

# Intelligente Agenten als „Personen“ im Strafrecht?

Prof. Dr. Gerhard Seher, Freie Universität Berlin

## A. Vorbemerkungen – Begriffliches – Gedankengang

I. Das Strafrecht ist dazu gemacht, Menschen dafür zu sanktionieren, dass sie grundlegende Regeln des rechtlichen Zusammenlebens verletzt haben. Seit jene alten Zeiten vergangen waren, in denen man – angeblich – auch Tiere für schädigendes Verhalten verurteilte und bestrafte,<sup>1</sup> gab es keinen Zweifel daran, dass die Normen des Strafrechts auf Menschen gemünzt sind. Das Gesetz scheidet als Täter aus, „wer bei Begehung der Tat noch nicht vierzehn Jahre alt war“ (§ 19 StGB<sup>2</sup>) oder wer „wegen einer krankhaften seelischen Störung, wegen einer tiefgreifenden Bewusstseinsstörung oder wegen Schwachsinn oder einer schweren anderen seelischen Abartigkeit unfähig ist, das Unrecht der Tat einzusehen oder nach dieser Einsicht zu handeln“ (§ 20 StGB) – Umstände, die nur mit Blick auf menschliche Täter von normativer Relevanz sein oder überhaupt nur auf Menschen zutreffen können. § 11 StGB enthält für besondere Gruppen von Straftätern (unter anderem) „Personenbegriffe“, die ebenfalls nur auf menschliche Akteure passen (Angehörige, Amtsträger, Richter usw.). Und die intensive Diskussion um eine Strafbarkeit von Wirtschaftsunternehmen hat jedenfalls in der deutschen Diskussion bisher sehr deutlich gezeigt, dass das Strafrechtssystem derzeit auf nicht-menschliche Adressaten nicht zugeschnitten ist.<sup>3</sup> Die Adressaten des Strafrechts sind *Menschen* als

- 
- 1 Dazu P. Dinzelbacher, Das fremde Mittelalter. Gottesurteil und Tierprozess, Essen 2006; zweifelnd gegenüber Strafverfahren gegen und Verurteilungen von Tieren jedenfalls für den deutschen Raum E. Schumann, in: B. Herrmann (Hrsg.), Beiträge zum Göttinger Umwelthistorischen Kolloquium 2008-2009, Göttingen 2009, S. 182 ff.
  - 2 Soweit nicht anders angegeben, beziehen sich alle Gesetzesangaben auf das deutsche Strafgesetzbuch.
  - 3 S. zur Diskussion um ein Unternehmensstrafrecht z.B. E. Kempf, K. Lüderssen, K. Volk (Hrsg.), Unternehmensstrafrecht, Berlin/Boston 2012; F. von Freier, Kritik der Verbandsstrafe, Berlin 1998; ders., Zurück hinter die Aufklärung: Zur Wiedereinführung von Verbandsstrafen, GA 2009, S. 98 ff.

Teilnehmer, Mitwirkende und Unterworfenen des Normensystems „Recht“ – *Personen im Recht*.

Nun aber treten im Zuge der rasanten digitalen Technisierung neue Akteure auf den Plan, die nicht menschlich sind, aber aufgrund komplexer Programmierung Aktionen ausführen, die nach herkömmlichem Verständnis nur Menschen bewältigen können: Roboter, selbstfahrende Autos, Drohnen; oder Computer, die Funktionen selbstständig ausführen, hinter denen der Laie intuitiv einen Menschen denkt, der den Computer steuert. Können solche „intelligenten Agenten“ als Adressaten des Strafrechts angesehen werden? Können sie unter die Rolle subsumiert werden, die man bislang exklusiv Menschen – als *Rechtspersonen* – vorbehalten hat? Oder sind und bleiben sie Werkzeuge in der Hand von Menschen, die sich ihrer bedienen, wie man sich eines Messers oder eines Fahrrades bedient?

Aus diesen Fragen ergeben sich zwei Perspektiven der Betrachtung: Man kann nach einer strafrechtlichen Verantwortlichkeit der intelligenten Agenten selbst fragen oder nach der Verantwortlichkeit von Menschen für die Aktionen solcher Agenten. Beide Aspekte werden zur Sprache kommen, und es wird sich erweisen, dass ihnen auf den verschiedenen Stufen der Straftat unterschiedliche Relevanz zukommt.

**II.** Im Hintergrund der folgenden Gedanken stehen „intelligente Agenten“. Eine allgemeine Definition für diesen Begriff ist nicht leicht zu finden. Deshalb sei vorausgesetzt, dass hier nur solche „Roboter“, „Maschinen“ oder „Software-Agenten“ von Interesse sind, die so programmiert sind, dass sie empfangene Informationen in einer Weise verknüpfen und für künftige Reaktionen auswerten können, die nicht vollständig durch die Programmierung festgelegt ist und deren Aktionen daher *ex ante* „offen“ und mithin nicht für jeden Einzelfall vorhersehbar sind. Nur hinsichtlich solcher „lernfähiger“ bzw. „selbst entscheidender“ Agenten kann überhaupt die Frage gestellt werden, ob ihnen das Strafrecht eine eigene, personale Rolle zuweisen kann. Alle anderen, „fest“ programmierten Geräte können – so will ich es unterstellen – nur als „Werkzeuge“ in der Hand menschlicher Agenten in Betracht kommen.

**III.** Im Folgenden sind die Stufen zu betrachten, in denen entschieden wird, ob ein Geschehen als schuldhaft begangene Straftat anzusehen ist. Dabei geht es um vier wohlbekanntes Fragestellungen: (1) Ist das Geschehen nach objektiv-normativen Kriterien einer Person als „ihre Tat“ zuzurechnen? – (2) Kann diese Person persönlich dafür verantwortlich gemacht werden, d.h. hat sie dieses Geschehen vorsätzlich geprägt oder aus Fahrlässigkeit eintreten lassen? – (3): Ist gegenüber dieser Person ein Schuld-

vorwurf zu erheben? – und (4): Ist eine Bestrafung „intelligenter Agenten“ denkbar?

## B. Handlung und Taterfolg

Dem Straftatabau folgend, stellt sich zunächst die Frage, wann ein Geschehen nach objektiv-normativen Kriterien als „die Tat“ einer bestimmten Person zu qualifizieren ist – wann sie also den objektiven Tatbestand einer Strafnorm erfüllt hat. Bei diesem ersten Schritt spielt es bereits eine wesentliche Rolle, ob *intelligente Agenten* menschlichen Akteuren gleichgestellt werden können oder nicht.

### I. Normtheoretische Grundlegung

Die Antwort auf diese Frage beginnt mit einer normtheoretischen Überlegung: Welchem Zweck dient das Recht, und auf welche Weise will das Recht diesen Zweck erreichen? – Das Recht will als staatliches Normensystem das Leben derjenigen, die ihm unterworfen sind, auf verlässliche Weise gestalten und dadurch bestimmte Werte – etwa Freiheit, Sicherheit, Frieden oder Wohlstand – verwirklichen. Diese Zwecke verfolgt es durch Verhaltensbefehle und durch die Anordnung bestimmter Folgen für die Befolgung oder Missachtung dieser Befehle.<sup>4</sup>

Wenn diese Verhaltensbefehle sinnvoll sein, d.h. eine Wirkung entfalten sollen, dann müssen sie sich an Personen richten, die diese Befehle *verstehen und befolgen können*. Das Recht kann also nur an Personen gerichtet sein, die *normativ ansprechbar* sind. Befehle lassen sich an niemanden

---

4 S. zu Verhaltensnormen als rechtlichem Instrument der Verhaltenssteuerung J. Hruschka, Verhaltensregeln und Zurechnungsregeln, *Rechtstheorie* 22/1991, S. 449 ff.; J. Renzikowski, Restriktiver Täterbegriff und fahrlässige Beteiligung, *Tübingen* 1997, S. 54-58.

richten, der sie nicht verstehen und deshalb nicht befolgen kann.<sup>5</sup> Wer aber kann Befehle verstehen und befolgen?

## II. Rechtliche Handlungsbegriffe

Die Antwort hierauf lautet in einer Jahrhunderte alten Tradition: Wer *handeln* kann. Der Begriff der Handlung ist damit das Eingangstor für die rechtliche Normenordnung.

„Handlung“ in diesem starken Verständnis des Wortes ist dann natürlich nicht jede Körperbewegung und erst recht nicht jede beliebige auf einen Menschen zurückführbare Veränderung in der Welt, sondern nur ein Verhalten, das durch Normbefehle *steuerbar* ist. Darin liegt ein terminologischer *Zuschnitt*: Der Begriff der Handlung wird so hergerichtet, dass er eine sinnvolle rechtliche Funktion erfüllen kann.<sup>6</sup>

Ein solcher normativ angereicherter Handlungsbegriff gehört allerdings nicht zum gesicherten Bestand der Strafrechtsdogmatik. Schon in den letzten Jahrzehnten des 19. Jahrhunderts setzte eine naturalistische, auf kausale Strukturen von Ursache und Wirkung fixierte Dogmatik beim physischen *Willensimpuls* an und fragte einzig, welche Körperbewegung dieser

---

5 Ein schönes Beispiel für die Grenzen von Befehlsgewalt bietet jene Parabel aus dem *Kleinen Prinzen*, in der dieser auf einem kleinen Planeten einen General besucht, der behauptet, über alle Planeten und Sterne befehlen zu können. Der Kleine Prinz wünscht sich einen Sonnenuntergang, aber der General will ihn nicht sofort befehlen. Er antwortet dem Kleinen Prinzen: „Wenn ich einem General geböte, nach der Art eines Schmetterlings von einer Blume zur anderen zu fliegen oder eine Tragödie zu schreiben ..., und wenn der General den erhaltenen Befehl nicht ausführte, wer wäre im Unrecht, er oder ich?“ ... „Man muss von jedem fordern, was er leisten kann. Die Autorität beruht vor allem auf der Vernunft.“ *A. de Saint-Exupéry*, *Der Kleine Prinz*, 54. Aufl., Düsseldorf 1999, S. 40.

6 Solche Begriffszuschnitte sind in der Rechtsdogmatik nichts Ungewöhnliches. Der „Besitz“ etwa umfasst nach § 868 BGB auch Rechtsbeziehungen eines Menschen zu einer Sache, die nach üblichem sozialen Verständnis nicht dazugehören („mittelbarer Besitz“), und eine zweifelhafte Berühmtheit hat die Rechtsprechung des Bundesgerichtshofes erlangt, nach der ein Täter einen Taterfolg auch dann *vorsätzlich* herbeiführt, wenn er ihn zwar nicht will, aber – vielleicht aus innerer Gleichgültigkeit – „billigend in Kauf nimmt“ (BGHSt 7, 363; BGH NSTZ 2007, 700, 701); auch dieser Rechtsbegriff ist vom normalsprachlichen Sinn eines „Vorsatzes“ spürbar entfernt.

Impuls bewirkt habe.<sup>7</sup> Damit wurde der maßgebliche normtheoretische Zusammenhang ausgeblendet, der in früheren Zeiten – in der Epoche des Naturrechts und ebenso bei *Kant* und *Hegel* – noch selbstverständlich gewesen war: dass es darauf ankomme, dass ein freier, sich der Normen von Recht und Moral *bewusster Wille* die Entscheidung für ein bestimmtes Verhalten getroffen habe, das eben dadurch nicht bloßes „Verhalten“, sondern eine „Handlung“ sei.<sup>8</sup>

So eindeutig der Begriff der Handlung stets als maßgebend für „rechtlich relevantes Verhalten“ angesehen wurde, so sehr variierte sein Bedeutungsgehalt in den verschiedenen Epochen seiner Dogmengeschichte. Diese Bewegungen – die, einer Sinuskurve gleich, hohe, geringe oder gar keine Ansprüche an die Freiheit des Willens stellten – brauchen hier nicht im Einzelnen nachgezeichnet zu werden. Entscheidend ist – mit Blick auf Aktionen intelligenter Agenten – allein, dass nur ein solches Verhalten als rechtlich relevante *Handlung* qualifiziert werden kann, das durch ein – aktuelles oder zumindest potentiell – *Normverstehen* geprägt ist.<sup>9</sup>

Das gilt selbst dann – einschlägig –, wenn man annimmt, dass menschliches Verhalten durchgängig determiniert ist, sich also kein in qualifizierter Weise freier Wille als Ursache menschlicher Handlungen ausmachen lässt. Denn auch ein neuronal vollständig determinierter Wille ist durch Normbefehle beeinflussbar, weil auch eine determinierte Psyche Befehle *verstehen* und reflektieren kann. Menschen können deshalb auch von

---

7 Paradigmatisch *F. von Liszt*, Lehrbuch des Deutschen Strafrechts, 6. Aufl., Berlin 1894, S. 99-103.

8 S. zum naturrechtlichen Handlungsbegriff *S. Pufendorf*, Elementorum Jurisprudentiae Universalis libri duo, Cygnaeae 1668, Lib. II, Axioma I (Handlung ist ein Verhalten, das an moralischen Normen ausgerichtet werden kann); und in dieser indeterministischen Tradition noch *I. Kant*, Metaphysik der Sitten, Königsberg 1797, S. 22 (Begriffe der Tat und der Person) und *G.W.F. Hegel*, Grundlinien der Philosophie des Rechts, Berlin 1821, §§ 113, 114: Handlung sei die Äußerung des zwecksetzenden, moralischen Willens.

9 Verzichtet man auf das Erfordernis des (zumindest potentiellen) Normverstehens, dann verzichtet man letztlich auf einen Handlungsbegriff insgesamt. Das verlagert die Frage auf die Schuldebene – und wirft das neue Problem auf, wie jemand Unrecht verwirklichen kann, ohne von rechtlichen Normbefehlen erreicht werden zu können. Jedenfalls mit der heute sehr verbreiteten personalen Unrechtslehre, die eine Normverletzung für die Unrechtsbegründung verlangt (dazu *K. Kühl*, Strafrecht AT, 7. Aufl., München 2012, § 5 Rn. 4) dürfte das schwer vereinbar sein. Aber dies im hiesigen Kontext nur am Rande.

einem strikt deterministischen Standpunkt aus als taugliche Adressaten von Rechtsnormen begriffen werden – und damit *handeln*.<sup>10</sup>

Die Fähigkeit, für Rechtsnormen *ansprechbar* zu sein, ist also notwendige Voraussetzung dafür, in rechtlich relevanter Weise *handeln* zu können. Wer Rechtsnormen in ihrem Sollgehalt inhaltlich nicht verstehen kann, scheidet als ihr Adressat aus. Es fehlt ihm die Minimalvoraussetzung dafür, als *Person* innerhalb des Systems des Rechts teilnehmen und anerkannt werden zu können.

Mit Blick auf „intelligente Agenten“ bedeutet dies, dass sie nur dann als eigenständige Adressaten des Rechts – und damit als *Rechtspersonen* – in Betracht kämen, wenn sie in der Lage wären, Normbefehle als solche wahrzunehmen, zu verstehen und in ihre Reaktionen mit einzubeziehen. Am „Verstehen“ von Normen als Befehlen dürfte es aber auf absehbare Zeit fehlen, da diese „Agenten“ allein aufgrund zuvor installierter Programmierung agieren, ohne die Rechtsnormen auf eine eigene, autonome Weise zu erfassen. Sie können zwar einprogrammierten Normen gemäß agieren, aber diese Programmierung ist immer vermittelt durch einen Menschen, der sie vornimmt. Allein dieser Mensch kann Adressat der jeweils einschlägigen Rechtsnormen sein, nicht aber die Entität, die darauf programmiert wird, den normgemäß gestalteten Algorithmus umzusetzen. Computer sind – soweit ich informiert bin – nicht in der Lage, von sich aus Normen *als solche* wahrzunehmen. Das mag man bei selbstfahrenden Autos möglicherweise anders sehen, soweit sie in der Lage sind, Verkehrsschilder und Ampelsignale richtig zu „lesen“, denn hier handelt es sich um schematisierte, symbolisierte Normen, die für einen Computer unmittelbar wahrnehmbar und umsetzbar sind. Anders sieht es dagegen bei der großen Menge der nicht symbolisch kommunizierten Normen aus: Ein Gerichtsurteil zum Beispiel, das ein abstraktes Gesetz für eine bestimmte Fallgruppe verbindlich konkretisiert, kann von einem „intelligenten Agenten“ nicht selbstständig in seiner Relevanz erkannt werden. Ändert etwa der Bundesgerichtshof den Mindestseitenabstand, den ein überholendes Fahrzeug zum überholten einhalten muss, dann muss diese neue Regel von einem Menschen in das Programm des „Agenten“ hineingeschrieben werden, bevor letzterer in der Lage ist, sie einzuhalten.

---

10 Dass die Handlungsfähigkeit nicht an die Annahme eines philosophischen Indeterminismus gebunden ist, ist allgemeine Meinung. Diese Frage wird allerdings heute bei der Schuldfähigkeit verortet (s. etwa C. Roxin, Strafrecht AT, Band I, 4. Aufl., München 2006, § 3 Rn. 55; Kühl, Strafrecht AT (Fn. 9), § 11, Rn. 3 f.).

„Rechtlich relevante Handlung“ ist also – ganz unabhängig aller Diskussionen über Willensfreiheit oder Determinismus – nur das Verhalten einer „Person“, die durch die Normen des Rechts unmittelbar angesprochen werden kann und die von ihrer Konstitution her in der Lage ist, den Inhalt der Normen *unmittelbar* zu verstehen und reflektierend in die Entscheidungen über nachfolgendes Verhalten mit einzubeziehen. Diese Fähigkeit fehlt intelligenten Agenten jedenfalls solange, bis man ihnen beigebracht hat, zu lesen und Fernsehnachrichten begreifend zu sehen.<sup>11</sup>

### III. Tathandlung, Kausalität und objektive Zurechnung

Neben der „Handlung“ als allgemeinster Voraussetzung bedarf es weiterer Schritte, um zu entscheiden, ob ein Geschehen als Erfüllung eines Straftatbestandes zu deuten ist. Dabei soll es hier nur um die Tatbestände des Kernstrafrechts gehen: solche, die schädigende Handlungsfolgen pönalisieren (Erfolgsdelikte). Und unter der Voraussetzung, dass „intelligenten Agenten“ bereits die strafrechtliche Handlungsfähigkeit fehlt, beschränkt sich der Blick nun auf das Verhalten der menschlichen Akteure, die einen „intelligenten Agenten“ in Betrieb setzen.

Jemand kann – von Unterlassungen abgesehen – nur dann für eine Handlungsfolge rechtlich zuständig sein, wenn er sie *verursacht* hat. Die *Kausalität* zwischen Handlung und schädigender Folge ist daher ein unbestrittenes Element der Straftat. Aber Kausalität ist eine deskriptive Denkkategorie: Wir setzen sie ein, um Ereignisse zu beschreiben und empirische Gesetzmäßigkeiten zwischen ihnen zu erkennen. Mangels normativer Dimension kann aber *rechtliche* Verantwortung auf Verursachung nicht gestützt werden. Es bedarf vielmehr eines *wertenden* Schrittes, der unter allen Ursachen einer schädigenden Geschehensfolge diejenige oder diejenigen auswählt, die als rechtlich relevant angesehen wird bzw. werden. Da es hierbei nicht darum geht, Umstände zu erkennen oder zu beweisen, sondern sie zu *bewerten*, spricht man von *Zurechnung*. Die Zurechnung auf der Ebene des Tatbestandes setzt voraus, dass jemand – erstens – eine Handlung ausgeführt hat, die deshalb strafrechtlich relevant ist, weil sie unter die Handlungsbeschreibung eines gesetzlichen Tatbestandes subsu-

---

11 Das sehen *Gless/Weigend*, ZStW 126 (2014), S. 571-573 vorsichtig anders, weil sie bereit sind, von einem normtheoretisch offenen, deskriptiven Handlungsbegriff auszugehen.

miert werden kann („töten“, „fälschen“, „nötigen“), und – zweitens – ihre spezifische Gefährlichkeit sich in einer Folge niedergeschlagen hat, die eben deshalb als „Erfolg“ dieser Handlung qualifiziert werden kann.<sup>12</sup>

Wenn eine direkte strafrechtliche Haftung „intelligenter Agenten“ mangels Handlungsfähigkeit nicht in Betracht kommt, rückt die Frage in den Mittelpunkt, unter welchen Voraussetzungen Menschen, die solche Agenten herstellen, besitzen oder nutzen, für schädigende Aktionen dieser Geräte strafrechtlich einstehen müssen. Einem Vorschlag von *Gless/Weigend* folgend<sup>13</sup>, bezeichne ich diese Personen pauschal als „Betreiber“.

Ein Betreiber kann also nach der Norm, mittels derer auf der Tatbestandsebene die Zurechnung vollzogen wird, für einen Schaden, den ein „intelligenter Agent“ verursacht hat, verantwortlich gemacht werden, wenn seine eigene Handlung eine Gefahr geschaffen hat, die sich in dem Schaden niedergeschlagen hat. Hier kommen natürlich – wie *Gless* und *Weigend* bereits treffend analysiert haben<sup>14</sup> – vor allem Fehlprogrammierungen und Bedienungsfehler in Betracht. Diese Fälle aber sind rechtlich nicht sehr spannend, denn die Gefährlichkeit der Handlung ist offensichtlich.

Interessant sind jene Fälle, in denen ein „intelligenter Agent“ aufgrund einer abstrakt korrekten Programmierung selbstständig Entscheidungen trifft, die zu einer Schädigung führen.<sup>15</sup> Hier stellt sich zunächst die Frage, worin die gefährliche Handlung des Betreibers liegen kann, denn „intelligente Agenten“ bergen strukturell das Risiko, dass einzelne ihrer Aktionen Schäden hervorrufen. Bei selbstfahrenden Autos und autonom agierenden Operationscomputern etwa liegt unmittelbar nahe, welche Körperschäden drohen können. Sie in Betrieb zu setzen, schafft daher dieses abstrakte Risiko auch dann, wenn Konstruktion und Programmierung fehlerfrei waren.

---

12 Verbreitet, aber verkürzend ist die sogenannte „Grundformel“ für die objektive Zurechnung, derzufolge ein Taterfolg dem Handelnden dann objektiv zurechenbar ist, wenn dessen Handlung eine für das geschützte Rechtsgut relevante (oder missbilligte) Gefahr geschaffen und diese Gefahr sich im eingetretenen Taterfolg realisiert hat (s. *Roxin*, Strafrecht AT (Fn. 10), § 11 Rn. 47 ff.; *J. Wessels/W. Beulke/H Satzger*; Strafrecht AT, 45. Aufl., Heidelberg 2015, Rn. 251 f.).

13 *S. Gless/Th. Weigend*, Intelligente Agenten und das Strafrecht, ZStW 126 (2014), S. 580 Fn. 80.

14 *Gless/Weigend* (Fn. 13), S. 579 ff.

15 Zu solchen Fällen s. *S. Kirn/C.-D. Müller-Hengstenberg*, Technische und rechtliche Betrachtungen zur Autonomie kooperativ-intelligenter Softwareagenten, Künstliche Intelligenz, Vol. 29, 1 (2015), S. 59, 62.



Dies ist vergleichbar mit dem Betrieb von herkömmlichen Autos, Eisenbahnen und Flugzeugen – Verkehrsmittel, die das Recht seit jeher als gefährliche Gegenstände einstuft.<sup>16</sup>

Führt aber – das ist die zweite hier einschlägige Frage – diese abstrakte Gefährlichkeit des Betriebens „intelligenter Agenten“ stets zur Zurechnung von Schäden, die durch diese ausgelöst werden? Nach der soeben zitierten allgemeinen Formel der tatbestandlichen Zurechnung scheint das ohne weiteres der Fall zu sein, denn wenn die offene Programmierung ein abstrakt unbeherrschbares Risiko birgt, sind diese Agenten eben in rechtlich relevanter Weise gefährlich, und wenn infolge dieser Programmierung ein tatbestandserfüllender Schaden eintritt, dann hat sich diese Gefährlichkeit in dem Taterfolg realisiert. Aber es gibt verschiedene Fallgruppen, in denen die Zurechnung ausgeschlossen wird.<sup>17</sup> Drei von ihnen können hier relevant werden:

1. Die schädigende Fehlaktion des „intelligenten Agenten“ könnte aus Sicht des gewissenhaften Betreibers ein *atypischer Kausalverlauf* sein – d.h. ein solcher, der nach allgemeiner Lebenserfahrung objektiv unvorhersehbar war.<sup>18</sup> Ob man singuläre Fehlentscheidungen eines an sich korrekt programmierten „intelligenten Agenten“ als unvorhersehbar ansehen kann, haben *Gless* und *Weigend* gründlich diskutiert – und mit überzeugender Begründung abgelehnt.<sup>19</sup> Zwar mag man von der konkreten Entscheidung des „Agenten“ überrascht sein, weil niemand sie für den konkreten Fall vorhersehen konnte. Die Lebenserfahrung gründet sich aber auf das Wissen um viele Fälle und Situationen, so dass richtigerweise zu fragen ist, ob *generell* vorhergesehen werden kann, dass lernfähige, selbst entscheidende Computer in Ausnahmefällen schädigende Aktionen ausführen – und das wird man ohne weiteres bejahen können. „Atypisch“ ist also eine schädigende Fehlaktion des „intelligenten Agenten“ aus Sicht eines Betreibers nicht.

---

16 Das bringen die Normen zur abstrakten Halterhaftung deutlich zur Geltung: für Autos § 7 StVG, für Flugzeuge § 33 LuftVG.

17 S. dazu allgemein *Wessels/Beulke/Satzger*, Strafrecht AT (Fn. 12), Rn. 254-301.

18 S. zum atypischen Kausalverlauf als Ausschlussgrund der objektiven Zurechnung *Kühl*, Strafrecht AT, (Fn. 9) § 4 Rn. 61 ff.

19 *Gless/Weigend* (Fn. 13), S. 581 f. – unter dem einschränkenden Aspekt subjektiver Vorhersehbarkeit.

2. Ein Zurechnungsausschluss kommt auch in Betracht im Falle des *eigenverantwortlichen Dazwischentretens Dritter*.<sup>20</sup> Diese Konstellation läge dann vor, wenn man davon ausginge, dass der „intelligente Agent“, dessen Aktion zwischen den Handlungen des Betreibers und dem Schadenseintritt liegt, zu eigenverantwortlichem Handeln in der Lage ist. Die Zweifel daran wurden bereits im Abschnitt über den Handlungsbegriff thematisiert. Wer aber nicht im Rechtssinne „handeln“ kann, dem ist auch die Möglichkeit eigenverantwortlichen Handelns abzusprechen.
3. Schließlich könnte sich das Betreiben des „intelligenten Agenten“ im Rahmen eines *erlaubten Risikos* bewegen. Mit dieser Rechtsfigur werden Bereiche gekennzeichnet, in denen abstrakt gefährliche Tätigkeiten von der Rechtsordnung aufgrund ihres überragenden sozialen Nutzens im Rahmen rechtlich normierter Voraussetzungen erlaubt werden.<sup>21</sup> Als Beispiel diene der Autoverkehr: Neue Kraftfahrzeuge müssen von einer Behörde für den öffentlichen Straßenverkehr zugelassen werden, und sie dürfen nur von Personen gefahren werden, die eine Fahrerlaubnis besitzen. Sind diese Voraussetzungen erfüllt, gilt das verkehrsgerechte Autofahren als „erlaubt“, solange keine Anzeichen für eine konkrete Gefährdung anderer Verkehrsteilnehmer erkennbar sind. Insofern wird die Verhaltensnorm, die die Verletzung Anderer verbietet, in ihrer Reichweite eingeschränkt. Man kann sagen: Das verkehrsgerechte Fahren wird im Rahmen der genannten Voraussetzungen als ungefährlich fingiert.

Es spricht viel dafür, dass der Gesetzgeber die Zulassung „intelligenter Agenten“ an derartige Erlaubnisverfahren knüpfen wird. Aktuell hat etwa der US-Bundesstaat Kalifornien Vorschriften für den Betrieb von selbstfahrenden Autos vorgeschlagen.<sup>22</sup>

Damit verlagert sich die rechtliche Regelungsfrage ins Verwaltungsrecht, denn von dort stammen die Vorschriften, die die Zulassung der

---

20 S. allgemein zum Eigenverantwortlichkeitsprinzip als Zurechnungsschranke *Kühl*, Strafrecht AT, (Fn. 9) § 4 Rn. 83 ff.

21 Treffend zur Kategorie des erlaubten Risikos als Ausschlussgrund für die objektive Zurechnung *U. Kindhäuser*, Erlaubtes Risiko und Sorgfaltswidrigkeit. Zur Struktur strafrechtlicher Fahrlässigkeitshaftung, GA 1994, S. 197 ff.

22 S. dazu den Beitrag in Spiegel online vom 17.12.2015: <http://www.spiegel.de/auto/aktuell/kalifornien-google-kritisiert-regeln-fuer-selbstfahrende-autos-a-1068260.html> (letzter Aufruf: 4.8.2016).

„Agenten“ in eine feste Form gießen. Es ist eine rechtspolitische und gegebenenfalls rechtsphilosophische Frage, welche „intelligenten Agenten“ man zulassen will und welche Bedingungen an einen erlaubten Betrieb derartiger Geräte geknüpft werden. Das Strafrecht ist insoweit akzessorisch: es übernimmt – nach dem Grundsatz der *Einheit der Rechtsordnung* – die andernorts erteilte Erlaubnis und bewertet sie als Ausschlussgrund für die tatbestandliche Zurechnung von Schäden, die durch Fehlaktionen solcher Geräte herbeigeführt werden.

Sollte es also zu einem durch Lizenzverfahren geregelten Betrieb „intelligenter Agenten“ kommen, sind die Betreiber, sofern sie über eine konkrete Erlaubnis verfügen und sich beim Betrieb an die erlassenen Verwaltungsregeln halten, straffrei.

### C. Vorsatz und Fahrlässigkeit

Nur kurz sei der subjektive Bereich des Tatbestandes angesprochen. Vorsätzliches Handeln des Betreibers spielt nur dort eine Rolle, wo nicht schon auf objektiver Ebene ein Zurechnungsausschluss eingegriffen hat, also vor allem in den Fällen, in denen sich ein Betreiber nicht innerhalb des Rahmens bewegt hat, der vom erlaubten Risiko abgedeckt ist. Handelt ein Betreiber in rechtlich relevanter Weise gefährlich und ist ihm dabei Vorsatz vorzuwerfen, wirft seine Strafbarkeit keine besonderen Fragen auf. Hier geht es um den Missbrauch „intelligenter Agenten“ für schädigende Zwecke.

Relevanter dürften Fälle sein, in denen dem Konstrukteur oder einem Nutzer ein Fahrlässigkeitsvorwurf gemacht werden kann, weil bei der Programmierung etwas übersehen worden ist oder die Bedienung des Geräts aufgrund von Unaufmerksamkeit über das Maß des erlaubten Risikos hinausgeht. Dabei stellen sich ähnliche Fragen der Vorhersehbarkeit wie beim objektiv-atypischen Kausalverlauf. Allerdings muss hier beachtet werden, dass eine generelle Vorhersehbarkeit einer schädigenden Aktion nicht ausreicht, sondern die Vorhersehbarkeit sich darauf beziehen muss, dass der Sorgfaltsfehler gerade zu dem konkreten Schaden bei Dritten geführt hat.<sup>23</sup> Ein fehlender oder nicht beweisbarer Zusammenhang zwischen

---

23 S. zum Maßstab der Vorhersehbarkeit bei Fahrlässigkeitstaten *Kühl*, Strafrecht AT, (Fn. 9) § 17 Rn. 14 ff.

Sorgfaltsmangel und schädigendem Erfolg schließt wiederum die Zurechnung aus.<sup>24</sup>

#### D. Schuld

Interessanter als Fragen von Vorsatz und Fahrlässigkeit ist das Problem der Schuld. Hier lohnt es sich, noch einmal auf den „intelligenten Agenten“ als mögliche Rechtsperson zurückzukommen. Sollten Programmierungen gelingen, die ein selbstständiges Normerkennen, -verstehen und -reflektieren ermöglichen, kann es ja durchaus in der Zukunft Agenten geben, denen man die strafrechtliche Handlungsfähigkeit nicht *per se* absprechen kann. Aber können sich derartige Entitäten schuldig machen?

Strafrechtliche Schuld ist nach heute recht einhelliger Auffassung keine psychische Eigenschaft des Täters – nichts, was dieser mit der Tat auf sich geladen hätte und nun abtragen müsste. Vielmehr gilt Schuld als ein Urteil, das die Rechtsgemeinschaft, vertreten durch die Gerichte – die hier sehr wörtlich „im Namen des Volkes“ sprechen –, gegen den Straftäter fällt. Sein Gegenstand ist ein sozialetischer Tadel für eine Tat, die nicht hingenommen werden kann. Der Grund für diesen Vorwurf steht allerdings strafrechtsphilosophisch in Streit. Während überwiegend – in mehr oder weniger loser Anlehnung an die berühmte Entscheidung BGHSt 2, 194 – vertreten wird, der Täter sei schuldig, wenn und weil er im Moment der Tat auch anders, nämlich rechtstreu hätte handeln können,<sup>25</sup> sieht die funktionalistische Strafrechtstheorie den Schuldvorwurf als Instrument zur Brandmarkung, soweit dies erforderlich sei, um sich von Tat und Täter zu

---

24 Konnte auch durch sorgfaltsgemäßes Verhalten der Taterfolg nicht verhindert werden, verdient der Handelnde keine Strafe – und zwar richtigerweise, weil er den tatbestandlichen Normbefehl nicht verletzt hat. Das gilt – *in dubio pro reo* – auch, wenn nicht beweisbar ist, ob sorgfaltsgemäßes Verhalten den Taterfolg verhindert hätte (so BGHSt 11, 1; BGH NJW 2010, 1087, 1091; *Wessels/Beulke/Satzger*, Strafrecht AT (Fn. 12), Rn. 295-297); Letzteres sehen die Vertreter der Risikoerhöhungstheorie (*Roxin*, Strafrecht AT (Fn. 10), § 11 Rn. 88 ff.; *G. Stratenwerth/L. Kuhlen*, Strafrecht AT, 6. Aufl., München 2011, § 8 Rn. 34-36 m.w.N.) anders.

25 S. zur herrschenden normativen Schuldtheorie *Kühl*, Strafrecht AT (Fn. 9) § 10 Rn. 2 ff.

distanzieren und den Widerspruch gegen den Normverstoß mit den Mitteln des Strafrechts zu artikulieren.<sup>26</sup>

**I.** Die verbreitet, wenn auch in verschiedenen Spielarten vertretene Konzeption, derzufolge dem Täter die Tat dann vorzuwerfen ist, wenn er die Unrechtshandlung kraft seines Willens hätte vermeiden können, nimmt unübersehbare Anleihen bei der traditionellen Philosophie des freien Willens. Soweit diese nicht mehr vertreten wird, tritt an die Stelle des freien Willens dessen Fiktion: Wir behandeln uns so, *als ob* wir autonom handelnde Persönlichkeiten wären. Dafür kann es gute Gründe geben, insbesondere die Unerträglichkeit der deterministischen Gegenposition, die uns zwänge, einander als neuronal programmierte Akteure zu behandeln, deren normwidriges Handeln nur mit bessernden oder sichernden Maßregeln beantwortet werden könnte, weil es ja unvermeidbar gewesen wäre und deshalb nur für die Zukunft einer Wiederholung vorgebeugt werden könnte.<sup>27</sup> Dieses Bild widerspricht einem verbreiteten menschlichen Selbstverständnis und der Idee einer Menschenwürde so sehr, dass es sich politisch-gesellschaftlich nicht durchsetzt – obwohl aus philosophischer wie aus neurowissenschaftlicher Sicht jeder Beweis für einen freien Willen fehlt.

Dieser Ansatz fordert die Frage heraus, ob die Rechtsgemeinschaft dazu neigen könnte, normverstehende Aktionen eines „intelligenten Agenten“ wie eine Handlung eines kompetenten Menschen anzusehen. Würden wir einem derartigen Roboter einen Vorwurf für sein schädigendes Verhalten machen? Das ist eher eine psychologische als eine rechtliche Frage – und es scheint derzeit schwer vorstellbar, dass sie eines Tages bejaht wird.

**II.** Auch der funktionalistische Ansatz führt zu keinem anderen Ergebnis. Er geht aus von der Theorie der positiven Generalprävention, derzufolge das Strafrecht dazu dient, die Fortgeltung der Rechtsnormen zu demonstrieren, wenn die Normgeltung durch eine Missachtung (in Form einer Straftat) ins Wanken geraten ist. Entscheidend ist hiernach das allgemeine Normgeltungsbewusstsein: Eröffnet eine Tat Zweifel an der Norm-

---

26 Dazu G. Jakobs, Strafrecht AT, 2. Aufl., Berlin 1991, 1/15 und 17; 17/18-22; *ders.*, Staatliche Strafe: Bedeutung und Zweck, Paderborn 2004, S. 31 f.

27 S. zu solchen Entwürfen vor allem die neurowissenschaftlichen Schriften, z.B. G. Roth, Willensfreiheit, Verantwortlichkeit und Verhaltensautonomie des Menschen aus Sicht der Hirnforschung, in: D. Dölling (Hrsg.), Jus Humanum. Grundlagen des Rechts und Strafrecht - *Festschrift* für Ernst-Joachim Lampe, Berlin 2003, S. 43-57, und W. Singer, Ein neues Menschenbild? Gespräche über Hirnforschung, Berlin 2003.

geltung, dann ist es geboten, den Täter schuldig zu sprechen, um das Vertrauen Aller in die Norm zu stützen. Nimmt die Bevölkerung ein Geschehen nicht als ernst zu nehmende Missachtung der Norm wahr, bedarf es keines Schuldspruchs.<sup>28</sup> – Die in § 20 StGB genannten massiven geistigen Defizite führen dazu, dass das Verhalten einer derartigen Person normativ nicht ernst genommen wird. Das Strafverfahren gegen den norwegischen Massenmörder Breivik im Jahre 2012 war ein sprechendes Beispiel dafür, dass es zuweilen nur darum geht, ob die Gesellschaft die Tat eines Menschen ernst nehmen will oder sie für einen irren Wahn hält. Letzterenfalls – so die funktionalistische Doktrin – bedarf es keiner Bestrafung, um die regelnde Kraft des Rechts zu demonstrieren.

Auf „intelligente Agenten“ übertragen, stellt sich aus dieser Sicht die Frage, ob deren schädigende Aktionen von der Rechtsgemeinschaft als Angriff auf die Geltung der Rechtsnormen aufgefasst würden. Das dürfte solange zu verneinen sein, wie Roboter, Drohnen und selbstfahrende Autos als „Sachen“ wahrgenommen werden, als Maschinen mit einer raffinierten Programmierung, die schematisch ihren Job erledigen – auch wenn es ihre Programmierung gestattet, unter mehreren Verhaltensoptionen auszuwählen. Die Gesellschaft betrachtet schädigende Fehlaktionen dieser Entitäten nicht als „Normbruch“ – und deshalb bedarf es keiner normstabilisierenden Schuldzuweisung.

Ein Schuldspruch gegen „intelligente Agenten“ ist mithin bei jetzigem Stand der Strafrechtsdogmatik und den gegenwärtigen Auffassungen in der Gesellschaft nicht geboten bzw. nicht einmal begründbar.

### E. Bestrafbarkeit

Ähnlich eindeutig fällt der Befund hinsichtlich der letzten Frage aus: ob man „intelligente Agenten“ – alle Strafbarkeitsvoraussetzungen einmal unterstellt – überhaupt bestrafen kann.<sup>29</sup> Maßstäbe für eine Antwort liefern die Straftheorien. Nach der Vergeltungstheorie dient die Strafe dazu, das durch den Täter verwirklichte Unrecht auszugleichen und dadurch ein Gerechtigkeitsdefizit zu beheben. Würde es aber in der Bevölkerung als geeignete Vergeltung angesehen werden, wenn an dem „intelligenten

---

28 *Jakobs*, Staatliche Strafe (Fn. 26), S. 31 f.

29 Auch hier haben *Gless/Weigend* (Fn. 13), S. 577 f. schon überzeugend Position bezogen.

Agenten“ eine Strafe vollstreckt würde? Selbst seine Zerstörung – also die „Todesstrafe“, wie *Gless* und *Weigend* so treffend formuliert haben<sup>30</sup> – gäbe doch Opfern oder Hinterbliebenen nicht das Gefühl, dass der Gerechtigkeit Genüge getan worden sei. Es mag zivil- oder verwaltungsrechtlich wichtig sein, dass ein „intelligenter Agent“, der einen Schaden verursacht hat, dauerhaft aus dem Verkehr gezogen wird. Aber mit Vergeltung hat das nichts zu tun.

Der heute verbreiteten positiven Generalprävention geht es um die Einübung von Normtreue und die Stabilisierung der Normgeltungserwartung in der Bevölkerung. Diese Ziele könnten nur dann durch eine Strafe gegen den „Agenten“ verfolgt werden, wenn dessen Aktionen als „Normbrüche“ wahrgenommen würden. Selbst wenn aber die Strafrechtsdogmatik einen Handlungsbegriff zugrunde legte, nach dem nicht-menschliche, computer-gelenkte Aktionen als Normbrüche verstanden werden könnten, wäre dieses Verständnis doch – so behaupte ich – auf den Insiderzirkel der Strafrjuristen beschränkt. In der Bevölkerung dürfte auf absehbare Zeit einhellig die Meinung vorherrschen, es handele sich um „Maschinen“ oder „Computer“. Normbrüche oder Normtreue schreibt man dagegen nur menschlichen Akteuren zu.

Spezialpräventive Ansätze wollen mit der Strafe die künftige Normtreue des konkreten Täters sicherstellen. Dieses Ziel könnte tatsächlich erreicht werden, wenn man einen „intelligenten Agenten“ bestraft, soweit er darauf programmiert ist, ein Scheitern seiner „Absicht“, richtig zu agieren, digital wahrzunehmen und selbst zu korrigieren. Aber es fehlte ihm doch das Bewusstsein, bestraft zu werden. Der Druck der strafenden Sanktion ist nur für ein Wesen mit ausgeprägten Emotionen erfahrbar – und gerade in diesem Bereich steckt die Technik der Programmierung noch in den Kinderschuhen oder wird gar nicht intendiert. Hier berührt sich die Frage einer strafenden Rechtsfolge mit der Eingangsfrage der Handlungsfähigkeit: Wer einen Normbefehl nicht als handlungsleitenden Druck wahrnehmen kann, kann auch ein Strafurteil nicht als „Tadel“ für einen Normbruch begreifen und die darauf folgende Sanktion nicht als „Übel“ bewerten.

„Strafe“ ist ein Begriff, der in den Kontext normativer Verständigung gehört. Zwischen „intelligenten Agenten“ und Menschen findet aber keine kommunikative Verständigung in einer normativen Sprache statt. Ansprechbar sind diese Agenten nur über eine binäre digitale Codierung. Sol-

---

30 *Gless/Weigend* (Fn. 13), S. 578.

che Systeme kann man umprogrammieren oder stilllegen, aber nicht bestrafen.

## F. Schluss: Fünf Thesen

Ich fasse den Gedankengang in fünf Thesen zusammen:

1. „Intelligente Agenten“ können nicht im strafrechtlichen Sinne „handeln“, weil ihnen die Fähigkeit fehlt, von sich aus Normen als Befehle wahrzunehmen und also von ihnen unmittelbar angesprochen zu werden. Normkenntnis erhalten sie nur durch die Vermittlung menschlicher Programmierer. Damit können sie keine Adressaten von Rechtsnormen sein.
2. „Intelligente Agenten“ können nicht schuldig gesprochen werden, weil sie im Moment ihrer schädigenden Aktion nicht anders konnten als so zu agieren, wie sie es taten. Sie schuldig zu sprechen, ergibt auch aus funktionalistischer Perspektive keinen Sinn, weil die Rechtsgemeinschaft die Aktionen dieser Agenten nicht als Normbrüche deutet.
3. „Intelligente Agenten“ können nicht bestraft werden, weil durch eine Nachteilszufügung gegenüber ihnen kein anerkannter Strafzweck bedient würde.
4. Die Betreiber „intelligenter Agenten“ können sich für Fehlhandlungen innerhalb ihrer rechtlichen Verantwortungsbereiche strafbar machen. Die Grenze einer solchen Strafbarkeit wird die Legislative wahrscheinlich durch Zulassungsverfahren festlegen, die den Betreibern einen straffreien Raum des *erlaubten Risikos* eröffnen.
5. Wer nicht im Rechtssinne handeln kann, wer nicht schuldig gesprochen werden kann und wer kein tauglicher Adressat einer Verurteilung und Bestrafung ist, ist keine „Person“ im Recht. „Intelligente Agenten“ werden also weiterhin als Maschinen oder Werkzeuge betrachtet werden, nicht als Rechtssubjekte.