert ist. Nur wenn und weil die menschliche Handlung Ausdruck eines zielgerichteten Willens des Handelnden ist, kann sie strafrechtlich relevant sein⁵¹. Wie sieht es danach mit der Handlungsfähigkeit Intelligenter Agenten aus? Sie sind aufgrund ihrer Lernfähigkeit in der Lage, selbständig die optimalen Zwischenschritte zu bestimmen, die unter den jeweiligen Umständen zum Erfolg ihrer Mission führen. So kann etwa – auf ganz einfacher Stufe – eine Internetsuchmaschine nach einer bestimmten Anzahl von Anfragen eine eigenständige Verknüpfung der verwendeten Suchbegriffe herstellen oder – auf sehr viel komplexerem Niveau – eine bewaffnete Drohne feindliche Objekte identifizieren und angreifen. Es ist allerdings eine Frage der (externen) Beurteilung, ob man annimmt, dass der Intelligente Agent seine Entscheidung aufgrund von „choices“ trifft, die ihm seine Programmierung bereits vorgibt, oder ob er „judgments“ – also „eigene“ Wertungen – vornimmt. Selbst wenn man letzteres bejaht, bedeutet dies jedoch noch nicht, dass sich ein Intelligenter Agent selbst „willentlich“ Ziele setzen und seine Handlungen nach diesen Zielsetzungen bestimmen kann. Dies würde ja mindestens voraussetzen, dass der Intelligente Agent selbst „erkennt“, was er tut, und dass er darüber hinaus weiß, dass sein Verhalten „sozialerheblich“ ist, also potentiell auf das Leben anderer Personen Einfluss nehmen kann. Ob Intelligente Agenten eine solche Art von Selbstbewusstsein haben oder erlangen können, ist Gegenstand der Diskussion um sog. Künstliche Intelligenz; bisher gibt es dazu aber noch keine positiven Antworten⁵². Wir nehmen heute (noch) nicht an, dass Intelligente Agenten wissen, was sie tun, und die Relevanz der eigenen Aktionen – über die Feststellung hinaus, ob sie sie dem programmierten Ziel näherbringen – wahrnehmen.

Letztlich dürfte die Frage nach der (strafrechtlichen) Handlungsfähigkeit Intelligenter Agenten eine Frage der Definition sein: Bei „kausalistischer“, bloß äußerlicher Betrachtung, die jede „willkürliche Körperbewegung“ als Handlung definiert, sind sie durchaus als Handelnde anzusehen. Je stärker man den Begriff der Handlung substantiell auflädt, je mehr an selbstbewusster Zielbestimmung

50 Welzel, Das deutsche Strafrecht, 11. Aufl. 1969, S. 33ff.; ders., JuS 1966, 423; aus heutiger Sicht Stratenwerth/Kuhlen, Strafrecht Allgemeiner Teil, 6. Aufl. 2011, S. 57f.

51 Stratenwerth, Schweizerisches Strafrecht. Allgemeiner Teil, Band I, 4. Aufl. 2011, S. 128f.

52 Siehe dazu Hofstadter (Anm. 15), S. 179ff., 195ff.

man in ihn hineinliest, desto weniger können Intelligente Agenten den Voraussetzungen der Handlungsfähigkeit genügen.

b) Schuldfähigkeit Intelligenter Agenten

Wenn man annimmt, dass Intelligente Agenten im strafrechtlichen Sinne Handlungen vornehmen, steht man bei der Frage nach ihrer möglichen strafrechtlichen Verantwortlichkeit alsbald vor der nächsten Hürde: Können sich Intelligente Agenten schuldig machen? Kann es einer Suchmaschine vorgeworfen werden, dass sie bei der Eingabe eines bestimmten Namens eine rufschädigende Begriffskette anzeigt, und kann man es einem intelligenten Skalpell vorwerfen, dass es gesundes Gewebe herausschneidet statt sich auf krankes zu beschränken? Für das Zivilrecht wird vorgeschlagen, Intelligente Agenten als Haftungssubjekte eigener Art zu verstehen, die – über einen zuvor eingerichteten Haftungsfonds – für die von ihnen verursachten Schäden in Anspruch genommen werden können⁵³. Auch für das Strafrecht wird erwogen, Intelligenzen mit Blick auf ihre Lern- und Erinnerungsfähigkeit Schuldfähigkeit zuzuschreiben⁵⁴. Nach dem verfassungsrechtlich verankerten Schuldgrundsatz⁵⁵ setzt Strafbarkeit jedoch eine persönliche Verantwortung des Täters für seine Tat in dem Sinne voraus, dass er sie bei genügender und ihm zumutbarer Anstrengung hätte vermeiden können; nur dann darf ihm sein unrechtes Verhalten vorgeworfen werden⁵⁶. Der BGH hat dies im Jahre 1952 in klassischer Formulierung so ausgedrückt: „Der innere Grund des Schuldvorwurfes liegt darin, daß der Mensch auf freie, verantwortliche, sittliche Selbstbestimmung angelegt und deshalb befähigt ist, sich für das Recht und gegen das Unrecht zu entscheiden, sein Verhalten nach den Normen des rechtlichen Sollens einzurichten und das rechtlich Verbotene zu vermeiden, sobald er die sittliche Reife erlangt hat (...)“⁵⁷.

53 Gruber (Anm. 22), S. 150 ff.

54 Zu Voraussetzungen der Entstehung von Rechtspersönlichkeit vgl. etwa Beck/Zabel, in: Friedrich/Zichy (Hrsg.), *Persönlichkeit – Neurowissenschaftliche und neurophilosophische Fragestellungen*, 2014, S. 49; Hildebrandt, in: Duff/Green, *Philosophical Foundations of Criminal Law*, Oxford 2014, S. 510.

55 Siehe BVerfGE 45, 187, 227; 95, 96, 140; besonders markant BVerfGE 123, 267, 413, wo das Schuldprinzip wegen der Verankerung in der Menschenwürde als unverfügbar i.S.v. Art. 79 Abs. 3 GG bezeichnet wird.

56 Allgemein zum Schuldgrundsatz Frister, *Strafrecht Allgemeiner Teil*, 6. Aufl. 2013, S. 30 ff.; Kühl, *Strafrecht Allgemeiner Teil*, 7. Aufl. 2012, S. 364 ff.; Roxin, *Allg. Teil I* (Anm. 47), S. 851 ff.; Stratenwerth/Kuhlen, *Allg. Teil* (Anm. 50), S. 159 ff.

57 BGHSt. 2, 194, 200.

Diese Definition scheint die Annahme strafrechtlicher Schuld einer noch so intelligenten Maschine kategorisch auszuschließen. Denn Intelligente Agenten werden jedenfalls derzeit nicht als auf sittliche Selbstbestimmung angelegte Personen wahrgenommen, sondern als computergestützte Technologien, die die ihnen gestellten Aufgaben gemäß ihrer Programmierung ausführen. Man kann ihnen daher keine Schuld in dem vom BGH proklamierten Sinne zuschreiben. Nun ist jedoch das 1952 noch (oder wieder) relativ unumstrittene Dogma der Willensfreiheit des Menschen – als wichtige Grundlage des traditionellen Schuldbegriffs – in jüngerer Zeit im Lichte bestimmter Erkenntnisse der Hirnforschung⁵⁸ vehement in Frage gestellt worden. Was die aktuelle Kontroverse von der idealistischen Vorstellung eines willensfreien Menschen übrig lässt, ist noch nicht abschließend geklärt⁵⁹. Dies ist allerdings – entgegen dem ersten Anschein – für die Diskussion um die mögliche strafrechtliche Schuld Intelligenter Agenten auch nicht von maßgeblicher Bedeutung. Denn schon vor der Diskussion um die Ergebnisse der jüngeren neurobiologischen Forschung hatte sich für das Strafrecht weitgehend die Erkenntnis durchgesetzt, dass es bei der „Feststellung“ strafrechtlicher Schuld nicht um die Ermittlung neuronaler Verknüpfungen und Abläufe, sondern um Erwartungen und Zuschreibungen in den aktuellen sozialen Verhältnissen geht⁶⁰.

Aus einem so verstandenen, von neuronalen Abläufen losgelösten Schuldgrundsatz ergeben sich durchaus denkbare Ansatzpunkte für die Annahme einer „Schuld“ intelligenter Agenten⁶¹. Aus einer funktionalen Perspektive wird Schuld demjenigen zugeschrieben, der als verantwortlicher Urheber des in der Straftat

58 Siehe etwa Roth/*Pauen*, Freiheit, Schuld und Verantwortung, 2008; *Singer*, in: *Geyer* (Hrsg.), Hirnforschung und Willensfreiheit, 2004, S. 30 ff.

59 Aus der umfangreichen Literatur zu dem Thema siehe nur *Duttge*, in: *ders.* (Hrsg.), Das Ich und sein Gehirn, 2009, S. 13; *Hillenkamp*, JZ 2005, 313; *Jäger*, GA 2013, 3; *G. Merkel*, Festschrift für Herzberg, 2008, S. 3; *R. Merkel*, Willensfreiheit und rechtliche Schuld, 2008; *Schiemann*, ZJS 2012, 774; *Streng*, Festschrift für Jakobs, 2007, S. 675; *Weißer*, GA 2013, 26.

60 *Frister*, Allg. Teil (Anm. 56), S. 32f.; *Roxin* (Anm. 47), S. 869 ff.; aus anderer Perspektive auch *Jakobs*, Strafrecht Allgemeiner Teil, 2. Aufl. 1991, S. 484f. Siehe auch bereits *Kant*, Kritik der reinen Vernunft, Werke in sechs Bänden, Bd. 2 (1956), S. 501: „Die eigentliche Moralität der Handlungen (Verdienst und Schuld) bleibt uns daher, selbst die unseres eigenen Verhaltens, gänzlich verborgen. Unsere Zurechnungen können nur auf den empirischen Charakter bezogen werden. Wie viel aber davon reine Wirkung der Freiheit, wie viel der bloßen Natur und dem unverschuldeten Fehler des Temperaments, oder dessen glücklicher Beschaffenheit (*merito fortunae*) zuzuschreiben sei, kann niemand ergründen, und daher auch nicht nach völliger Gerechtigkeit richten.“

61 Siehe hierzu auch *Gruher* (Anm. 44), S. 157; *ders.* (Anm. 21), S. 150; *Teubner*, Zeitschrift für Rechtssoziologie 27 (2006), S. 15 ff.

ausgetragenen sozialen Konflikts identifiziert wird⁶². Auch die Idee des „freien Willens“ wird damit fungibel: Die „Freiheit zu selbstbestimmtem Handeln“ ist kein Naturphänomen, sondern Produkt einer Zuschreibung, die bestimmten gesellschaftlichen Zwecken dient, nämlich der Organisation der Gesellschaft unter der Prämisse wechselseitig anzuerkennender Freiheit⁶³. Allerdings können Schuld und die darauf basierte rechtliche Verantwortlichkeit unter den Bedingungen der gegenwärtigen Gesellschaft nicht nach Belieben zugeschrieben werden; andernfalls würde man die soziale Funktionalität des Strafrechts aufs Spiel setzen. Bei den Mindestvoraussetzungen dafür, dass wir jemandem Verantwortlichkeit für sein Verhalten als Person zuschreiben, gelangen wir wieder zu der oben schon erörterten Bedingung der Möglichkeit zur Selbstreflexion⁶⁴. Wer nicht in der Lage ist, sich selbst und seine früheren Entscheidungen nach einem ethischen Referenzsystem zu bewerten, in der Alltagssprache: „ein gutes oder schlechtes Gewissen“ zu haben⁶⁵, der kommt auch als Rechtsperson, als Objekt einer Zuschreibung von Verantwortung und Schuld nicht in Betracht; denn mit ihm ist kein ethischer Dialog möglich, er kann auf einen ethischen Vorwurf keine „Antwort“ geben. Ein solcher Agent kann – wie ein wildes Tier oder ein unter schweren Wahnvorstellungen leidender Mensch – nur durch Maßnahmen der Gefahrenabwehr kontrolliert, aber nicht strafrechtlich sanktioniert werden.

Allerdings „entscheiden“ Intelligente Agenten bereits heute auf einer Grundlage, die über das Bewerten der eigenen früheren Handlungsweisen in einem einfachen Raster von Erfolg oder Misserfolg hinausgeht. Je schwieriger die dem Intelligenten Agenten übertragenen Aufgaben sind, umso komplexer müssen ihre Lernstrukturen werden; dies gilt etwa für ein selbstfahrendes Auto, das für das Erreichen seines Ziels vor der Entscheidung steht, eine Bundesstraße zu benutzen, auf der sich der Verkehr staut, oder eine für den Durchgangsverkehr gesperrte Spielstraße zu durchfahren. Es ist absehbar, dass Intelligente Agenten künftig in vermehrtem Umfang auch ethische Aspekte für ihr Entscheidungsverhalten werden berücksichtigen müssen. Das gilt z.B. bei selbstfahrenden Autos für notwendige Ausweichmanöver, die zur Verletzung von Rechtsgütern Dritter führen können⁶⁶; ethische Dilemmata können aber auch beim Einsatz ziviler Droh-

62 *Jakobs* (Anm. 60), S. 471ff.

63 *Frister* (Anm. 56), S. 32f.

64 Siehe auch *Beck/Zabel* (Anm. 54), S. 62ff.

65 Siehe *Seelmann* in: *Heer* u. a. (Anm. 37), S. 575.

66 Vor diesem Hintergrund erschließt sich die lebhafteste Diskussion im Internet darüber, wie ein Intelligenter Agent das „Trolley Problem“ oder das „Fat Man Problem“ (in der deutschen Debatte als „Weichensteller“-Fälle bekannt) lösen könnte; siehe: <http://marginalrevolution.com/marginalrevolution/2012/06/the-google-trolley-problem.html> (besucht am 18.07.2014).

nen für den Versand online bestellter Waren vorkommen⁶⁷. Wenn solche (schon konkreten) Pläne umgesetzt werden, müssten Liefer-Drohnen bei der Erfüllung ihrer Mission regelmäßig zwischen mehreren Handlungsmöglichkeiten wählen und sich dabei an einem (einprogrammierten) Bonus-Malus-Schema orientieren, das auch die Einhaltung normativer Regeln berücksichtigt. Wie sollte sich eine Drohne verhalten, die eine bestimmte Zielvorgabe (z.B. Auslieferung der Ware beim Empfänger vor 17 Uhr) und gleichzeitig den Auftrag erhält, ihre Mission in sozialverträglicher Weise auszuführen, etwa Lärmschutzauflagen zu erfüllen und die Privatsphäre der Menschen auf ihrem Weg zu respektieren? Zu diesem Zweck müssten ihr detaillierte Verhaltensregeln mitgegeben werden, z.B. nicht in niedriger Höhe Krankenhäuser oder Penthouse-Wohnungen zu überfliegen. Für die Einhaltung dieser Regeln erhielte die Drohne Bonus-Punkte, während deren Verletzung mit Malus-Punkten belegt wird, quantifiziert je nach der Schwere des Regelverstößes⁶⁸. Aus der Bonus-Malus-Bilanz, die nach der Erfüllung der Mission gezogen wird, kann die Drohne „lernen“, ihre Ziele künftig mit einer möglichst günstigen Punktbilanz zu erreichen⁶⁹. Bei einem solchen System stellt sich die Frage, ob Regelverletzungen nicht nur zu einer Anpassung des zukünftigen Verhaltens der Intelligenten Agenten führen sollen, sondern ob man der Drohne nicht auch Schuld zuschreiben könnte, wenn sie sich normativ „falsch“ verhält, also z.B. eine mit Malus-Punkten belegte Route über mehrere Krankenhäuser wählt.

Denkbar wäre eine solche Schuld-Zuschreibung gewiss – aber hätte sie auch Sinn für die Bewältigung des sozialen Interessenkonflikts zwischen der möglichst raschen Warenauslieferung und dem Ruhebedürfnis der Patienten? Man kann diese Frage wohl nur dann bejahen, wenn man voraussetzt, dass ein komplexes Bonus-Malus-System für einen Intelligenten Agenten funktional der ethischen Selbstbewertung und Selbststeuerung eines Menschen entspricht. Nur unter dieser Voraussetzung kann es sinnvoll sein, Intelligenten Agenten einen strafrechtlichen Schuldvorwurf zu machen. Wenn man dies möchte, müsste man die Schuldvoraussetzung der Fähigkeit zur Unrechtseinsicht den Bedingungen eines Intelligenten Agenten anpassen: Es müsste dann genügen, dass der (menschliche

⁶⁷ Siehe dazu: <http://www.zeit.de/wirtschaft/2013-12/amazon-will-ware-per-drohne-liefern> (besucht am 18.07.2014).

⁶⁸ *Pfützenreuter*, Intelligentes Missionsmanagement für autonome mobile Systeme, 2005.

⁶⁹ So erhielte eine Drohne, die sich dafür entscheidet, zwecks möglichst rascher Auslieferung der Ware zwei Krankenhäuser zu überfliegen, die Rückmeldung, dass sie für den Regelverstoß eine hohe Zahl von Malus-Punkten erhalten hat; das soll sie dazu veranlassen, beim nächsten Versuch eine andere Route zu wählen, die vielleicht etwas länger ist, aber die Lärmbelastung der Kranken vermeidet.

oder Intelligente) Agent die Unerwünschtheit eines Verhaltens erkennen und dieses Verhalten deshalb „intern“ negativ bewerten kann⁷⁰. Mit solchen Überlegungen beschäftigt sich der seit Beginn dieses Millennium wachsende Bereich der *Machine Ethics*. Hier geht es vor allem darum, ob und gegebenenfalls wie man Robotern ein ethisch begründetes Regelsystem implementieren kann⁷¹. Bisher stecken diese Projekte aber noch in den Kinderschuhen, da die Formalisierung ethischer Entscheidungen generell große Schwierigkeiten bereitet, so dass die Programmierung solcher Entscheidungen für Roboter noch in weiter Ferne liegen dürfte⁷².

c) Bestrafung Intelligenter Agenten

Wenn man gleichwohl davon ausgeht, dass einem Intelligenzen Agenten – zumindest in einer absehbaren Zukunft – strafrechtliche Verantwortlichkeit zugeschrieben werden kann, stellt sich vor der Einführung eines Roboter-StGB allerdings noch die weitere Frage, ob Intelligente Agenten bestraft werden können. Unser Strafrecht ist – jedenfalls bis jetzt – auf Menschen bezogen und schließt deshalb die Vorstellung aus, dass andere als menschliche Wesen für ihr Fehlverhalten bestraft werden können oder sollen⁷³. Dieser Grundsatz ist in verschiedenen Rechtsordnungen insoweit durchbrochen, als dort Kriminalstrafen gegen juristische Personen verhängt werden können⁷⁴. Auch diese (zumeist finanziellen) Sanktionen sollen aber letztlich nicht das Abstraktum „juristische Person“ treffen, sondern diejenigen natürlichen Personen, die wirtschaftlich – etwa

⁷⁰ Zur parallelen Problematik bei einer Strafbarkeit juristischer Personen *Frisch* (Anm. 45), S. 349, 353.

⁷¹ *Wallach/Allen*, *Moral Machines: Teaching Robots Right from Wrong*, Oxford 2008; *Anderson/Anderson* (Hrsg.), *Machine Ethics*, Cambridge 2011; *Storrs Hall*, *Beyond AI: Creating the Conscience of the Machine*, 2007; *Georges*, *Digital Soul: Intelligent Machines and Human Values*, Cambridge, MA 2003.

⁷² Siehe dazu etwa *Muehlhauser/Helm*, *Intelligence Explosion and Machine Ethics*, 2012, <http://intelligence.org/files/IE-ME.pdf> (besucht am 18.07.2014). Zur Auseinandersetzung um die mögliche technische Umsetzung siehe *Bostrom/Yudkowsky*, *Cambridge Handbook of Artificial Intelligence*, Cambridge 2011; *Santos-Lang*, in: *van Rysewyk/Pontier* (Hrsg.), *Machine Medical Ethics*, New York 2014, S. 74.

⁷³ Siehe dazu *Matthias* (Anm. 42); *Wallach/Allen* (Anm. 71).

⁷⁴ Siehe dazu die Beiträge von *Beale* (USA), *Hilf* (Österreich) und *Tricot* (Frankreich) in *ZStW* 126 (2014), S. 27 ff., 73 ff., 55 ff.; zum Schweizer Recht siehe Art. 102 schw. StGB und dazu *Stratenwerth* (Anm. 51), S. 447 ff.

als Anteilseigner oder Gesellschafter – hinter der juristischen Person stehen⁷⁵. Die Möglichkeit einer „Bestrafung“ Intelligenter Agenten, die dem Sinn und den Zwecken der Strafe gegenüber Menschen entspräche, kann man sich dagegen derzeit kaum vorstellen. Da ein Intelligenter Agent kein eigenes Vermögen hat (jedenfalls kann er nicht *begreifen*, dass er ein Vermögen hat, selbst wenn es ihm zivilrechtlich irgendwie zugeordnet werden sollte)⁷⁶, müsste eine ihm etwa auferlegte Geldstrafe von seinem Betreiber oder von einem Haftungsfonds beglichen werden, so dass die Sanktion nicht den Intelligenzen Agenten träge; und eine „persönliche“ Bestrafung etwa durch physische Zerstörung oder durch vollständige Neuprogrammierung des Intelligenzen Agenten mutet zwar beim Blick von außen wie eine Form der Todesstrafe an, trifft aber den Agenten doch nicht in gleicher Weise wie einen Menschen – jedenfalls solange Intelligente Agenten nicht mit einem menschlichen (Über-)Lebenswillen begabt sind. Aus diesem Grund scheitert eine kriminalrechtliche Bestrafung von Intelligenzen Agenten beim gegenwärtigen Stand der Entwicklung daran, dass sie eine Strafe nicht als solche empfinden können. Der Mensch kann die Zerstörung eines Roboters als tadelnde Sanktion für dessen Fehlverhalten verstehen – dem Roboter selbst erschließt sich dieser Sinn der Maßnahme nicht. Dennoch kann die Annahme einer (wenngleich „imperfekten“) strafrechtlichen Verantwortlichkeit eines Intelligenzen Agenten von rechtlicher Bedeutung sein – nämlich insofern, als sie einen menschlichen Agenten im Hintergrund möglicherweise von dessen Verantwortlichkeit zu entlasten vermag⁷⁷.

4. Zwischenergebnis

Solange Intelligente Agenten nur im Rahmen ihrer Programmierung „entscheiden“, fehlt es ihnen – je nach dem zugrunde gelegten Begriff der Handlung – möglicherweise schon an der Fähigkeit zum strafrechtlich relevanten „Handeln“. Wo diese angenommen wird, stößt man auf die Grundsatzfrage, ob man Intelli-

⁷⁵ Unter welchen Voraussetzungen eine solche Haftung der wirtschaftlich Betroffenen für das Fehlverhalten anderer Personen zulässig, angemessen oder gar geboten ist, ist eine schwierige, im Kern ungeklärte Frage, die man gerne durch Analogien zum Zivilrecht zu überspielen versucht. Siehe hierzu *Forster*, Die strafrechtliche Verantwortlichkeit des Unternehmens nach Art. 102 StGB, 2006, S. 38f.; *Kremnitzer/Ghanayim*, ZStW 113 (2001), S. 538, 542ff.; *Seelmann*, in: *Ackermann/Donatsch/Rehberg* (Hrsg.), Wirtschaft und Strafrecht, 2001, S. 182ff.; *Triig*, StraFo 2011, 471, 473; *Weigend*, Journal of International Criminal Justice 6 (2008), S. 927, 931f.

⁷⁶ Siehe hierzu auch *Koops/Hildebrandt/Jaquet-Chifelle* (Anm. 36), S. 527ff.

⁷⁷ Dazu noch unten III.

genten Agenten Schuld zuschreiben darf. Dies ist nicht nur fraglich, wenn sie einen berechneten oder mindestens berechenbaren Entscheidungsablauf vollziehen, sondern auch dann, wenn sie aus ihrer Umgebung aufgenommene Daten selbständig – unter Rückgriff auf ihre Erfahrung – interpretieren und ihren eigenen Weg bestimmen. Denn schuldfähiges Subjekt kann nach unserem heutigen Verständnis nur jemand sein, der ein Gewissen hat, also zu ethisch bezogener Selbstreflexion imstande ist. Das ist bei Intelligenten Agenten nach dem derzeitigen Stand der Entwicklung nicht ersichtlich; eine mindestens analoge Fähigkeit zur „moralischen“ Bewertung des eigenen Verhaltens ist für zukünftige Intelligente Agenten aber durchaus vorstellbar⁷⁸. Der letzte Schritt läge dann in der Herstellung einer Straffähigkeit Intelligenter Agenten, also in deren Konstruktion derart, dass sie mit bestimmten äußeren Veränderungen die moralische Missbilligung ihres Verhaltens assoziieren können. Dahin dürfte es aber noch ein langer Weg sein.

III. Der Mensch hinter dem Intelligenten Agenten

Der Gedanke, dass Intelligente Agenten massive Schäden anrichten können, ohne dass dem eine adäquate strafrechtliche Sanktionsmöglichkeit gegenüber steht, hinterlässt Unbehagen; und dieses Unbehagen nimmt zu, je mehr Intelligente Agenten als scheinbar eigenständige Akteure in unserem Leben in Erscheinung treten. Solange wir Intelligenten Agenten keine eigene strafrechtliche Verantwortlichkeit zuschreiben können, stellt sich umso drängender die Frage, unter welchen Voraussetzungen die Menschen hinter der Maschine für Fehlhandlungen Intelligenter Agenten zur Verantwortung gezogen werden können und müssen. Haftet beispielsweise der Entwickler einer Computertechnologie oder der Betreiber einer Internet-Suchmaschine, wenn diese eine verleumderische Wortkombination anbietet? Macht sich der Konstrukteur oder der Verkäufer strafbar, wenn eine Drohne auf harmlose Personen feuert oder wenn ein fahrerloses Auto einen anderen Verkehrsteilnehmer verletzt oder tötet?

1. Vorsatzstrafbarkeit

Keine Probleme bietet die strafrechtliche Verantwortlichkeit des Betreibers dann, wenn er den Intelligenten Agenten gezielt oder wissentlich so programmiert, dass

⁷⁸ Beck (Anm. 36); Hildebrandt, in: Hildebrandt/Gaakeer (Anm. 19), S. 27f.

dieser strafrechtlich relevante Erfolge herbeiführt: Setzt ein militärischer Befehlshaber eine Drohne bewusst dazu ein, Zivilpersonen zu töten, oder wird ein selbstfahrendes Auto vorsätzlich so programmiert, dass es Fahrradfahrer, die sich auf der Straße bewegen, überfährt, so haben wir keine Schwierigkeit, den Befehlshaber oder den Konstrukteur des Roboterautos als verantwortliche Urheber der schädlichen Folgen zu identifizieren und als Täter des relevanten Straftatbestandes verantwortlich zu machen, denn sie haben sich eines – gleichgültig wie intelligenten – Agenten in einer Weise bedient, wie sich auch sonst Menschen eines Werkzeugs nach ihrem Willen bedienen⁷⁹. Es handelt sich also um einen Fall unmittelbarer vorsätzlicher Täterschaft der Person, die die Maschine nach ihrem Willen steuert.

2. Fahrlässigkeitsstrafbarkeit

Sehr viel häufiger dürften aber die Fälle sein, in denen der Einsatz eines Intelligen Agenten zu einem *ungewollten* Schaden führt: Die Drohne soll nach dem Willen ihres Betreibers⁸⁰ nur das vorgegebene militärische Ziel treffen, weicht jedoch aufgrund einer Fehlfunktion von ihrer Route ab und tötet Zivilisten; das selbstfahrende Auto soll regelgemäß Vorfahrt gewähren, missachtet diese aber und verletzt an einer Kreuzung einen Radfahrer. In diesen Fällen kommt lediglich eine Fahrlässigkeitsstrafbarkeit des Betreibers in Betracht, wenn er den tatbestandsrelevanten Erfolg (also den Tod der Zivilisten oder die Verletzung des Radfahrers) hätte voraussehen können und nicht die notwendige Sorgfalt aufgewandt hat, um diesen Erfolg zu vermeiden⁸¹. Außerdem muss der Erfolgseintritt dem Betreiber zuzurechnen sein. Bei all diesen Punkten stellen sich für den Einsatz Intelligenter Agenten neue Fragen.

⁷⁹ In diese Kategorie fällt etwa auch die so genannte *Google Bomb*, die Google-Suchergebnisse für eine bestimmte Webseite durch vielfaches Setzen von Links manipuliert und diese Seite damit unbrauchbar machen kann. Google ändert heute die Algorithmen der Suchmaschine immer wieder, um solche Manipulationen zu erschweren.

⁸⁰ Der Begriff des „Betreibers“ wird im Folgenden zusammenfassend für alle Personen verwendet, die als Hersteller, Programmierer, Konstrukteure, Verkäufer oder verantwortliche Nutzer möglicherweise strafrechtlich für Intelligente Agenten haften.

⁸¹ Überlegungen zur Fahrlässigkeitshaftung für Intelligente Agenten bei *Hilgendorf*, in: *Hilgendorf/Hötitzsch*, Das Recht vor den Herausforderungen der modernen Technik. Beiträge zur 1. Würzburger Tagung zum Technikrecht, 2014, S. 27f.

a) Vorhersehbarkeit

Bei der Herstellung und dem Betreiben Intelligenter Agenten wirft bereits die Vorhersehbarkeit des Erfolgsintritts als Element der Fahrlässigkeit Probleme auf⁸². Denn Intelligente Agenten werten die aus ihrer Umgebung aufgenommenen Informationen selbständig aus und reagieren dann auf die Ergebnisse dieser Auswertung ohne weiteren menschlichen Einfluss so, dass sie ihren vorgegebenen Auftrag optimal ausführen. Selbst wenn der Betreiber die (Re-)Aktionen eines komplexen Intelligenzen Agenten ständig beobachtet, könnte er also nicht im Einzelnen voraussehen, welche Muster der Roboter in den aufgenommenen Daten erkennen, wie er diese interpretieren und wie er auf sie reagieren wird. Daraus folgt weiter, dass die in einem solchen offenen System vorgegebenen Entscheidungsregeln regelmäßig lückenhaft sind, weil nicht alle Lebenssituationen vorausgesehen und durch Algorithmen in Handlungsanweisungen übersetzt werden können und weil das System im Laufe des Einsatzes außerdem eigenständig hinzulernt.

Damit sind eine gewisse Unberechenbarkeit des Intelligenzen Agenten und die mit ihr verbundenen Gefahren für Dritte „vorprogrammiert“. So kann es etwa sein, dass ein selbstfahrendes Auto ein leicht verschmutztes Verkehrsschild, das „Vorfahrt gewähren“ anordnet, nicht richtig „lesen“ kann und deshalb einem Fahrradfahrer die Vorfahrt nimmt und ihn verletzt. Der Betreiber des selbstfahrenden Autos wird möglicherweise gegen den Vorwurf der fahrlässigen Körperverletzung vorbringen, dass er dieses Fehlverhalten nicht habe vorhersehen können, da das Auto die aufgenommenen Daten autonom verarbeitet habe. Aus diesem Befund kann man zwei konträre Schlüsse ziehen: Man kann entweder sagen, dass den Betreiber eines autonom agierenden Agenten *nie* eine Fahrlässigkeitsstrafbarkeit treffen kann, da das schädliche Tun der Maschine aufgrund deren eigenständiger Informationsverarbeitung nicht vorhersehbar ist. Oder man kann den genau umgekehrten Schluss ziehen, dass nämlich diejenigen, die Intelligente Agenten betreiben, mit „allem“ rechnen müssen, dass also jede Art von Schädigung prinzipiell vorhersehbar ist⁸³.

Die erste Variante hat wenig für sich; denn wer einen von ihm nicht sicher beherrschbaren Intelligenzen Agenten in die Welt setzt, kann seine Haftung für dessen nicht vorhergesehene schadenstiftende Fehlreaktion nicht einfach mit dem Hinweis auf die Unberechenbarkeit des Intelligenzen Agenten verneinen – er kann dies ebenso wenig wie der Zoodirektor, der einen Tiger in die Freiheit

82 Duttge, in: Münchener Kommentar zum StGB, 2. Aufl. 2011, § 15 Rdn. 107 ff.

83 Siehe z. B. Beck, Journal for Robotics and Autonomous Systems (erscheint 2015).

entlässt und gegenüber einem Passanten, der von dem Tiger angefallen und gebissen wird, auf die unberechenbare Natur des wilden Tieres verweist. Wollte man anders entscheiden, so entstünde eine normativ nicht hinnehmbare Verantwortungsdiffusion zwischen Mensch und Intelligentem Agenten, die die tatbestandliche Verletzung des zufälligen Opfers durch das Fehlgehen des Agenten ohne jede Reaktion lässt. Wenn also alle denkbaren Schadensfolgen als für den Betreiber vorhersehbar gelten müssen, dann bietet die Voraussetzung der Vorhersehbarkeit der Tatbestandsverwirklichung keine Hürde gegenüber einer umfassenden Fahrlässigkeitshaftung des Betreibers bei jedwedem schadenstiftenden Verhalten des Intelligenten Agenten.

b) Sorgfaltspflichtverletzung

Fahrlässigkeitsstrafbarkeit setzt neben der Vorhersehbarkeit der Tatbestandsverwirklichung weiter voraus, dass der Täter eine Sorgfaltspflicht verletzt hat⁸⁴. In der Regel folgt die Pflichtverletzung aus der Vorhersehbarkeit des Erfolges: Wer erkennen kann, dass sein Handeln die Verletzung eines geschützten Interesses herbeiführen kann, handelt normalerweise sorgfaltswidrig, wenn er die fragliche Handlung dennoch vornimmt⁸⁵. Speziell für die Herstellung und den Vertrieb von potentiell gefährlichen Produkten hat die Rechtsprechung besonders strenge Sorgfaltsregeln entwickelt, die weitgehend aus dem Zivilrecht ins Strafrecht transferiert worden sind⁸⁶. Danach darf ein sorgfältiger Hersteller nur solche Produkte auf den Markt bringen, deren Sicherheit dem Stand von Wissenschaft und Technik entspricht und vor dem Vertrieb durch ausreichende Tests überprüft worden ist. Auch nachdem der Produzent das Produkt in den Verkehr gebracht hat, muss er es anhand von Rückmeldungen der Verbraucher fortlaufend beobachten. Zeigen sich dabei unerwartete Schäden oder Risiken, so ist der Hersteller zur Warnung der Verbraucher und nötigenfalls zum Rückruf des Produkts verpflichtet⁸⁷.

⁸⁴ Siehe dazu *Vogel*, in: Leipziger Kommentar, 12. Aufl. 2007, § 15 Rdn. 164 ff.

⁸⁵ Siehe dazu *Duttge*, in: MK (Anm. 82), § 15 Rdn. 121 ff.

⁸⁶ Grundlegend BGHSt. 37, 106 („Lederspray“). Zum Verhältnis von zivilrechtlicher und strafrechtlicher Produkthaftung *Gless*, recht 2013, 57 f.; *Kuhlen*, in: *Achenbach/Ransiek*, Handbuch Wirtschaftsstrafrecht, 2012, Rdn. 29 ff.; *ders.*, JZ 1994, 1142; *Schlutz*, DStR 1994, 1811, 1814; *Vogel*, in: LK (Anm. 84), § 15 StGB Rdn. 277–281. Zu der „Lederspray“-Entscheidung siehe *Kuhlen*, NStZ 1990, 566; *Hilgendorf*, NStZ 1994, 561. Grundsätzliche Kritik etwa bei *Hassemer*, Produktverantwortung im modernen Strafrecht, 1994, S. 74.

⁸⁷ Siehe *Kuhlen*, in: *Achenbach/Ransiek* (Anm. 86), Rdn. 40 ff.; *Weigend*, in: LK (Anm. 84), § 13 Rdn. 53 f.

Diese Grundsätze gelten prinzipiell auch (und gerade) in Bezug auf innovative Produkte⁸⁸ wie Intelligente Agenten; hier sind – wie gezeigt – die Risiken besonders hoch und schwer zu kontrollieren. Die drohende strafrechtliche Haftung könnte dann allerdings dazu führen, dass die Herstellung und der Vertrieb von (jedenfalls: manchen) Intelligenten Agenten wegen ihrer fehlenden Steuerbarkeit ganz unterbleiben müsste. Mit dieser Konsequenz wäre zwar ein Gewinn an Sicherheit, andererseits aber der Verlust einer Vielzahl innovativer Anwendungen verbunden, die den Menschen das Leben wesentlich erleichtern könnten. Fahrlässigkeitsstrafbarkeit kann jedoch nicht bedeuten, dass man auf jede möglicherweise riskante Handlung schlechthin verzichten müsste⁸⁹. Wenn man es für sinnvoll, ja sogar vielleicht für geboten hält, dass die Menschen die Dienste Intelligenter Agenten nutzen, so bleibt als Ausweg nur eine Reduktion der gebotenen Sorgfalt⁹⁰. Man müsste bei deren Bestimmung dann die vom Betreiber geschaffenen Risiken in Relation zum Nutzen der betreffenden Technologie setzen. Denn es erschiene unfair, wenn die „Gesellschaft“, die den Nutzen aus der Innovation zieht, das unweigerlich mit ihr verbundene Risiko über das Strafrecht vollständig auf den Betreiber abwälzen könnte.

Betreiber von Intelligenten Agenten könnten sich dann gegenüber dem Vorwurf mangelnder Sorgfalt möglicherweise darauf berufen, dass sie mit der Entwicklung oder Bereitstellung einer lernfähigen und damit nur noch begrenzt kalkulier- und steuerbaren Technologie einem gesellschaftlichen Bedürfnis dienen. Damit ist die entscheidende Frage gestellt, in welchem Umfang die Gesellschaft um des gesteigerten Gesamtnutzens willen bereit ist, die besonderen Gefahren hinzunehmen, die von Intelligenten Agenten für einzelne (oder auch viele) Menschen ausgehen. Die Antwort darauf ist kompliziert, weil es gerade die risikostiftende selbständige Entscheidungsfähigkeit des Intelligenten Agenten ist, die eine vollkommene Kontrolle durch Menschen ausschließt, aber gleichzeitig seinen erhöhten Nutzen begründet und deshalb ökonomisch und sozial erwünscht ist.

Eine eindeutige Auflösung gibt es nicht: Einerseits ist es sicher nicht so, dass die gegenwärtige Gesellschaft sämtliche Entwicklungen in Richtung auf eine maschinengesteuerte Welt kritiklos hinnimmt und die damit verbundenen massiven Risiken – man denke nur an außer Kontrolle geratene bewaffnete Drohnen – ohne Weiteres als Preis des Fortschritts akzeptiert. Andererseits zeigt sich an der öffentlichen Förderung der Forschung an Intelligenten Agenten, dass ein gesellschaftliches Interesse am Einsatz solcher moderner Technologien besteht. Man

88 Kühlen, in: Achenbach/Ransiek (Anm. 86), Rdn. 34 ff.

89 Stratenwerth (Anm. 51), S. 171.

90 Siehe zu möglichen Lösungswegen Eisele, in: Schönke/Schröder, StGB, 29. Aufl. 2014, vor § 13 Rdn. 91; Puppe in: NK (Anm. 47), vor § 13 Rdn. 157 ff.; Vogel, in: LK (Anm. 84), § 15 Rdn. 214 ff.

kann auch annehmen, dass bestimmte Risiken, die mit dem Einsatz von Intelligenen Agenten verbunden sind – etwa die unzulässige Verwendung von Daten, die von Suchmaschinen im Internet gespeichert werden –, bereits allgemein akzeptiert werden. Soweit die Allgemeinheit bereit ist, mit den Risiken der neuen Technologie zu leben, müsste sie auch die Konsequenz akzeptieren, dass unerwartete Reaktionen Intelligenter Agenten zu (generell) vorhersehbaren Schädigungen bei zufällig betroffenen Personen führen, ohne dass die Betreiber strafrechtlich sanktioniert werden⁹¹. Man mag für diesen Fall Versicherungen oder eine zivilrechtliche Gefährdungshaftung vorsehen⁹²; eine umfassende *strafrechtliche* Verantwortlichkeit des Betreibers für generell vorhersehbare Schädigungen müsste man jedoch konsequenterweise ausschließen⁹³.

Andererseits kann der strafrechtsfreie Bereich für die Schaffung unvorhersehbarer Risiken für Leben und Gesundheit Anderer nicht groß sein⁹⁴. Denn die Gesellschaft ist zwar bereit, bestimmte Restrisiken beim Umgang mit Intelligenen Agenten zu akzeptieren; sie muss aber Wert darauf legen, dass diese Gefahren durch sorgfältige Konstruktion und Programmierung, eingehende Tests und kontinuierliche Beobachtung so weit wie möglich minimiert werden. Die entscheidende Frage ist also, welche Anforderungen an die möglichst effektive und weitgehende Begrenzung der mit Intelligenen Agenten verbundenen Risiken zu stellen sind⁹⁵. Dabei kommt es für das Maß des notwendigen Aufwandes maßgeblich darauf an, wie hoch einerseits der Nutzen des jeweiligen Intelligenen Agenten ist und wie gravierend andererseits die Gefahren sind, die von ihm ausgehen können⁹⁶. Bewaffnete Drohnen bedürfen unter beiden Aspekten weitreichender Vorsichtsmaßnahmen, während unerwartete Reaktionen bei Einkäufen im Internet eher hingenommen werden, auch wenn diese etwa zu kleineren Vermögensschäden bei den Benutzern führen können. Je nach ihrem Nutzen wird

91 Kühlen, in: Achenbach/Ransiek (Anm. 86), Rdn. 27 ff.

92 Gless (Anm. 86), S. 57.

93 An welcher Stelle im Delikttaufbau man diesen Ausschluss vornimmt, ist eine zweitrangige Frage. Jedenfalls geht es um eine Begrenzung des Unrechts des jeweiligen Fahrlässigkeitstatbestandes, nicht der individuellen Schuld des Betreibers. Als „dogmatische“ Anknüpfungspunkte kommen eine Begrenzung der Sorgfaltspflicht, eine Verneinung der „unerlaubten“ Risikosteigerung als Voraussetzung (oder Teil) der objektiven Zurechnung des Erfolges (siehe dazu etwa Frisch, GA 2003, 719, 722–724; Puppe, in: NK [Anm. 47], vor § 13 StGB Rdn. 153 ff., 205 ff.) oder eine tatbestands- oder unrechtausschließende „Sozialadäquanz“ der Schaffung eines Risikos in Betracht.

94 Siehe zu einzelnen Fallgruppen Vogel, in: LK (Anm. 84), § 15 Rdn. 216–218.

95 Beck (Anm. 36) sowie aus zivilrechtlicher Sicht Spindler, NJW 2004, 3146 f.

96 Prittwitz, Strafrecht und Risiko, 1993, S. 297 ff.; von Bar, in: Lieb (Hrsg.), Produktverantwortung und Risikoakzeptanz, 1998, S. 40 ff.

man auch für Intelligente Agenten gewisse Entwicklungs- und Anwendungsrisiken in Kauf nehmen dürfen. So kann es zulässig sein, nach hinreichenden Tests ein selbstfahrendes Auto auch dann in den Verkehr zu bringen, wenn es etwa in seltenen Einzelfällen plötzlich stehengeblieben ist, wenn es auf eine ungeplante Situation stieß. Da selbstfahrende Autos für den Einzelnen und für die Lösung allgemeiner Verkehrsprobleme großes Potential aufweisen, können hier großzügigere Maßstäbe gelten als sie einerseits bei technologisch schlichteren Produkten (wie etwa Staubsaugern oder Fahrrädern), andererseits bei Intelligenzen Agenten von geringerem sozialem Mehrwert (etwa Spielzeug oder Kampfdrohnen) angewandt werden müssen⁹⁷.

Allerdings verlangt es die Sorgfalt gegenüber den geschützten Rechtsgütern der betroffenen Menschen, auch bei nützlichen Intelligenzen Agenten die Risiken in größtmöglichem Umfang auszuschließen. So bestehen auch bei einem akzeptierten Innovationsrisiko beispielsweise umfassende Warn- und Hinweispflichten gegenüber den Benutzern sowie eine strikte Verpflichtung der Hersteller, die Erfahrungen mit dem Produkt zu beobachten und umgehend auf Schadensmeldungen zu reagieren⁹⁸. Falls sich etwa zeigt, dass schädliche Verhaltensweisen eines selbstfahrenden Autos nicht nur in ganz seltenen Konstellationen oder bei grob missbräuchlicher Benutzung auftreten, so darf das Auto nicht weiter vertrieben, sondern muss gegebenenfalls zurückgerufen werden⁹⁹. Bleibt der Betreiber untätig, so macht er sich beim Eintritt von Schäden nach dem jeweils anwendbaren Tatbestand strafbar¹⁰⁰. Diese Handlungspflicht des Betreibers ergibt sich – wie auch sonst im Produkthaftungsrecht – aus einer „schlafenden Ingerenz“¹⁰¹: Durch die Einführung oder den Einsatz eines der Natur der Sache nach nicht vollkommen beherrschbaren Intelligenzen Agenten schafft der Betreiber (auch im eigenen wirtschaftlichen Interesse) ein erhöhtes Risiko für die Allgemeinheit, was seine Pflicht verschärft, das Produkt fortlaufend zu überwachen und die notwendigen Konsequenzen aus (nachträglich) erkannten Gefahren zu ziehen¹⁰². Diese

97 Vgl. dazu etwa die Überlegungen zur U.S.-amerikanischen Gesetzgebung zu selbstfahrenden Autos: *Autonomous Vehicles in California*; *Statement of Policy on Automated Vehicles*, U.S. Department of Transportation, <http://www.dmv.ca.gov/vr/autonomous/auto.htm> (besucht am 18.07.2014).

98 Kuhlen, in: *Achenbach/Ransiek* (Anm. 86), Rdn. 34 ff.

99 Gless (Anm. 86), S. 59 ff.

100 Zur Übertragbarkeit von Haftungsüberlegungen BGHSt. 37, 106, 115; dazu *Schmidt-Salzer*, *Produkthaftung*, Bd. I, 2. Aufl. 1987, Rdn. 1.023 ff.

101 Gless (Anm. 86), S. 60 f.; vgl. auch *Hoyer*, GA 1996, 175 f.; *Schmidt-Salzer*, NJW 1996, 6.

102 Siehe für die Schweiz BGE 134 IV 255, 260 f.; für Deutschland *Arzt*, JA 1980, 712, 714; *Freund*, *Erfolgsdelikt und Unterlassen*, 1990, S. 199 ff.; *Jakobs* (Anm. 60), S. 812 f.; *Kuhlen*, NSTZ 1990, 566, 568; *Maiwald*, JuS 1981, 482.

Pflicht ist eine Verkehrssicherungspflicht. Sie folgt aus dem an sich rechtmäßigen Eingehen eines gesteigerten Risikos, nicht wie die Ingerenzhaftung aus einem pflicht- oder rechtswidrigen Vorverhalten¹⁰³.

c) Objektive Zurechnung

Wenn ein Betreiber Intelligenter Agenten unter den erläuterten Bedingungen seine Sorgfaltspflicht verletzt hat, so stellt sich noch die Frage, ob ihm ein eingetretener tatbestandlicher Erfolg („objektiv“) zugerechnet werden kann – oder ob nicht doch „die Maschine schuld“ ist.

aa) Intelligente Agenten als allgemeines Lebensrisiko

Die Zurechnung des eingetretenen Erfolges zum (sorgfaltswidrigen) Verhalten des Betreibers könnte zunächst mit der Begründung ausgeschlossen werden, dass ein „Regressverbot“ bestehe, wenn der unmittelbar Verursachende (d. h. der Intelligente Agent) selbst wegen Vorsatztat strafrechtlich verantwortlich ist¹⁰⁴. Ein so beschriebenes allgemeines „Regressverbot“ wird heute allerdings überwiegend abgelehnt¹⁰⁵, da es prinzipiell möglich ist, dass zwei Personen unabhängig voneinander für ein und denselben Erfolg haften müssen, wenn ihnen dieser Erfolg jeweils objektiv zurechenbar ist¹⁰⁶.

103 *Donatsch/Tag*, Strafrecht I, 8. Aufl. 2006, S. 308; *Cassani*, in: CR, Art. 11 Nr. 44; *Stratenwerth* (Anm. 51), S. 467 f.

104 Zu der Lehre vom Regressverbot eingehend *Puppe*, in: NK (Anm. 47), vor § 13 Rdn. 167 ff. mit zahlreichen Nachweisen; siehe auch *Caro John*, Das erlaubte Kausieren verbotener Taten – Regressverbot, 2007, S. 55 ff.

105 Vgl. *Eisele*, in: *Schönke/Schröder* (Anm. 90), vor § 13 Rdn. 77, 100 ff.; *Frisch*, Tatbestandsmäßiges Verhalten und Zurechnung des Erfolges, 1988, S. 62 f.; *Puppe*, in: NK (Anm. 47), vor § 13 StGB Rdn. 236–238.

106 So kann sich wegen fahrlässiger Tötung strafbar machen, wer eine Schusswaffe vorschriftswidrig ungesichert einer anderen Person zugänglich macht, woraufhin diese Person – für den Waffenbesitzer vorhersehbar – vorsätzlich einen Menschen mit der Waffe erschießt; siehe BGH StraFo 2012, 173 mit Anm. *Berster*, ZIS 2012, 623 (Fall Winnenden). Der kritische Punkt in solchen Fällen ist die Frage, ob der Zurechnungszusammenhang zwischen der fehlenden Sicherung der Waffe und dem Tod des Opfers durch „autonomes“ Dazwischentreten des Schützen unterbrochen wird. Die Antwort auf diese Frage hängt davon ab, ob das vom Ersttäter geschaffene Risiko durch den Zweittäter in eine andere Bahn gelenkt wird (wenn etwa in dem Beispielsfall der Zweittäter die Pistole als tödliches Schlagwerkzeug verwendet) oder ob sich durch den tödlichen Schuss des Zweittäters genau das Risiko verwirklicht, das durch die Pflicht zur sorgfältigen Sicherung der Waffe vor unbefugtem Gebrauch verhindert werden soll.

Mehr Erfolg verspricht der Ansatz, die Zurechnung des von einem Intelligen- ten Agenten hervorgerufenen Erfolges zu dessen Betreiber mit der Begründung abzulehnen, der Erfolgseintritt (z. B. die Verletzung eines Menschen durch ein selbstfahrendes Auto) sei ein „natürliches“, nicht von einem Menschen gesteuertes Ereignis, in dem sich lediglich das allgemeine Lebensrisiko manifestiert hat¹⁰⁷. Dies würde jedoch voraussetzen, dass Intelligente Agenten als normale Erscheinungen des täglichen Lebens wahrgenommen werden, mit deren Risiken jeder vertraut ist und auf die er sich einstellen kann, vergleichbar etwa dem Risiko herabfallender Äste bei einem Waldspaziergang bei Sturm. In der Gegenwart und der nahen Zukunft dürften potentiell schadenstiftende Intelligente Agenten jedoch noch eine Ausnahmerecheinung sein, so dass derjenige, der sie in den Verkehr bringt, ein *außergewöhnliches* Risiko schafft und daher auch für ihre (allgemein vorhersehbaren) Fehlleistungen im Rahmen der Fahrlässigkeitsstrafbarkeit einzustehen hat. Anders ist dies aber wohl schon heute bei Software-Agenten im Internet, die dort als Suchmaschinen oder *recommendation systems* oder in anderer Form allgegenwärtig sind. Obwohl wir um die Gefahr einer Verletzung unserer Persönlichkeitsrechte wissen, bestellen wir weiter Waren bei Internet-Händlern – und akzeptieren als allgemeines Lebensrisiko die Möglichkeit, dass das System unsere persönlichen Daten in einer Weise verarbeitet und verwendet, die in nicht durchschaubarer und nicht kontrollierbarer Weise in unsere Persönlichkeitssphäre eingreift¹⁰⁸. Es spricht also viel dafür, dass der Betreiber eines Software-Agenten im Internet, der andere Menschen den „normalen“ Risiken für den Persönlichkeitsschutz aussetzt, die mit den Algorithmen von Suchmaschinen verbunden sind, für die Betroffenen bereits heute nur ein allgemeines Lebensrisiko aktualisiert. Es kann gut sein, dass eine solche Bewertung von Intelligen- ten Agenten als „normale“ Interaktionspartner der Menschen in der Zukunft nicht auf relativ simple Maschinen im Internet beschränkt bleibt, sondern auch andere Lebensbereiche erreicht. Sobald fahrerlose U-Bahn-Züge und Autos zum selbstverständlichen Bild des Schienen- und Straßenverkehrs gehören, wird auch die Begegnung mit solchen Fahrzeugen für den Menschen zum „Normalrisiko“, und der Betreiber haftet dann (strafrechtlich) nicht für jegliche generell vorhersehbare Fehlfunktion, sondern nur noch für die schädlichen Folgen vermeidbarer Konstruktions- oder Programmierfehler.

107 Siehe *Puppe*, in: NK (Anm. 47), vor § 13 Rdn. 236 ff.; *Walter*, in: LK (Anm. 84), vor § 13 Rdn. 103 ff.

108 Zu dieser Problematik etwa *Ramakrishnan/Keller/Mirza*, Privacy risks in Recommender Systems, abrufbar unter: <http://people.cs.vt.edu/~naren/papers/ppp.pdf> (besucht am 18.07.2014); umfassender zur Frage des Schutzes der Privatsphäre angesichts der Verbreitung Intelligenter Agenten *Nadakavukaren Schefer*, SchwZStrR 2014, 259.

bb) Intelligente Agenten als autonom handelnde Dritte

Eine andere Variante des Ausschlusses der objektiven Zurechnung kann dann gegeben sein, wenn die „eigene“ Entscheidung des autonom agierenden Roboters den Zurechnungszusammenhang zwischen der Tätigkeit des Betreibers und dem Eintritt des strafrechtlich relevanten Erfolges unterbricht¹⁰⁹. Dies liegt hier nicht fern, erscheint doch etwa die Körperverletzung durch ein selbstfahrendes Auto in gewisser Weise nicht als Tat des Konstrukteurs, sondern „des Autos selbst“. Wir hatten zwar oben (II.) festgestellt, dass eine Sanktionierung des Intelligenten Agenten nach „Menschenstrafrecht“ derzeit nicht in Frage kommt. Doch dieser Umstand bedeutet nicht zwingend, dass der Einsatz autonom agierender Technologie für die Zurechnung des Erfolges zu dem Menschen hinter der Maschine irrelevant wäre. Wenn eine Aktion eines Intelligenten Agenten als autonome Handlung eines „Vordermannes“ qualifiziert werden kann, dann kann der Zurechnungszusammenhang zu dem „Hintermann“ (dem Betreiber) abbrechen. Von autonom schädigendem Handeln des Intelligenten Agenten können wir jedoch nicht sprechen, solange wir diesem nicht die Fähigkeit zur eigenen Willensbildung zuschreiben. Dies ist derzeit noch nicht möglich. Oben hatten wir jedoch bereits angedeutet, dass eine solche Zuschreibung bei lern- und erinnerungsfähigen Intelligenten Agenten für die Zukunft nicht ausgeschlossen werden kann¹¹⁰. Doch selbst wenn man bei der Entwicklung anthropomorpher Wesen bis an diesen Punkt gelangt, sollte man den strafrechtlichen Regress auf den fahrlässigen Betreiber zulassen, solange Intelligente Agenten nicht auch bestrafungsfähig werden; andernfalls wäre das Opfer einer schadensbringenden Tat nämlich in der oben bereits angesprochenen unbefriedigenden Situation einer Verantwortungsdiffusion¹¹¹ zwischen Mensch und Intelligentem Agenten, die jede strafrechtliche Ahndung selbst schwerster Schädigungen ausschliesse.

IV. Fazit

Das Vordringen mehr oder weniger segensreicher Intelligenter Agenten in unser Leben ist eines jener Phänomene, das die Rechtsordnung und speziell das Strafrecht vor praktische, aber auch vor grundsätzliche Fragen stellt, die sich mit den hergebrachten Lehren nicht wirklich beantworten lassen. In letzter Konsequenz wirft die Existenz Intelligenter Agenten die Frage auf, was genau das Menschliche

¹⁰⁹ Hierzu eingehend *Roxin*, Allg. Teil I, S. 401ff.

¹¹⁰ Siehe dazu die Nachweise in Anm. 44.

¹¹¹ *Beck* (Anm. 38).

am Menschen ist – eine Frage, die insbesondere für das Strafrecht als das den Menschen höchstpersönlich adressierende Rechtsgebiet von höchster Relevanz ist.

Der strafrechtlichen Zurechnungslehre fehlt bisher das adäquate Instrumentarium, um bei dem Zusammenspiel zwischen dem (möglicherweise wenig intelligenten) Menschen und der (möglicherweise sehr intelligenten) Maschine über die sachadäquate Zuweisung von Verantwortung zu entscheiden. Dabei dürfte die Einführung einer Strafbarkeit Intelligenter Agenten für die voraussehbare Zukunft ein bloßes Denkmodell bleiben. Intelligente Agenten verfügen weder über eine – für einen Menschen nachvollziehbare – Strafeempfindlichkeit noch über Verständnis für den mit der Bestrafung verbundenen ethischen Vorwurf, so dass der Einsatz des Strafrechts gegen sie „persönlich“ nicht sinnvoll wäre. Anders als bei einer Strafbarkeit juristischer Personen ließe sich eine Bestrafung Intelligenter Agenten auch nicht als mediatisierte Bestrafung der im Hintergrund stehenden natürlichen Personen für deren „Organisationsverschulden“ interpretieren. Jedenfalls könnte man ein solches individuelles Verschulden der Betreiber des Intelligenzen Agenten ebenso wirksam durch die unmittelbare Bestrafung der Betreiber wegen ihres eigenen vorsätzlichen oder fahrlässigen Kausalbeitrags zu der von dem Agenten verursachten Schädigung sanktionieren. Es ist zwar durchaus denkbar, den Intelligenzen Agenten als selbständiges Subjekt zivilrechtlicher Haftung zu konstruieren (und ihn zu diesem Zweck mit einer Haftpflichtversicherung oder einem Haftungsfonds auszustatten)¹¹², und man kann auch erwägen, die Zerstörung oder Re-Programmierung eines außer Kontrolle geratenen Intelligenzen Agenten als strafrechtliche Sanktion gegen seinen Eigentümer anzuordnen; aber eine genuin strafrechtliche Verantwortlichkeit des Agenten selbst kommt nicht ernsthaft in Betracht, solange der Intelligente Agent nicht auch zum Moralischen Agenten geworden ist.

Schwerer ist die Frage nach der Strafbarkeit des „Menschen hinter der Maschine“ zu beantworten. Ruft die Aktion eines Intelligenzen Agenten einen strafrechtlich relevanten Schaden hervor, so haftet der Mensch jedenfalls dann als unmittelbarer Täter, wenn er den Agenten in der Absicht oder in dem Wissen benutzt hat, dass mit seiner Hilfe ein Straftatbestand verwirklicht wird oder werden kann. Denn der Mensch kann die Schuld nicht auf die Maschine abschieben, wenn er sie gezielt als Werkzeug einsetzt oder wenn er jedenfalls weiß und für sich akzeptiert, dass der Intelligente Agent bei dem konkret vorgesehenen Einsatz einen bestimmten strafrechtlich relevanten Erfolg verwirklichen kann.

112 Siehe Gruber (Anm. 44), S. 157; ders., in: Beck (Anm. 21), S. 150 ff.

Auch wenn die Aktion eines Intelligenten Agenten ohne Wissen und Willen des Betreibers zur objektiven Verwirklichung eines Straftatbestandes führt, kann die Schuld des Betreibers nicht mit einem lapidaren Verweis auf den „Automatismus“ von Algorithmen verneint werden. Der Finger, der auf einen – durch mathematische Handlungsanweisungen vorgeschriebenen – Automatismus zeigt, weist letztlich auf den Menschen zurück, der das Verhalten des Intelligenten Agenten und damit die Möglichkeit der Schädigung programmiert oder durch den Vertrieb des Agenten ermöglicht hat. Im Fall bloßer Fahrlässigkeit tauchen mögliche Probleme auf der Ebene der Vorhersehbarkeit und der Erfolgszurechnung auf. Komplexe Intelligente Agenten sind in ihrem Entscheidungsverhalten nicht vollständig berechenbar, so dass man die Vorhersehbarkeit ihres Fehlverhaltens für den Betreiber in Frage stellen kann. Auch kann man ihr Verhalten in gewissem Umfang als selbstbestimmt ansehen. Dies sollte einem strafrechtlichen Rückgriff auf eine Fahrlässigkeit des Betreibers freilich nur dann im Wege stehen, wenn das Verhalten des Intelligenten Agenten nach den allgemeinen Regeln über die Zurechnung von Erfolgen den Zurechnungszusammenhang zwischen dem Verhalten des Betreibers und dem Erfolgseintritt unterbricht. Dies ist etwa dann denkbar, wenn das Tätigwerden Intelligenter Agenten in einem bestimmten Lebensbereich so sehr zur Normalität geworden ist, dass auch eine ungeplante Fehlfunktion zum allgemeinen Lebensrisiko des Betroffenen zu rechnen ist.

Eine ganz andere Frage ist es, ob ein kleines („Rest“-)Risiko, das sich aus der Lernfähigkeit eines Intelligenten Agenten ergibt, um des großen sozialen Nutzens eines solchen Agenten willen hingenommen werden kann. Diese Frage stellt sich bereits auf der Ebene des Fahrlässigkeitsunrechts, nämlich als mögliche Begrenzung der Sorgfaltspflicht des Betreibers. Diese Problematik dürfte die Diskussion um die strafrechtliche Haftung für Intelligente Agenten in der näheren Zukunft in erster Linie bestimmen. Über Intelligente Agenten hat der Mensch – anders als bei anderen Produkten und Erfindungen – keine perfekte Kontrolle; er kann die mit ihnen verbundenen Risiken nicht vollkommen planen und beherrschen. Je vielseitiger, lernfähiger und „intelligenter“ Roboter werden, desto mehr entziehen sie sich einer vorausschauenden Planung ihres Betreibers und desto weniger wird man einzelne Konstruktions- oder Programmierungsfehler identifizieren können, die eine Fahrlässigkeitsstrafbarkeit der beteiligten Personen zu begründen vermöchten. Will man sich nicht wegen der letztlich nicht beherrschbaren Risiken auf ein Totalverbot Intelligenter Agenten verständigen, so dürfte keine andere Möglichkeit bestehen, als die nicht verlässlich durch Programmierung und verantwortlichen Einsatz steuerbaren Gefahren der „Gesellschaft“ aufzubürden, indem man auf eine strafrechtliche Fahrlässigkeitsverantwortlichkeit verzichtet, also die geschädigten Personen als Opfer nicht-menschlichen Verhaltens definiert. Dabei ist es nicht notwendig, eine allgemeine und umfassende Freistel-

lung der Menschen hinter der Maschine festzuschreiben; vielmehr bietet es sich an, das Strafrecht nur in dem Maße zurückzunehmen, wie der allgemeine soziale Nutzen des Intelligenen Agenten die Hinnahme der mit ihm verbundenen Risiken für die Allgemeinheit rechtfertigt.

Rechtsmethodisch kann man dieses Ergebnis durch die Einführung flexibler Generalklauseln erreichen, etwa indem man die erforderliche Sorgfalt des Betreibers durch die Gestattung eines „Entwicklungsrisikos“ beschränkt oder indem man einen Rechtfertigungsgrund des „sozialadäquaten Einsatzes“ von Intelligenen Agenten schafft. Im Sinne der Gesetzesbestimmtheit wäre allerdings der mühsamere Weg vorzuziehen, die Maßstäbe für das erlaubte Risiko bei verschiedenen Typen und Einsatzbereichen Intelligenter Agenten durch Gesetz oder Verordnung festzulegen. Damit ließe sich in einem demokratisch legitimierten Prozess ein angemessener Ausgleich schaffen zwischen dem Bedürfnis, die vielversprechenden Möglichkeiten selbständig agierender Helfer des Menschen ohne Furcht vor unübersehbaren strafrechtlichen Konsequenzen einzusetzen, und der Schaffung notwendiger Anreize zu einer möglichst weitgehenden Eindämmung der beherrschbaren Risiken in der Schönen Neuen Welt¹¹³ der Intelligenen Agenten.

¹¹³ Huxley, *Brave New World*, 1932, siehe: https://archive.org/details/ost-english-brave_new_world_aldous_huxley (besucht am 18.07.2014).