



Dr. Janina Sombetzki, Wien

## Verantwortung und Roboterethik – ein kleiner Überblick

*Die Robotik und KI-Forschung stellt uns vor zahlreiche Herausforderungen. Eine davon besteht in der Transformation traditionell dem Menschen vorbehaltenen Kompetenzen wie Autonomie, Handlungsfähigkeit und Verantwortung in ihrer Übertragung auf künstliche Systeme. Um die Rolle und Funktion des Verantwortungskonzepts innerhalb der Mensch-Maschine-Interaktion und Roboterethik in den Blick nehmen zu können, definiere ich zunächst eine etymologische Minimaldefinition der Verantwortung und stelle ihre Relationselemente vor. In einem zweiten Schritt betrachte ich die beiden Arbeitsfelder der Roboterethik, um zuletzt die Rolle der Verantwortung mit Blick auf künstliche Systeme als moralische Handlungssubjekte und als Wertträger\*innen zu veranschaulichen.*

S. 10

- HFR 03/2016 S. 1 -

### <sup>1</sup> A. Einleitung

Neben zahlreichen Herausforderungen, mit denen uns der rasante Fortschritt in Robotik und KI-Forschung gegenwärtig konfrontiert, sehen wir uns vor die Aufgabe gestellt, traditionell nur dem Menschen vorbehaltene Kompetenzen – Vernunft, Autonomie, Urteilskraft, um nur einige zu nennen – in ihrer Übertragung auf künstliche Systeme<sup>1</sup> zu transformieren. In diesem Artikel widme ich mich dem Phänomen der Verantwortung<sup>2</sup>, um am Beispiel dieser Kernkompetenz des Menschen einen Überblick über das Arbeitsfeld der Roboterethik zu geben. Dafür frage ich zunächst danach, was unter Verantwortung traditionell verstanden wird, und schlage eine Minimaldefinition von „Verantwortung“ vor, welche nur die wesentlichen etymologischen Komponenten und damit den „kleinsten gemeinsamen Nenner“ jeder Rede von Verantwortung enthält.<sup>3</sup> In einem zweiten Schritt erläutere ich, was es mit der philosophischen Disziplin der Roboterethik auf sich hat, um zuletzt die Rolle der Verantwortung innerhalb derselben näher in den Blick zu nehmen.

<sup>1</sup> Der Begriff „Roboter“ geht auf das tschechische Wort „robota“ zurück, was so viel bedeutet wie Arbeit, Frondienst und Zwangsarbeit. 1920 wurde der Begriff „robot“ von dem Künstler Josef Čapek geprägt, und sein Bruder Karel Čapek gebrauchte den Begriff „labori“ in seinem Theaterstück *R.U.R. (Rossum's Universal Robots, 1921)* für humanoide Apparaturen, die Serviceleistungen und Arbeit an des Menschen Stelle übernehmen. Catrin Misselhorn versteht einen Roboter als eine besondere Art von elektro-mechanischer Maschine, als spezifische Apparatur, die aus einer Einzeleinheit (einem Prozessor) besteht, aus Sensoren, die Daten oder Informationen über die Welt sammeln und aus einem Effektor oder Aktor, der Signale in zumeist mechanische Abläufe übersetzt. Das Verhalten eines Roboters ist oder wirkt zumindest bis zu einem gewissen Grad autonom (Catrin Misselhorn, *Robots as Moral Agents*, in: Frank Rövekamp/Friederike Bosse (Hrsg.), *Ethics in Science and Society, German and Japanese Views*, 2013, 30-42). Roboter können in einer Weise auf die Umgebung Einfluss nehmen und in sie hinein wirken, in der es Computer nicht in der Lage sind (vgl. hierzu auch Daniel Ichbiah, *Roboter, Geschichte – Technik – Entwicklung*, 2005). Im Folgenden gebrauche ich die Begriffe „Roboter“ und „artifizielles System“ synonym. Nicht bei allen künstlichen Systemen, wohl aber bei den für die hier angestellten Überlegungen relevanten, handelt es sich um Roboter – bereits beim gebräuchlichen Computer fängt der Graubereich an, der von Technikphilosoph\*innen ausgeleuchtet zu werden verdient und in dem bereits Isaac Asimov die roboternahe Sphäre vermutet hat, in der wir neben Computern auch Maschinen, Automaten und weiteren Verwandten und Bekannten der Roboter begegnen (Isaac Asimov, *The Complete Robot, The Definitive Collection of Robot Stories*, 1982, 53).

<sup>2</sup> Zur Verantwortung vgl. Ludger Heidbrink/Claus Langbehn/Janina Sombetzki (Hrsg.), *Handbuch Verantwortung*, 2016 (im Erscheinen).

<sup>3</sup> Vgl. ausführlich Janina Sombetzki, *Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse*, 2014.

- 2 Die Geschichte der Verantwortung ist verhältnismäßig kurz, insofern das Wort „verantwortlich“ erst im 13. Jahrhundert in Frankreich und im 17. Jahrhundert in Deutschland auftaucht. Sie umfasst also im Großen und Ganzen bis zur Gegenwart knapp 800 Jahre, wobei ein reflektierter und systematischer Gebrauch des Begriffs erst seit Beginn des 19. Jahrhunderts zu beobachten ist.<sup>4</sup> Verantwortung lässt sich als Lösungskonzept für solche Kontexte begreifen, in denen die Übersicht über die Akteur\*innen durch große Komplexität und undurchschaubare Hierarchieverhältnisse erschwert wird. Sie dient der Ordnung und Aufhellung intransparenter Verhältnisse und kann als Werkzeug begriffen werden, das die bislang üblichen (wie bspw. die Konzepte der Pflicht- und Schuldzuschreibung) mindestens ergänzt, wenn nicht gar ersetzt. Da Handlungsabläufe mit der Zeit über die Zwischenschaltung von Instanzen immer vermittelter werden, sodass die Einzelnen die Folgen ihres Tuns nicht mehr vollständig zu überblicken in der Lage sind (Industrielle Revolution), greifen direkte Pflicht- und Schuldzuschreibungen immer häufiger zu kurz.<sup>5</sup> Verantwortung wird als Aufgabe verstanden, die die involvierten Handlungssubjekte mit der Lösung von Zuschreibungsproblemen konfrontiert und ihnen aufgibt, den Herausforderungen deprimierender Selbstwahrnehmung aufgrund eines Gefühls des Verlustes von Handlungskontrolle zu begegnen. Mit dem vergangenen Jahrhundert traten neue potenzielle Anwärt\*innen in die Arena der zur Verantwortung angesprochenen Wesen: Roboter.

## S. 11

- HFR 03/2016 S. 2 -

- 3 Die Roboterethik als verhältnismäßig junge Teildisziplin der Philosophie sieht sich immer wieder zwei Vorwürfen ausgesetzt, die ihren Status als Bereichsethik in Frage stellen: Zum einen habe sie keinen spezifischen Gegenstand, da sich Ethik nicht mit Unbelebtem beschäftige. Doch selbst wenn künstliche Systeme zu Recht in den Fokus der ethischen Reflexion geraten würden, ließen sich – so der zweite Einwand – mit ihnen im Blick keine neuen, sondern in anderen ethischen Arenen längst formulierte und ausgetragene Fragen stellen. In aller Knappheit ließe sich dem ersten Einwand damit begegnen, dass, falls sich herausstellen sollte, dass Roboter selbst keine moralischen Handlungssubjekte sein, sie dennoch ihren gerechtfertigten Platz im moralischen Universum einnehmen könnten. Schließlich existieren eine ganze Reihe von (teil-)unbelebten Entitäten, denen wir einen Wert zuzusprechen gewillt sind – Landschaften, Ökosystemen, dem Planeten Erde, aber auch Häusern, Autos und Smartphones. Um was für eine Art von Wert es sich im Falle künstlicher Systeme handelt, bleibt freilich zu diskutieren, doch wo, wenn nicht in der Ethik, sollte eine solche Diskussion geführt werden? Dem zweiten Einwand ist nicht viel entgegenzusetzen, jedoch handelt es sich dabei nicht um eine spezifisch die Roboterethik als philosophische Bereichsethik treffende Kritik. Diese ließe sich letztlich auf alle Bereichsethiken (wie bspw. auch die Klimaethik und die Medizinethik) beziehen, solange wir den Menschen als Ausgangs-, Dreh- und Angelpunkt ethischen Nachdenkens begreifen und dieser in allen Sphären der ethischen Überlegung anzutreffen ist.
- 4 Die Roboterethik stellt durchaus traditionelle philosophische Fragen, gibt einigen Herausforderungen, vor die sich andere Bereichsethiken (wie bspw. die Tierethik) bereits gestellt sahen, ein neues Gewand und wirft den Menschen letztlich auf sich selbst zurück.<sup>6</sup> Welche Kompetenzen erachten wir bspw. als grundlegend, um als Handlungs-

<sup>4</sup> Vgl. Hans Lenk, Matthias Maring, Verantwortung, in: Joachim Ritter (Hrsg.), Historisches Wörterbuch der Philosophie, Band 11, 2007, 566-575; Kurt Bayertz, Eine kurze Geschichte der Herkunft der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 3-71; Richard McKeon, The Development and the Significance of the Concept of Responsibility, in: *Revue Internationale De Philosophie*, 6/39/1957, 3-32.

<sup>5</sup> Vgl. Hans Lenk, Matthias Maring, Wer soll Verantwortung tragen? Probleme der Verantwortungsverteilung in komplexen (soziotechnischen-sozioökonomischen) Systemen, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 1995, 241-286, 242 ff.

<sup>6</sup> Die folgenden aktuellen Sammelbände, Monographien und Artikel bieten einen Einstieg in die Roboterethik: Janina Sombetzki, Roboterethik – ein kritischer Überblick, in: Matthias Maring (Hrsg.), Zur Zukunft der Bereichsethiken, Herausforderungen durch die Ökonomisierung der Welt, 2016 (im Erscheinen); Eric Hilgendorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014; Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012; Michael Anderson/Susan Leigh Anderson (Hrsg.), Machine Ethics, 2011; Patrick Lin/Keith Abney/George Bekey, Robot Ethics, Mapping the Issues of a

subjekte gelten zu können? Was ist darüber hinaus Bedingung für moralische Akteursfähigkeit? Mit welchen moralischen Prinzipien und Werten sollten wir artifizielle Systeme ausrüsten? Auf was für ein moralisches Selbstverständnis lässt es schließen, wenn wir Roboter ‚schlecht‘ behandeln?<sup>7</sup> In welchem Nahbereich des Menschen – Industrie-, Militär-, Medizin-, Altenpflege- und Servicerobotik, um nur einige zu nennen – wollen wir uns auch weiterhin nur oder zumindest in einem signifikanten Ausmaß auf menschliche und nicht auf artifizielle Expertise verlassen? Dass hier auch Fragen der Verantwortungszuschreibung, -teilung und ggf. -delegation eine Rolle spielen, leuchtet intuitiv ein.

S. 12

- HFR 03/2016 S. 3 -

## 5 B. Was ist Verantwortung?

Eine umfangreiche etymologische Untersuchung würde zeigen, dass unser Verständnis von Verantwortung auf drei Säulen fußt: Verantwortung bedeutet *erstens*, dass jemand Rede und Antwort steht, und *zweitens*, dass dies kein rein deskriptives, sondern immer ein zumindest auch normatives Geschehen darstellt. Gerne erklärt man zwar bspw. den Regen verantwortlich für das Nass-Sein der Straße. Doch hier ist von Verantwortung in übertragenem Sinn als Verursachung die Rede. Mit Verantwortung als Rede-und-Antwort-Stehen meinen wir mehr, was uns ein Fall wie bspw. der, dass Person x für den Tod von Person y verantwortlich ist, vor Augen führt. Zwar denken wir hier vielleicht auch daran, dass Person x den Tod von Person y in einem rein kausalen Sinne verursacht hat. Darüber hinaus jedoch ließen sich mit einer solchen Bemerkung auch Erwartungen und Forderung verknüpfen. *Drittens* rekurriert die Rede von Verantwortung auch immer auf bestimmte Kompetenzen, die wir der oder dem Angesprochenen implizit zuschreiben. Wir unterstellen, dass die fragliche Person integer, bedacht und reflektiert das Anliegen der Verantwortung in Angriff nimmt.

- 6 Aus dieser Minimaldefinition ergeben sich fünf Relationselemente der Verantwortung, auf die ich in den folgenden Abschnitten B.I bis B.V näher eingehe. Es bedarf eines *Subjekts* bzw. einer Trägerin oder eines Trägers der Verantwortung. Darüber hinaus ist ein *Objekt* oder Gegenstand zu definieren. *Drittens* gilt es, die *Instanz*, vor der man sich verantwortlich zeigt, auszumachen. *Viertens* tragen wir gegenüber einer *Adressatin* bzw. einem *Adressaten* Verantwortung. Schließlich geben *normative Kriterien* den Maßstab und die Richtlinien dafür ab, in welcher Weise Verantwortung zuzuschreiben ist. Ein Dieb (= Subjekt/Träger) steht bspw. für einen Diebstahl (= Objekt/Gegenstand) vor Gericht (= Instanz). Adressat\*in ist die oder der Bestohlene und die normativen Kriterien stellen das Strafrecht (also strafrechtliche Normen) dar. Innerhalb der Verantwortungsforschung ist es Alfred Schütz, dem zugesprochen wird, als erster explizit auf die Stelligkeit des Verantwortungsbegriffs eingegangen zu sein, indem er diesem zwei Relationselemente – Objekt und Instanz – zuschreibt.<sup>8</sup> Später variiert die Anzahl der Relationselemente von drei über vier<sup>9</sup> bis hin zu sechs<sup>10</sup> und sogar sieben Relationselementen<sup>11</sup>.

Mechanized World, in: Artificial Intelligence, 175/2011, 942-949; Klaus Mainzer, Leben als Maschine, Von der Systembiologie zur Robotik und Künstlichen Intelligenz, 2010.

<sup>7</sup> Ähnlich wie sich bereits Immanuel Kant in § 17 des zweiten Teils seiner *Metaphysik der Sitten* gegen Tierquälerei ausspricht, da diese zu einer Verrohung des Menschen führe, plädiert Kate Darling für Roboterrechte, da es dem Menschen dann eher gelinge „menschlich“ zu bleiben.

<sup>8</sup> Alfred Schütz, Einige Äquivokationen im Begriff der Verantwortlichkeit, in: Alfred Schütz (Hrsg.), *Gesammelte Aufsätze*, Band 2. Studien zur soziologischen Theorie, 1972, 256-258, 256; zu Verantwortungsdiskurs und Verantwortungsforschung siehe Janina Sombetzki, *Historische Beiträge zu einer Minimaldefinition von „Verantwortung“*, in: *Archiv für Begriffsgeschichte*, 56/2014, 197-219, 200-202.

<sup>9</sup> Vgl. bspw. Kurt Bayertz, *Eine kurze Geschichte der Herkunft der Verantwortung*, in: Kurt Bayertz (Hrsg.), *Verantwortung, Prinzip oder Problem?*, 3-71, 16; Peter Fischer, *Politische Ethik. Eine Einführung*, 2006, 105; Hans Freyer, *Verantwortung – Heute*, in: Hans Freyer (Hrsg.), *Gedanken zur Industriegesellschaft*, 1970, 195-221, 197-198; Gertrud Nunner-Winkler, *Verantwortung*, in: Georges Enderle (Hrsg.), *Lexikon der Wirtschaftsethik*, 1993, 1185-1192, 1187; Bernhard Schlink, *Die Zukunft der Verantwortung*, in: *Merkur. Deutsche Zeitschrift für europäisches Denken*, 738/64/1010, 1047-1058, 1048; Bernhard Waldenfels, *Antwort und Verantwortung*, in: *Friedrich Jahresheft*, 10/1992, 130-132, 130.

## S. 13

- HFR 03/2016 S. 4 -

- 7 Die Bedingungen für die Möglichkeit einer Zuschreibung von Verantwortung lassen sich in drei Kompetenzgruppen differenzieren: Kommunikationsfähigkeit, Handlungsfähigkeit und Urteilkraft. Auf diese komme ich in Abschnitt B.VI zu sprechen. Alle Kompetenzen als Voraussetzung für die etwaige Zuschreibung von Verantwortung und mit ihr die Verantwortung selbst sind graduell bestimmbar; man kann von mehr oder weniger Kommunikations- und Handlungsfähigkeit sprechen und abhängig davon von mehr oder weniger Verantwortung.<sup>12</sup>

8 **B.I. Das Subjekt der Verantwortung**<sup>13</sup>

Das Subjekt der Verantwortung, auch Träger\*in der Verantwortung genannt, ist die bzw. derjenige, die oder der Rede und Antwort stehen kann. Welche Fähigkeiten und Kompetenzen dieser Jemand mitbringen muss, um als potenzielle\*r Verantwortungsträger\*in in Betracht zu kommen, wird in B.VI ausgeführt. Abhängig von den Bedingungen, die für eine etwaige Zuschreibung von Verantwortung erfüllt sein müssen, lässt sich der Frage nachgehen, ob nur „gesunde und erwachsene“ Menschen oder auch Kinder für Anforderungen zu einer Übernahme von Verantwortung ansprechbar sind. Sind es gar auch (einige) Tiere und auch Pflanzen, vielleicht sogar unbelebte Dinge (wie bspw. einige künstliche Systeme)? Denn schließlich sprechen wir doch auch vom Regen, der verantwortlich dafür ist, dass die Straße nass wird. Innerhalb des Verantwortungsdiskurses ist man sich jedoch darüber einig, dass Verantwortung traditionell ein individualistisches Prinzip darstellt und in deutlicher Nähe zu (wenn auch nicht unbedingt gleichbedeutend mit) dem Konzept der Personalität gesehen wird.<sup>14</sup> Dieser Kernbestand des klassischen Verantwortungskonzepts wird erst in der Gegenwart in Frage gestellt. Hieraus erhellt, warum jeder rein deskriptive oder kausale Gebrauch des Verantwortungsbegriffs, d. h. ein solcher, der einen normativen Gebrauch gar nicht erlaubt, wie in dem obigen Regen-Beispiel, nur metaphorisch gemeint sein kann.<sup>15</sup> Der Regen ist nicht in der Lage, Rede und Antwort zu stehen und Pflanzen sind es wohl ebenso wenig. Ob (einige) Tiere und Kinder als Anwärter\*innen auf die Träger\*innenposition der Verantwortung in Frage kommen, darüber kann hier kein Urteil gefällt werden,<sup>16</sup> wohl aber komme ich auf künstliche Systeme als potenzielle Verantwortungssubjekte in Abschnitt D.I zu sprechen.

- 9 Ein genaues Verständnis von Verantwortung in einer fraglichen Situation verlangt ein Urteil darüber, ob es sich bei dem fraglichen Verantwortungssubjekt um ein Individuum

<sup>10</sup> Hans Lenk, Matthias Maring, Verantwortung, in: Joachim Ritter (Hrsg.), Historisches Wörterbuch der Philosophie, Band 11, 2007, S. 570.

<sup>11</sup> Günter Ropohl, Das Risiko im Prinzip Verantwortung, Ethik und Sozialwissenschaften. Streitforum für Erziehungskultur, 1/5/1994, 109-120, 111-114.

<sup>12</sup> Zur Verantwortung ausführlich Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014 – zur Minimaldefinition 33 ff., zu den Relationselementen 63 ff. und zu den Bedingungen für die Möglichkeit zur Verantwortungsübernahme 43 ff.

<sup>13</sup> Zu den einzelnen Relationselementen, die nun vorgestellt werden, gibt es in Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, jeweils ein eigenes Unterkapitel.

<sup>14</sup> Vgl. Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 66-69.

<sup>15</sup> Vgl. Hans Lenk, Matthias Maring, Wer soll Verantwortung tragen? Probleme der Verantwortungsverteilung in komplexen (soziotechnischen-sozioökonomischen) Systemen, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 1995, 242; Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 37-39; Micha H. Werner, Verantwortung, in: Marcus Düwell/Christoph Hübenal/Micha H. Werner (Hrsg.), Handbuch Ethik, 2006, 541-548, 542.

<sup>16</sup> Vgl. zum Lernen von Verantwortung und zur Gradualität des Verantwortungskonzepts Peter Fauser, Kann die Schule zur Verantwortung erziehen?, in: Friedrich Jahresheft, 10/1992, 7-9; John Martin Fischer/Mark Ravizza, Responsibility and Control, A Theory of Moral Responsibility, 1998, 208-210, 238; Julian Nida-Rümelin, Politische Verantwortung, in: Ludger Heidbrink/Alfred Hirsch (Hrsg.), Staat ohne Verantwortung? Zum Wandel der Aufgaben von Staat und Politik, 2007, 55-85, 63; Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 69-71, 143-150.

oder um ein Kollektiv handelt.<sup>17</sup> Im Falle eines Individuums ist zwischen alleiniger, persönlicher, Eigen- und Selbstverantwortung zu differenzieren. Sprechen wir einem Kollektiv Verantwortung zu, geht es um Mechanismen der Verantwortungsdelegation oder -teilung, die einem Mitglied des Kollektivs entweder eine partielle, volle oder gar keine Verantwortung für den fraglichen Verantwortungsgegenstand zuschreiben. In Abschnitt D.II führe ich den Terminus des Verantwortungsnetzwerks ein<sup>18</sup>, um die unterschiedlichen Funktionen in den Blick zu bekommen, die die involvierten Parteien insbesondere im Falle von Mensch-Maschine-Interaktionen haben

S. 14

- HFR 03/2016 S. 5 -

## 10 B.II. Das Objekt der Verantwortung

Man steht nicht einfach so Rede und Antwort, sondern immer für etwas oder jemanden. Die Rede von Verantwortung verlangt neben einem Subjekt auch ein Objekt bzw. einen Gegenstand, wofür Verantwortung übernommen wird. Auf den ersten Blick lassen sich zahlreiche Arten von Verantwortungsobjekten in grob vier Kategorien unterscheiden: Wir sind für *Lebewesen* wie bspw. Kinder verantwortlich, für *Gegenstände und Dinge* wie z. B. die Brille, die wir für unseren Vater vom Optiker abholen sollen. Weiterhin stehen wir für *Ereignisse* Rede und Antwort wie etwa für einen Raubüberfall. Und schließlich sind es *Handlungen und Handlungsfolgen*, für die man uns Verantwortung zuschreibt. Bei genauerer Betrachtung allerdings wird deutlich, dass jede Verantwortlichkeit in eine solche für Handlungen und Handlungsfolgen übersetzt werden kann. Die Verantwortung der Eltern für ihre Kinder ist eine Verantwortung für das Wohlergehen der Kinder, was durch die Handlungen xy gefördert und erreicht wird. Auch der Raubüberfall als Ereignis stellt ein Arrangement an Handlungen dar, die zusammengenommen die Verantwortung für den Raubüberfall ausmachen. In welcher Weise Verantwortungssubjekt und -objekt miteinander verknüpft sind, wird über die Rollen deutlich, die unseren Alltag gewissermaßen ordnen. Rollen geben uns die Objekte, für die wir verantwortlich sind, vor – unbenommen, dass Rollen unterschiedlich scharfe Grenzen aufweisen und verschieden klar definiert sein können.<sup>19</sup>

11 Verantwortungsobjekte sind immer vergangene oder zukünftige – sind Teil retrospektiver oder prospektiver Verantwortungskonstellationen. Die Prospektivität oder Retrospektivität wird auch Ex-post- und Ex-ante-Verantwortung genannt und im Englischen forward looking und backward looking responsibility.<sup>20</sup> Um den Zeitpunkt einer Verantwortungsübernahme ausmachen zu können, muss das Objekt, für das jemand Rede und Antwort stehen soll, bereits bekannt bzw. zumindest oberflächlich definiert worden sein. In den meisten Situationen wird uns das ohne größere Schwierigkeiten gelingen; so stellt die Verantwortung des Angeklagten für den Diebstahl offensichtlich eine retrospektive Verantwortlichkeit dar, denn um als Angeklagter vor Gericht zitiert werden zu können, muss der Gegenstand, um den es geht (in diesem Fall der Diebstahl), bereits in der Vergangenheit liegen. Doch in anderen Fällen erscheint die zeitliche Einordnung des Verantwortungsgegenstands in Vergangenheit oder Zukunft ohne ausführende Erklärungen weniger eindeutig auf der Hand zu liegen, wie in Sätzen wie „Die

<sup>17</sup> Vgl. Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 71-94.

<sup>18</sup> Ursprünglich stammt der Terminus von Christian Neuhäuser, der ihn allerdings nicht genauer definiert hat (Christian Neuhäuser, Roboter und moralische Verantwortung, in: Eric Hilgendorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, 269-286).

<sup>19</sup> Vgl. R. A. Duff, Responsibility, in: Edward Craig (Hrsg.), Routledge Encyclopedia of Philosophy, 1998, 290-294, 292; Johannes Schwartländer, Verantwortung, in: Hermann Krings/Hans Michael Baumgartner/Christoph Wild (Hrsg.), Handbuch philosophischer Grundbegriffe, Band 6, Transzendenz – Zweck, 1974, 1577-1588, 1584; Micha H. Werner, Verantwortung, in: Marcus Düwell/Christoph Hübenenthal/Micha H. Werner (Hrsg.), Handbuch Ethik, 2006, 543.

<sup>20</sup> Vgl. Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 103-104; Dieter Birnbacher, Grenzen der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 1995, 143-183, 145-146; Ibo van de Poel, The Relation Between Forward-Looking and Backward-Looking Responsibility, in: Nicola A. Vincent/Ibo van de Poel/Jeroen van den Hoven (Hrsg.), Moral Responsibility, Beyond Free Will and Determinism, 2001, 37-52.

Menschen sind für den Klimawandel verantwortlich“. In diesem Fall kann, ohne dass dazu Näheres gesagt wird, mit „Klimawandel“ bspw. auf den jetzigen Stand des Klimas rekuriert werden, was die fragliche Verantwortung zu einer retrospektiven erklären würde. Insofern sind die Menschen für alle Handlungen verantwortlich zu machen, die zu dem Klimawandel bzw. dem jetzigen Stand des Klimas geführt haben. Allerdings ließe sich die Verantwortung der Menschen für den Klimawandel auch prospektiv ausdeuten, indem man darunter eine Beeinflussung des Klimas hin zu einem anderen, wünschenswerteren Zustand versteht. In jedem Fall müssen weitere Angaben zur Prospektivität und Retrospektivität für eine Konkretisierung der in dem obigen Satz genannten Verantwortlichkeit gemacht werden, damit man weiß, wann in diesem Fall ein Rede- und-Antwort-Stehen gefordert ist.

S. 15

- HFR 03/2016 S. 6 -

### 12 **B.III. Die Instanz der Verantwortung**

Die Instanz stellt neben Subjekt und Objekt die bekannteste und am wenigsten hinterfragte Relation der Verantwortung dar. Das liegt vielleicht auch daran, dass der Verantwortungsbegriff zuerst im Rechtsbereich auftrat,<sup>21</sup> wo ein\*e Angeklagte\*r für gewöhnlich vor der Instanz des Gerichts Rede und Antwort zu stehen hat. Solange die Fähigkeit, Verantwortung tragen zu können, an Personalität geknüpft ist (selbst dann, wenn Personalität ebenso wie Verantwortung als graduelles Konzept verstanden wird), kommen im eigentlichen Sinne weder unbelebte Gegenstände, Pflanzen, Tiere noch Kleinkinder als potenzielle Instanzen in Betracht.<sup>22</sup>

13 Innerhalb des Verantwortungsdiskurses wird zwischen externen und internen Instanzen differenziert, wobei das Gericht eine externe und das persönliche Gewissen wohl die berühmteste interne Instanz darstellt. Verantwortungsinstanzen lassen sich entweder absolut verstehen, und sind damit unhintergehbare, unhinterfragbare, „letzte“<sup>23</sup> Instanzen, oder ihnen wird nur der Status vorletzter Instanzen, die in ihrer Kompetenz und Reichweite begrenzt sind, zugesprochen. Tradierte Anwarter\*innen auf die Rolle der absoluten Instanz, die innerhalb der Verantwortungsforschung höchstumstritten ist,<sup>24</sup> sind bspw. Gott<sup>25</sup> und die Geschichte<sup>26</sup>.

14 Die Verantwortungsinstanz, die in der Regel eine externe und vorletzte Autorität darstellt, ist entweder eine öffentliche oder private. Öffentlich ist eine Instanz intersubjektiv geteilt, was ggf. mit einer gewissen Sanktionskraft einhergeht. Öffentlichkeit meint

---

<sup>21</sup> Zu Verantwortung als rechtlicher Begriff vgl. Jan Henrik Klement, Verantwortung im Recht, in: Ludger Heidbrink/Claus Langbehn/Janina Sombetzki (Hrsg.), Handbuch Verantwortung, 2016 (im Erscheinen); Klaus Günther, Voluntary Action and Criminal Responsibility, in: Sabine Maasen (Hrsg.), Voluntary Action, Brains, Minds, and Sociality, 2007, 263-280; Jan Henrik Klement, Verantwortung, Funktion und Legitimation eines Begriffs im Öffentlichen Recht, 2006; Lukas Gschwend, Verantwortung und Strafrecht, in: Heinrich Schmidinger/Clemens Sedmak (Hrsg.), Der Mensch – ein freies Wesen? Autonomie – Personalität – Verantwortung, 2005, 289-306; Ulfried Neumann/Lorenz Schulz (Hrsg.), Verantwortung in Recht und Moral, Referate der Tagung der Deutschen Sektion der Internationalen Vereinigung für Rechts- und Sozialphilosophie vom 2. bis zum 3. Oktober 1998 in Frankfurt am Main, 2000.

<sup>22</sup> Vgl. Wilhelm Weischedel, Das Wesen der Verantwortung, Ein Versuch, 1972, 26.

<sup>23</sup> Christian Müller, Verantwortungsethik, in: Annemarie Pieper (Hrsg.), Geschichte der neueren Ethik, 1992, 103-131, 114.

<sup>24</sup> Vgl. Matthias Kaufmann, Die Grenzen der Zurechnung, in: Matthias Kaufmann (Hrsg.), Zurechnung als Operationalisierung von Verantwortung, 2004, 283-293, 288-289; Christian Müller, Verantwortungsethik, in: Annemarie Pieper (Hrsg.), Geschichte der neueren Ethik, 1992, 103-131, 114; Micha H. Werner, Verantwortung, in: Marcus Düwell/Christoph Hübenthal/Micha H. Werner (Hrsg.), Handbuch Ethik, 2006, 541-548, 546-547.

<sup>25</sup> Vgl. Jann Holl, Historische und systematische Untersuchungen zum Bedingungsverhältnis von Freiheit und Verantwortlichkeit, 1980, 46, 83.

<sup>26</sup> Vgl. Horst Dreier, Verantwortung im demokratischen Verfassungsstaat, in: Ulfried Neumann/Lorenz Schulz (Hrsg.), Verantwortung in Recht und Moral. Referate der Tagung der Deutschen Sektion der Internationalen Vereinigung für Rechts- und Sozialphilosophie vom 2. bis zum 3. Oktober 1998 in Frankfurt am Main, 2000, 9-38, 18; Georg Picht, Der Begriff der Verantwortung, in: Georg Picht (Hrsg.), Wahrheit, Vernunft, Verantwortung. Philosophische Studien, 1969, 318-342, 328-329; Bernhard Schlink, Die Zukunft der Verantwortung, in: Merkur. Deutsche Zeitschrift für europäisches Denken, 738/64/1010, 1047-1058, 1047.

damit nicht mediale Präsenz, sondern die Anerkennung durch die Beteiligten, woraus ein etwaiger psychosozialer Druck folgt, dem sich die verantwortliche Person ausgesetzt sieht. Es handelt sich nicht um ein inhaltliches Gutheißen z. B. eines Gesetzes oder der Entscheidungen eines Richters. Anerkennung der Instanz rekuriert auf die Akzeptanz des legitimen Zustandekommens bspw. eines Gesetzes, woraus ggf. resultiert, dass die Angesprochenen dem Geforderten nachzukommen haben und tatsächlich nachkommen. In diesem Sinne genießt der Bestseller einer bekannten Autorin breite mediale Präsenz, ohne dass das in ihrem Buch zum Ausdruck Gebrachte notwendig anerkannt wird. Privatheit hingegen bezieht sich nicht auf eine etwaige Vagheit oder Ambiguität der Instanz. Konventionen sind ein Beispiel für öffentliche, jedoch nicht notwendig eindeutig definierte Maßstäbe. Private Normen werden nicht intersubjektiv geteilt und haben keine bzw. wenig Sanktionskraft.

**S. 16**

- HFR 03/2016 S. 7 -

- 15 Ob die Verantwortungsinstanz privat oder öffentlich ist, hängt je nach Kontext bspw. von den spezifischen Kriterien innerhalb eines Verantwortungsbereichs ab. In den meisten Fällen wird es zwar so sein, dass bspw. eine Richterin eine öffentliche Instanz darstellt, deren Rolle an den Bereich der strafrechtlichen Verantwortung geknüpft ist. Kompetenzen wie Unparteilichkeit und Objektivität kommen ihr hingegen nicht nur als Richterin in einem *Amt* (bspw. am BGH) zu, sondern charakterisieren ihre *Rolle* als Richterin (die sowohl eine berufliche wie auch eine moralische sein kann). Deshalb wird eine Person mit dem Beruf Richter\*in auch in anderen Kontexten als Verantwortungsinstanz gerne zu Rate gezogen, wenn auch dann nicht in ihrer formalen Funktion. Stellen wir uns bspw. den Streit einiger Freunde vor, die für die Lösung ihres Konflikts jemanden suchen, der unparteiisch entscheidet und über ein klares Urteilsvermögen verfügt. Eine Richterin des BGH als eigentlich strafrechtliche und öffentliche Verantwortungsinstanz kann auch in moralischen Fragen als private Autorität zu Rate gezogen werden, wenn auch nicht in ihrer formalen oder beruflichen Funktion, wohl aber dank der die Rolle Richter\*in definierenden (oben genannten) Kompetenzen. Das bedeutet, dass Richter\*innen nicht als generell entweder formelle oder informelle Instanzen zu begreifen sind, sondern in einigen Kontexten formelle Instanzen aufgrund der Normen, die den Handlungsbereich in diesem Fall definieren, in anderen informelle darstellen.

**16 B.IV. Die Adressatin bzw. der Adressat der Verantwortung**

Innerhalb des Verantwortungsdiskurses stellt die Adressatin bzw. der Adressat der Verantwortung ein häufig umstrittenes Relationselement dar und die in meinen Augen in der Tat unterschätzteste Relation der Verantwortung. Oftmals wird zwischen Instanz und Adressat\*in nicht hinreichend differenziert, da letztere\*r entweder für überflüssig gilt<sup>27</sup> oder beide Relationen synonym gebraucht werden<sup>28</sup>. Selten wird in der Adressatin ein eigenes und genuines Relationselement gesehen.<sup>29</sup> Die Adressatin ist das Gegenüber des Verantwortungssubjekts. Sie ist die Betroffene der fraglichen Verantwortlichkeit und definiert den Grund für das Vorhandensein derselben. Diesen Sachverhalt drückt Albs knapp und äußerst treffend mit der Phrase aus, dass „[w]o kein Kläger, da kein Richter“<sup>30</sup> zu finden sei. Im Falle eines Diebstahls stellt die oder der Bestohlene den Grund dafür dar, dass in diesem Kontext überhaupt von Verantwortung

<sup>27</sup> Vgl. Kurt Bayertz, Eine kurze Geschichte der Herkunft der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 3-71, 16, Fußnote; Christian Müller, Verantwortungsethik, in: Annemarie Pieper (Hrsg.), Geschichte der neueren Ethik, 1992, 103-131, 127.

<sup>28</sup> Vgl. Ludger Honnefelder, Was soll ich tun, wer will ich sein? Vernunft und Verantwortung, Gewissen und Schuld, 2007, 43; Alfred Schütz, Einige Äquivalenzen im Begriff der Verantwortlichkeit, in: Alfred Schütz (Hrsg.), Gesammelte Aufsätze, Band 2. Studien zur soziologischen Theorie, 1972, 256-258, 256; Wilhelm Weischedel, Das Wesen der Verantwortung, Ein Versuch, 1972, 26.

<sup>29</sup> Vgl. bspw. Birgit Albs, Verantwortung übernehmen für Handlungen und deren Folgen, 1997, 27; Hans Lenk, Matthias Maring, Verantwortung, in: Joachim Ritter (Hrsg.), Historisches Wörterbuch der Philosophie, Band 11, 2007, 566-575, 570. Auf die Gründe dafür, dass die Adressatin innerhalb von Verantwortungskonstellationen häufig in ihrer Eigenständigkeit als Relationselement schlecht wegkommt, kann ich hier nicht eingehen.

<sup>30</sup> Birgit Albs, Verantwortung übernehmen für Handlungen und deren Folgen, 1997, 27.

gesprachen wird, in einem weiteren Sinne vielleicht noch die Bürger\*innen Deutschlands und die verletzte Norm selbst. Hieran zeigt sich überdies, dass die Adressatin bzw. der Adressat – im Gegensatz zur Verantwortungsinstanz – nicht selbst potenzielle\*r Verantwortungsträger\*in sein können muss, nicht selbst die Kompetenzen für die Möglichkeit einer Verantwortungszuschreibung mitzubringen hat. Als Grund für das Bestehen einer Verantwortlichkeit können sowohl unbelebte wie belebte Wesen dienen und auch Abstrakta wie Normen und Werte – eben all das, welchem wir einen Wert zuschreiben. Die Adressatin bzw. der Adressat als Warum der Verantwortung bzw. als Gegenüber der oder des Verantwortlichen lässt sich aus der Minimaldefinition von Verantwortung, nämlich aus der Fähigkeit des Rede-und-Antwort-Stehens als ein normatives Konzept ableiten. Denn rein deskriptive Zuschreibungssachverhalte bestehen faktisch, normative Verhältnisse hingegen bedürfen einer Begründung, die durch die Adressatin bzw. den Adressaten gewährleistet wird.

S. 17

- HFR 03/2016 S. 8 -

- 17 Auch die Adressatin ist– ebenso wie die Instanz – privat oder öffentlich. Stellen wir uns eine Politikerin vor, die in einer Rede verlauten lässt: „Unsere Kinder, ja, auch meine eigenen Kinder, leiden unter den mangelhaften Ausbildungsmöglichkeiten, und dafür tragen wir alle, auch ich, die Verantwortung“. Hier sind die Kinder nicht nur moralische private Adressat\*innen innerhalb der Verantwortung der Eltern für sie, sondern auch politische öffentliche Adressat\*innen. Über die Kinder der Politikerin wird dabei in derselben Weise geurteilt wie über die Kinder der Zuhörer\*innen, ihnen kommt derselbe Status zu.

## 18 **B.V. Die normativen Kriterien der Verantwortung**

Als letztes Relationselement, das aus der oben gegebenen Minimaldefinition abgeleitet werden kann, nehme ich nun die normativen Kriterien in den Blick. Sie stellen das Inwiefern, den Maßstab und normativen Bezugsrahmen dar, nach dem in einem gegebenen Kontext darüber geurteilt wird, ob die fragliche Person verantwortlich gehandelt hat. Die Verantwortung bringt ihre Normen nicht mit, sondern muss mit ihnen gewissermaßen erst „angereichert“ werden; sie ist „parasitär“<sup>31</sup> gegenüber ihrem eigenen normativen Fundament. Normative Kriterien nehmen unterschiedliche Gestalt an – sie begründen eine Verantwortlichkeit bspw. als Werte, Prinzipien, Gebote, Maximen, Gesetze, Regeln, Befehle, Aufgaben oder Anweisungen. Doch es sind wohl die Pflichten, die traditionell eine für Verantwortungskonstellationen wesentliche Rolle spielen<sup>32</sup>; wohl nicht zuletzt auch aus dem Grund, dass sie neben der Schuld eines der berühmten Vorgängerkonzepte der Verantwortung darstellen.<sup>33</sup> Über die Einspeisung von Pflichten, die Verantwortlichkeit ausformulieren und begründen, lässt sich der Verantwortungsbegriff in das Vokabular des deontologischen Denkens einordnen.

- 19 Hier begegnen wir notwendig dem „Grundlegungsproblem“ oder „Letztbegründungsdefizit“ der Verantwortung, insofern man in jeder Rede von Verantwortung „gezwungen [ist], eine Metaphysik nachzureichen, die allererst zeigen kann, weshalb die Verpflich-

<sup>31</sup> Kurt Bayertz, Eine kurze Geschichte der Herkunft der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 3-71, 65.

<sup>32</sup> Vgl. Pascal Delhom, Staat und Politik aus Verantwortung, in: Ludger Heidbrink/Alfred Hirsch (Hrsg.), Staat ohne Verantwortung? Zum Wandel der Aufgaben von Staat und Politik, 2007, 193-214, 196; Horst Dreier, Verantwortung im demokratischen Verfassungsstaat, in: Ulfried Neumann/Lorenz Schulz (Hrsg.), Verantwortung in Recht und Moral, Referate der Tagung der Deutschen Sektion der Internationalen Vereinigung für Rechts- und Sozialphilosophie vom 2. bis zum 3. Oktober 1998 in Frankfurt am Main, 2000, 9-38, 11; Matthias Kaufmann, Die Grenzen der Zurechnung, in: Matthias Kaufmann (Hrsg.), Zurechnung als Operationalisierung von Verantwortung, 2004, 283-293, 54; Hans Ryffel, Verantwortung als sittliches Phänomen, Ein Grundzug der Moderne, in: Der Staat. Zeitschrift für Staatslehre, öffentliches Recht und Verfassungsgeschichte, 3/6/1967, 275-292, 286.

<sup>33</sup> Vgl. Janina Sombetzki, Historische Beiträge zu einer Minimaldefinition von „Verantwortung“, in: Archiv für Begriffsgeschichte, 56/2014, 197-219, 202-206.



tung der Verantwortung gerade vor dieser Instanz [...] besteht".<sup>34</sup> Jede Verantwortung muss sich wieder auf Normen berufen, die einer Begründung bedürfen. In der Verantwortung lässt sich letztlich ein Begründungszirkel feststellen, der bereits in der ersten vollen Verantwortlichkeit der autonomen Person – in ihrer Selbstverantwortung – auszumachen ist, da in dieser die Normen zur Beurteilung des eigenen Lebens und Daseins aus dem Individuum selbst stammen.<sup>35</sup>

S. 18

- HFR 03/2016 S. 9 -

20 Die normativen Kriterien definieren Verantwortungsbereiche, in denen jemand – begrenzt durch Normen – Rede und Antwort steht. Je nach Kontext, abhängig von dem Set an Kriterien, die der fraglichen Verantwortlichkeit zugrunde liegen, handelt es sich dabei um bspw. den strafrechtlichen, politischen, moralischen oder wirtschaftlichen Raum und demzufolge um eine strafrechtliche, politische, moralische oder wirtschaftliche Verantwortung. Der Verantwortungsbereich ordnet eine für sich genommen unbestimmte Norm als politische oder moralische Norm ein und präzisiert auf diese Weise den Handlungsrahmen der fraglichen Verantwortlichkeit. So kann ein Satz wie „Man darf niemanden töten“ je nach Äußerungskontext bspw. als moralisches Prinzip, als religiöses Gebot oder gar als Gesetz interpretiert werden.<sup>36</sup>

21 Was unter Privatheit und Öffentlichkeit eines Relationselements zu verstehen ist, habe ich bereits in B.III und B.IV ausgeführt. Mit Rekurs auf die normativen Kriterien lassen sich private und öffentliche Normen eher graduell differenzieren; man könnte bspw. moralische Normen als tendenziell private, politische Normen als eher öffentliche und strafrechtliche Normen als öffentliche Normen mit einer besonderen Sanktionskraft begreifen. Nicht alle moralischen Normen sind gleichermaßen privat und nicht alle strafrechtlichen Normen gleichermaßen öffentlich. Die Normen des Strafgesetzbuchs stellen zunächst zwar eindeutig öffentliche Normen dar, doch sind bspw. zahlreiche strafrechtliche Normen völkerrechtlicher Verträge aufgrund ihrer mangelnden Sanktionskraft in einem sehr viel schwächeren Sinne öffentlich. Auch der Beispielsatz „Man darf niemanden töten“ kann dies veranschaulichen, denn als Erziehungsprinzip, gegenüber Kindern geäußert, hat er einen sehr viel privateren Status, als wenn er bspw. als theologisches Lehrprinzip während der Ostermesse vorgetragen wird.

## 22 **B.VI. Bedingungen für die Möglichkeit zur Verantwortungsübernahme**

Zu Beginn von Abschnitt B habe ich bereits vorweggenommen, dass sich die Bedingungen, die erfüllt sein müssen, damit jemandem Verantwortung zugeschrieben werden kann, in drei Kompetenzgruppen bündeln lassen: Ein potenzielles Verantwortungssubjekt (und im Übrigen ebenso die Instanz, wie in Abschnitt B.III festgestellt) bedarf der Kommunikationsfähigkeit, der Handlungsfähigkeit sowie Urteilskraft.<sup>37</sup> Alle Kompetenzen und mit ihr die Verantwortung selbst sind graduell bestimmbar – im Sinne von mehr oder weniger Verantwortung.

23 Insofern jede Verantwortlichkeit als ein Akt der Kommunikation aufzufassen ist, bedarf die Fähigkeit, Rede und Antwort stehen zu können, der Möglichkeit zu kommunizieren. Innerhalb des Verantwortungsdiskurses spielt in der Tat Sprachlichkeit als Voraus-

<sup>34</sup> Alle Zitate in Christian Müller, Verantwortungsethik, in: Annemarie Pieper (Hrsg.), Geschichte der neueren Ethik, 1992, 103-131, 106.

<sup>35</sup> Vgl. Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 161-171.

<sup>36</sup> Für eine detailliertere Auseinandersetzung mit dem Verantwortungsbereich vgl. Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 125-128.

<sup>37</sup> Ausführlich in Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, Kapitel 2.

setzung für Verantwortung eine große Rolle.<sup>38</sup> Auch Weischedel bemerkt, dass „Verantwortung [...] nur da anzutreffen sein [wird], wo Sprechen möglich ist“<sup>39</sup>. Da anzunehmen ist, dass auch eine taubstumme Person in der Lage ist, Rede und Antwort zu stehen, bietet sich Kommunikations- sehr viel besser als Sprachfähigkeit als Verantwortungsbedingung an.

S. 19

- HFR 03/2016 S. 10 -

- 24 Handlungsfähigkeit – vielleicht noch besser mit Autonomie benannt – umfasst vier weitere Kompetenzen, nämlich Folgenbewusstsein (Wissen), Kontextwahrnehmung als Geschichtlichkeit, die Wahrnehmung des Individuums als Einheit (Personalität) sowie Einflussmöglichkeit. Innerhalb des Verantwortungsdiskurses ist man sich hinlänglich darüber einig, dass überdies Freiheit eine der notwendigen Verantwortungsbedingungen ist, ja, dass das „gesamte Rechtssystem [...] auf der Voraussetzung [basiert], daß es so etwas wie Willensfreiheit und persönliche Verantwortung gibt“<sup>40</sup>. Willensfreiheit scheint intuitiv für die Erklärung von Phänomenen wie Handlung, Schuld und Rechtfertigung zu überzeugen.<sup>41</sup> Nur eine Minderheit unter den Verantwortungsforscher\*innen widerspricht der Überzeugung, Freiheit stelle eine Bedingung für die Möglichkeit zur Verantwortungsübernahme dar.<sup>42</sup> Ich verstehe Freiheit nicht direkt als Bedingung für Verantwortung, sondern – wenn überhaupt – als Voraussetzung für Handlungsfähigkeit bzw. sehe wiederum in Handlungsfähigkeit umfassende Autonomie, die mit Wissen, Kontextwahrnehmung, Personalität und Einflussmöglichkeit über die für eine etwaige Verantwortungszuschreibung wesentlichen Momente der Selbstbestimmung verfügt.
- 25 Verantwortung als Rede-und-Antwort-Stehen stellt eine gesteigerte oder zweckgebundene Form des bloßen Antwortens dar, weshalb Normen als Maßstab für die Entscheidung, ob Verantwortungsübernahme eingefordert werden kann, benötigt werden. Verantwortung beruht daher auch auf Urteilskraft, die der oder dem Verantwortlichen die Gründe und Kriterien an die Hand reicht, mit deren Hilfe sie bzw. er sein Handeln als Rede-und-Antwort-Stehen beurteilen kann.<sup>43</sup> Verantwortung als normatives Konzept, für das ein reines Kausalverhältnis nicht hinreichend ist, meint, dass sie „stets mit einem Werturteil verknüpft ist“<sup>44</sup>, wofür die Verantwortliche „auf ein jeweils bestimmtes System von Normen und Werten“<sup>45</sup> zurückgreifen kann. Um ein solches Normensystem auszubilden, muss geurteilt werden können. Urteilskraft umfasst dabei sowohl weitere kognitive Kompetenzen wie Reflexion und Rationalität als auch zwischenmenschliche Institutionen wie Versprechen, Vertrauen und Verlässlichkeit. Nida-

<sup>38</sup> Vgl. Rainer Piepmeier, Zum philosophischen Begriff der Verantwortung, in: Friedrich Hermanni/Volker Steenblock (Hrsg.), Philosophische Orientierungen, Festschrift zum 65. Geburtstag von Willi Oelmüller, 1995, 85-102, 86; Johannes Schwartländer, Verantwortung, in: Hermann Krings/Hans Michael Baumgartner/Christoph Wild (Hrsg.), Handbuch philosophischer Grundbegriffe, Band 6, Transzendenz – Zweck, 1974, 1577-1588, 1580.

<sup>39</sup> Wilhelm Weischedel, Das Wesen der Verantwortung, Ein Versuch, 1972, 15.

<sup>40</sup> Michael Pauen, Freiheit und Verantwortung, Wille, Determinismus und der Begriff der Person, in: Allgemeine Zeitschrift für Philosophie, 26/2001, 23-44, 23. Dieser Hinweis ist insofern interessant, als der Verantwortungsbegriff historisch zuerst im rechtlichen Bereich auftauchte. Gerade im juristischen Bereich „[bestehe] hinter allen ungelösten Streitfragen [ein] Konsens [...], dass Freiheit den Menschen für sein Wollen und Handeln im rechtlichen Sinne verantwortlich macht.“ (Ernst-Joachim Lampe, Einleitung, Teil II: Juristische Beiträge, in: Ernst-Joachim Lampe/Michael Pauen/Gerhard Roth (Hrsg.), Willensfreiheit und rechtliche Ordnung, 2008, 16-37, 30)

<sup>41</sup> Vgl. auch Michael Pauen, Freiheit, Schuld und Strafe, in: Ernst-Joachim Lampe/Michael Pauen/Gerhard Roth (Hrsg.), Willensfreiheit und rechtliche Ordnung, 2008, 41-71, 43.

<sup>42</sup> Vgl. bspw. John Martin Fischer, Responsiveness and Moral Responsibility, in: Ferdinand David Schoeman (Hrsg.), Responsibility, Character, and the Emotions, New Essays in Moral Psychology, 1988, 81-106; Achim-Lohmar, Moralische Verantwortlichkeit ohne Willensfreiheit, 2005; Stephen J. Morse, The Non-Problem of Free Will in Forensic Psychiatry and Psychology, in: Public Law and Legal Theory Research Paper Series, 2007, 203-220.

<sup>43</sup> Urteilskraft als das Vermögen, Besonderes unter Allgemeines (Begriffe und Regeln) zu subsumieren (bestimmende Urteilskraft), sowie in Ausnahmesituationen neue Regeln zu definieren (reflektierende Urteilskraft), wurde maßgeblich von Kant in der *Kritik der reinen Urteilskraft* untersucht.

<sup>44</sup> Kurt Bayertz, Eine kurze Geschichte der Herkunft der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 3-71, 13.

<sup>45</sup> Kurt Bayertz, Eine kurze Geschichte der Herkunft der Verantwortung, in: Kurt Bayertz (Hrsg.), Verantwortung, Prinzip oder Problem?, 3-71, 15.

Rümelin folgert, dass „[e]rst die Festlegung auf Gründe [...] das Handeln und die Verständigungspraxis [strukturiert] und [es] erlaubt [...], Personen als sich in der Zeit durchhaltende Akteure wahrzunehmen“<sup>46</sup>. Das Festlegen auf Normen, zu denen jemand durch Urteilskraft gelangt, gewährleistet Personalität und garantiert überdies Versprechen, Vertrauen und Verlässlichkeit sowie insbesondere die Möglichkeit von Verantwortungsübernahme. Die Verknüpfung von Urteilskraft mit den hier genannten zwischenmenschlichen Institutionen ist darin zu sehen, dass durch Erstere, bzw. durch die Ausbildung von handlungsleitenden Gründen, erst ein Normalfall konstituiert wird, aus dem Versprechen, Vertrauen und Verlässlichkeit in der sozialen Praxis folgen.

S. 20

- HFR 03/2016 S. 11 -

26 Nachdem ich in Abschnitt C einen kurzen Überblick über die philosophische Disziplin der Roboterethik gegeben habe, nehme ich das Verantwortungsphänomen in der Mensch-Maschine-Interaktion, insbesondere unter Fokussierung ihrer fünf Relationselemente als auch der für die Verantwortungsübernahme notwendigen Kompetenzen, in den Blick.

27 **C. Was ist Roboterethik?**

Innerhalb der Disziplin der Roboterethik sind zwei Felder zu unterscheiden: Die einen fragen danach, inwiefern Roboter als sogenannte „moral patients“ zu verstehen sind, also passiv als Träger\*innen moralischer Rechte, bzw. inwiefern ihnen ein moralischer Wert zukommt. Die anderen interessieren sich dafür, ob und ggf. inwiefern Roboter sogar „moral agents“ sein könnten, also aktiv Träger\*innen moralischer Pflichten bzw. moralische Handlungssubjekte.<sup>47</sup> Beide Arbeitsbereiche ergänzen einander und Verantwortung ist in beide jeweils unterschiedlich einzubinden, wie in D.I und D.II gezeigt werden soll. Die Gruppe der moral agents ist gegenüber der der moral patients exklusiver; für gewöhnlich zeichnen wir nur Menschen und längst nicht alle mit Moralfähigkeit im genuinen Sinne des Wortes aus – einige Menschen wie bspw. Kinder und solche mit spezifischen geistigen und körperlichen Einschränkungen können temporär oder sogar generell von ihrer Moralfähigkeit ganz oder teilweise entschuldigt werden. Auch bestimmte Umstände und Unfälle erlauben eine Freisprechung von Moral.

28 Einer ganzen Reihe von Wesen und Dingen wie z.B. Tieren, Pflanzen, aber auch Gegenständen wie dem teuren Auto, dem Smartphone oder Haus wird indes ein moralischer Wert zugeschrieben – zumindest in dem Sinn, dass diese Entitäten moralisch bedenkenswert sind, wenn ihnen vielleicht auch kein Eigen- sondern nur ein hoher instrumenteller Wert beigemessen wird. Als moralisches Handlungssubjekt hat man zugleich einen Platz im Kreis der Wertträger\*innen – dies gilt allerdings nicht umgekehrt. Lebewesen und Gegenständen kann man abhängig von der Perspektive einen moralischen Wert zuschreiben; eine anthropozentrische Position argumentiert bspw. dafür, dass nur dem Menschen ein Eigenwert zukommt, weitere Ansätze stellen der Patho-, der Bio- und der Physiozentrismus dar.<sup>48</sup>

S. 21

- HFR 03/2016 S. 12 -

29 Innerhalb des Arbeitsbereichs zu Robotern als Wertträger\*innen wird das menschliche Verhalten gegenüber artifiziellen Systemen in den Blick genommen. Hier geht es darum, wie mit Robotern umzugehen ist und inwiefern ihnen (ggf. analog zu Tieren und

<sup>46</sup> Julian Nida-Rümelin, Politische Verantwortung, in: Ludger Heidbrink/Alfred Hirsch (Hrsg.), Staat ohne Verantwortung? Zum Wandel der Aufgaben von Staat und Politik, 2007, 55-85, 71.

<sup>47</sup> Luciano Floridi/J. W. Sanders, On the Morality of Artificial Agents, in: Minds and Machines, 14/2004, 349-379, 349.

<sup>48</sup> Vgl. Angelika Krebs, Naturethik im Überblick, in: Angelika Krebs (Hrsg.), Naturethik, Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion, 1997, 337-379, 345. Interessant ließe sich die Überlegung an, inwiefern ein Bedenken von artifiziellen Systemen als mit einem Eigenwert ausgestattete Phänomene eine weitere Perspektive eröffnet, die all das mit einem Eigenwert bemisst, das in einer spezifischen Weise gesteuert oder programmiert bzw. lernfähig ist – ein Mathenozentrismus sozusagen (von griech. „matheno“, lernen).

kleinen Kindern) ein moralischer Wert zukommt, selbst wenn man sich darüber einig sein sollte, dass sie selbst nicht zu moralischem Handeln in der Lage sind. In dieses Themenfeld fallen alle Fragen, die artifizielle Systeme als Werkzeuge oder als Ergänzungen des Menschen verstehen, wie bspw. bei der Formulierung von Ethikkodizes in Unternehmen<sup>49</sup>, die Frage, inwiefern Beziehungen zu und mit Robotern denkbar und wünschenswert sind,<sup>50</sup> inwiefern man Roboter ‚versklaven‘ kann<sup>51</sup> und wie der Einsatz von artifiziellen Systemen zu Therapiezwecken zu beurteilen ist<sup>52</sup>. Innerhalb dieses Arbeitsbereichs verbleibt die moralische Kompetenz und Kompetenzkompetenz bei den menschlichen Designer\*innen (und u.U. auch Nutzer\*innen) artifizieller Systeme. Die menschlichen ‚Eltern‘ entscheiden über die Moral ihrer Geschöpfe und darüber, wer im Falle eines Unfalls Verantwortung trägt. Sicher ist, dass den künstlichen Kreaturen keinerlei oder zumindest nur wenig Verantwortung zuzuschreiben ist, da es ihnen an den Kompetenzen mangelt, die als Bedingung für die Möglichkeit von Verantwortungszuschreibung gelten. Auf die Möglichkeiten einer Verantwortungszuschreibung in solchen Fällen reduzierter oder nicht vorhandener Verantwortungskompetenzen (Kommunikations- und Handlungsfähigkeit sowie Urteilskraft, vgl. Abschnitt B.VI) komme ich unter dem Schlagwort der Verantwortungsnetzwerke in Abschnitt D.II zu sprechen. Verantwortung wird in der Mensch-Maschine-Interaktion häufig, wenn auch nicht immer, als rechtliche Kategorie diskutiert.<sup>53</sup>

- 30 Innerhalb des Arbeitsfelds zu Robotern als moralischen Handlungssubjekten fragt man insbesondere danach, inwiefern Roboter zu moralischem Handeln in der Lage sind, und folglich, über welche Kompetenzen sie in welchem Maße dafür verfügen müssen. Interessieren sich die einen in diesem Bereich eher für die Zuschreibung von Freiheit<sup>54</sup> als Bedingung für moralisches Handeln, befassen sich andere eher mit kognitiven Kompetenzen<sup>55</sup> (Denken, Verstehen, Geist, Intelligenz, Bewusstsein, Wahrnehmung und Kommunikation) und wieder andere mit Empathie und Emotionen<sup>56</sup>.

---

<sup>49</sup> Vgl. Elisa May, Robotik und Arbeitsschutzrecht, in: Eric Hilgendorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, 99-118.

<sup>50</sup> Vgl. David Levy, The Ethics of Robot Prostitutes, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 223-231; Matthias Scheutz, The Inherent Danger of Unidirectional Emotional Bonds between Humans and Social Robots, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 205-221; Blay Whitby, Do You Want a Robot Lover? The Ethics of Caring Technologies, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 233-247.

<sup>51</sup> Vgl. Steve Petersen, Designing People to Serve, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 283-298; Steve Petersen, The Ethics of Robot Servitude, in: in Journal of Experimental & Theoretical Artificial Intelligence, 19/2007, 43-54.

<sup>52</sup> Vgl. Heidrun Boeker, Robotik in Betreuung und Gesundheitsvorsorge, in: Technikfolgenabschätzung, Theorie und Praxis, 22/2013, 62-64; Catrin Misselhorn/Ulrike Pompe/Mog Stapleton, Ethical Considerations Regarding the Use of Social Robots in the Fourth Age, in: Geropsych 26/2013, 121-133; Jason Borenstein/Yvette Pearson, Robot Caregivers, Ethical Issues Across the Human Lifespan, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 251-265; Noel Sharkey/Amanda Sahrkey, The Rights and Wrongs of Robot Care, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 267-282; Edoardo Datteri/Guglielmo Tamburrini, Ethical Reflections on Health Care Robotics, in: Rafael Capurro/Michael Nagenborg (Hrsg.), Ethics and Robotics, 2009, 35-47.

<sup>53</sup> Vgl. Gert-Jan Lokhorst/Jeroen van den Hoven, Responsibility for Military Robots, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), Robot Ethics, The Ethical and Social Implications of Robotics, 2012, 145-156; Matthias Maring, Mensch-Maschine-Interaktion, Steuerbarkeit – Verantwortbarkeit, in: Christoph Hubig/Peter Koslowski (Hrsg.), Maschinen, die unsere Brüder werden, Mensch-Maschine-Interaktion in hybriden Systemen, 2008, 113-129; Peter M. Asaro, Robots and Responsibility from a Legal Perspective, in: Proceedings of the IEEE Conference in Robotics and Automation, Workshop on Roboethics, 2007, 20-24; Luciano Floridi/J. W. Sanders, On the Morality of Artificial Agents, in: Minds and Machines 14/2004, 349-379.

<sup>54</sup> Vgl. Georg Dorffner, Rationalität, Emotionalität und Körperlichkeit, Können Maschinen Menschen verstehen – und umgekehrt?, in: Heinrich Schmidinger/Clemens Sedmak (Hrsg.), Der Mensch – ein freies Wesen? Autonomie – Personalität – Verantwortung, 2005, 102-112

<sup>55</sup> Vgl. Anthony Clark, Artificial Intelligence and the Many Faces of Reason, in: Stephen P. Stich/Ted A. Warfield (Hrsg.), The Blackwell Guide to Philosophy of Mind, 2003, 309-321; Anthony Clark, Towards a Cognitive Robotics, in: Adaptive Behavior, 7/1999, 5-16; Daniel C. Dennett, Cognitive Wheels, The Frame Problem of AI, in: Kenneth M. Ford/Clark Glymour/Patrick J. Hayes (Hrsg.), Thinking About Android Epistemology, 2006, 147-169; Daniel C. Dennett, Brainchildren, Essays on Designing Mind, 1998.

- 31 Beiden Arbeitsfeldern innerhalb der Roboterethik liegt die Frage zugrunde, was Moral bzw. was Ethik ist und wie moralische Urteile gefällt werden.<sup>57</sup> Auch hier lassen sich (allerdings nicht in diesem Artikel) verschiedene Positionen unterscheiden; in einem ersten Schritt könnte man vorschlagen, allen Wesen Moralfähigkeit zuzuschreiben, die in Situationen geraten, in denen moralische Entscheidungen zu treffen sind. So gehen bspw. Wendell Wallach und Colin Allen in ihrem Werk *Moral Machines. Teaching Robots Right from Wrong* (2009) vor. Sie beschreiben vor der Reflexionsfolie von Philippa Foot's klassischem Gedankenexperiment der Trolley Cases den Fall von „driverless“ train systems<sup>58</sup>, in denen in London, Paris und Kopenhagen bereits seit Mitte der 1960er Jahre Menschen nur als Fahrgäste anzutreffen sind.<sup>59</sup> Eine moralische Entscheidung wird – so Wallach und Allen – bereits dann gefällt, wenn sich auf den Gleisen Menschen befinden, die der Zug zu überrollen droht. Der Zug ‚urteilt‘, indem er dazu programmiert ist, immer dann unverzüglich zu stoppen, wenn sich Menschen auf den Gleisen aufhalten, selbst wenn damit ggf. Unfälle im Zuginnern in Kauf genommen werden müssen.
- 32 Um diesem Gedanken, dass all denjenigen Wesen (rudimentäre) Moralfähigkeit zuzuschreiben ist, die wie Menschen in Situationen geraten, in denen moralische Entscheidungen zu treffen sind, noch ein wenig zu folgen, stellen wir uns einen selbstfahrenden Krankentransport in einer Notfallsituation vor, der sich mit bspw. einer schwer verletzten Jugendlichen auf dem Weg ins nächste Krankenhaus befindet. Absolut unvorhersehbar überquert nun unmittelbar vor dem Krankentransport eine weitere Jugendliche die Straße. Ein plötzliches Bremsen hätte den Tod der Transportierten zur Folge, ein Weiterfahren den Tod der Fußgängerin. Sowohl eine Entscheidung für die eine als auch für die andere Option stellt ein moralisches Urteil dar bzw. beruht auf einem moralischen Prinzip. Es kann zunächst keine Rede davon sein, dass ein autonomer Krankentransport, ausgerüstet mit einer spezifischen algorithmischen Struktur, im genuinen Sinne des Wortes moralisch handelt. Allerdings ähnelt diese Situation äußerlich einer solchen, in der sich auch ein Mensch befinden könnte. In ihrer von außen beobachtbaren phänomenologischen Qualität gleicht die Maschine – so Wallach und Allen – rudimentär einem Menschen. Das genügt, um zumindest ein Nachdenken über Roboter als moral agents nachvollziehbar erscheinen zu lassen, ohne, dass man sich gleich zu schließen gezwungen fühlen müsste, dass artifizielle Systeme per se, in derselben Weise und in demselben Ausmaß zu moralischem Handeln befähigt seien wie Menschen. Wallach und Allen beschreiben mit ihrem Ansatz – auf den im nun folgenden Abschnitt D.I ausführlich eingegangen wird – eine Version der schwachen KI-These.

<sup>56</sup> Vgl. Mark Coeckelbergh, *Moral Appearances, Emotions, Robots, and Human Morality*, In: *International Journal of Social Robotics*, 1/2009, 217-221.

<sup>57</sup> Die schwedische Serie *Real Humans – Echte Menschen (Äktamänniskor)*, 2012 von Lars Lundström vereint beide Arbeitsfelder miteinander, indem es auf der einen Seite um die Einführung von Hubots (für *Human Robots*) geht, die als hochkomplexe Serviceroboter im Dienstleistungssektor, in der Industrie und in Privathäusern arbeiten. Auf der anderen Seite treten sog. „freie“ Hubots auf, die über einen besonderen „Code“ verfügen (also über eine spezielle algorithmische Struktur), durch die sie zu autonomen Handlungssubjekten werden.

<sup>58</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 14.

<sup>59</sup> Hieran schließt sich die Debatte um autonome Fahrassistenzsysteme; vgl. Wulf Loh/Janina Sombetzki, *Autonomy and Responsibility in Hybrid Systems – The Example of Autonomous Cars*, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), *Robot Ethics*, 2017 (im Erscheinen); Alexander Hevelke/Julian Nida-Rümelin, *Intelligente Autos im Dilemma*, in: *Spektrum der Wissenschaft*, 10/2015, 82-85; Markus Maurer/J. Christian Gerdes/Barbara Lenz/Hermann Winner (Hrsg.), *Autonomes Fahren, Technische, rechtliche und gesellschaftliche Aspekte*, 2015; G. Both/J. Weber, *Hands-Free Driving? Automatisiertes Fahren und Mensch-Maschine Interaktion*, in: Eric Hilgendorf (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, 171-187; S. Hötitzsch/E. May, *Rechtliche Problemfelder beim Einsatz automatisierter Systeme im Straßenverkehr*, in: Eric Hilgendorf (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, 189-210; Peter M. Knoll, *Prädikative Fahrassistenzsysteme – Bevormundung des Fahrers oder realer Kundennutzen?*, in: Christoph Hubig/Peter Koslowski (Hrsg.), *Maschinen, die unsere Brüder werden. Mensch-Maschine-Interaktion in hybriden Systemen*, 2008, 159-171.

S. 23

- HFR 03/2016 S. 14 -

33 **D. Verantwortung in der Mensch-Maschine-Interaktion**

Häufig wird die Möglichkeit einer Verantwortungsübernahme von artifiziellen Systemen mit dem Verweis auf die Kompetenzen als Bedingung für die Zuschreibung von Verantwortung bestritten, die bei den fraglichen Maschinen nicht vorlägen: Roboter würden weder über Urteilskraft, Handlungsfähigkeit, Autonomie noch über sonstige Fähigkeiten verfügen, die für die Übernahme von Verantwortung eine Rolle spielten. Wendell Wallach und Colin Allen formulieren in ihrem Buch *Moral Machines. Teaching Robots Right from Wrong* (2009) den Ansatz der funktionalen Äquivalenz, mit dem dieses Problem mangelnder Kompetenzen bei artifiziellen Systemen umgangen werden kann. Die beiden folgenden Abschnitte nehmen die Rolle der Verantwortung in den beiden Arbeitsfeldern der Roboterethik – Roboter als moral agents bzw. Handlungssubjekte (D.I mit Bezugnahme insbesondere auf B.VI) und Roboter als moral patients bzw. Wertträger\*innen (D.II unter Rekurs insbesondere auf B.I-B.V) – in den Blick.

34 **D.I. Roboter als Handlungssubjekte und die Bedingungen für die Möglichkeit zur Verantwortungsübernahme – Wallachs und Allens Ansatz funktionaler Äquivalenz<sup>60</sup>**

Indem Wallach und Allen die Frage stellen, inwiefern Roboter als artifizielle moralische Akteure zu verstehen sind, definieren sie „moral agency“ als graduelles Konzept mit zwei Bedingungen, nämlich Autonomie und Empfänglichkeit bzw. Empfindlichkeit für moralische Werte („sensitivity to values“<sup>61</sup>). Menschen gelten als moralische Akteure im genuinen Sinne, allerdings sind einige Maschinen – bspw. ein Autopilot, oder das artifizielle System Kismet – operationale moralische Akteure zu nennen. Sie sind autonomer und ethisch empfänglicher als manch ein anderes nicht-mechanisches Werkzeug wie z. B. ein Hammer, und dennoch verbleiben sie immer noch „totally within the control of [the] tool’s designers and users“<sup>62</sup>. In diesem Sinne sind operationale moralische Akteure „direct extensions of their designers’ values“<sup>63</sup>. Nur besondere artifizielle Systeme haben bereits den Status funktionaler moralischer Akteursfähigkeit – so wie bspw. das medizinische ethische Expertensystem MedEthEx<sup>64</sup>. Funktionale Moralität bedeutet, dass das fragliche artifizielle System insofern entweder autonomer und/oder Werte-sensitiver ist, als ein operationaler moralischer artifizieller Akteur, als funktionale moralische Maschinen „themselves have the capacity for assessing and responding to moral challenges“<sup>65</sup>. Wallachs und Allens Ansatz einer graduellen Zuschreibung von Kompetenzen und Fähigkeiten gründet auf dem Gedanken der funktionalen Äquivalenz:

„Just as a computer system can represent emotions without having emotions, computer systems may be capable of functioning as if they understand the meaning of symbols without actually having what one would consider to be human understanding.“<sup>66</sup>

S. 24

- HFR 03/2016 S. 15 -

35 Mit ihrem Ansatz funktionaler Äquivalenz beschreiben Wallach und Allen eine Version der schwachen KI-These<sup>67</sup>, der an der Simulation spezifischer Kompetenzen in artifiziellen Systemen gelegen ist und nicht daran, Maschinen tatsächlich im genuinen Sinne des Wortes mit Intelligenz, Bewusstsein und Autonomie zu konstruieren (starke KI, irr-

<sup>60</sup> Die Überlegungen in diesem Abschnitt können in zahlreichen meiner Artikel nachgelesen werden – am aktuellsten vgl. Wulf Loh/Janina Sombetzki, *Autonomy and Responsibility in Hybrid Systems – The Example of Autonomous Cars*, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), *Robot Ethics*, 2017 (im Erscheinen).

<sup>61</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 25.

<sup>62</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 26.

<sup>63</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 30.

<sup>64</sup> Michael Anderson/Susan Leigh Anderson/C. Armen, *An Approach to Computing Ethics*, in: *Intelligent Systems*, IEEE 4/2006, 2-9.

<sup>65</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 9.

<sup>66</sup> Wendell Wallach/Colin Allen, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 69.

<sup>67</sup> John R. Searle, *Minds, brains and programs*, in: *Behavioral and Brain Sciences*, 3/3/1980, 417-157.

tümlich auf Turing zurückgeführt<sup>68</sup>). Wallach und Allen zufolge stellt eine Annahme starker KI und der daran geknüpften Kompetenzen keine notwendige Bedingung für artifizielle moralische Akteure dar, sondern sie fokussieren die Zuschreibung von funktional äquivalenten Bedingungen und Verhaltensweisen. Funktionale Äquivalenz bedeutet, dass spezifische Phänomene verstanden werden, „als ob“ sie – um das Kant'sche Vokabular der regulativen Ideen zu nutzen – kognitiven, emotionalen oder anderen Kompetenzen und Fähigkeiten entsprechen. Es sei an dieser Stelle daran erinnert, dass dieses Argument bereits auf Menschen zutrifft, spätestens allerdings auf Tiere.<sup>69</sup> Bezüglich anderer Menschen sind wir generell bereit, prima facie Fähigkeiten wie Vernunft, Bewusstsein und Willensfreiheit zu statuieren, wenn auch keine Garantie dafür besteht, dass die fraglichen Individuen tatsächlich mit besagten Kompetenzen ausgestattet sind.<sup>70</sup> Die Frage, inwiefern artifizielle Systeme irgendwann in der Tat intelligent, bewusst oder autonom im Sinne der starken KI-These genannt werden können, wird durch die Frage ersetzt, in welchem Ausmaß und Umfang die fraglichen Kompetenzen der Funktion entsprechen, die sie innerhalb der moralischen Evaluation spielen – in diesem Fall dem Verantwortungskonzept. Vernünftiges oder autonomes Handeln eines artifiziellen Systems wird nur insofern in den Blick genommen, als es als Voraussetzung dafür zu deuten ist, Verantwortung zuzuschreiben.

- 36 Obwohl Wallach und Allen den Übergang von operationaler über funktionaler bis hin zu voller bzw. genuiner Moralzuschreibung in Abhängigkeit von den vorliegenden Kompetenzen (Autonomie und moralische Sensitivität) graduell denken, bleibt es doch schwer vorstellbar, wie (zumindest in der nahen Zukunft) ein artifizielles System ein funktionales Äquivalent zu der genuin menschlichen Fähigkeit „second order volitions“<sup>71</sup> zu bilden entwickeln könnte bzw. zu der Fähigkeit, als „self-authenticating sources of valid claims“<sup>72</sup> über selbst gesetzte moralische Prämissen und Prinzipien zu reflektieren. Hilfreich erscheint an dieser Stelle Darwalls Unterscheidung zwischen vier Formen von Autonomie: „personal“, „moral“, „rational“ und „agential“ Autonomie.<sup>73</sup> Während persönliche Autonomie die Fähigkeit umfasst, Werte, Ziele und letzte Zwecke zu definieren, zielt moralische Autonomie auf die Möglichkeit, selbst gesetzte Prinzipien und ethische Überzeugungen zu reflektieren. Diese beiden Formen von Autonomie werden wohl noch für eine lange Zeit menschlichen Handlungssubjekten vorbehalten bleiben. Rationale Autonomie hingegen, die Darwall zufolge auf Handlungen auf der Basis der „weightiest reasons“ gründet, scheint prima facie auch für artifizielle Akteure erreichbar, insofern die besagten Gründe funktional äquivalent bspw. in Form von Algorithmen repräsentiert werden können. Und erst recht scheint Akteursautonomie Maschinen zugeschrieben werden zu können, die darin besteht, ein spezifisches Verhalten als „genuine Handlung“ – also nicht vollständig durch externe Faktoren determiniert – zu identifizieren. „Agential“ Autonomie kann funktional äquivalent durch die Fähigkeit simuliert werden, interne Zustände eines artifiziellen Systems ohne externe Stimuli zu ändern.

## S. 25

- HFR 03/2016 S. 16 -

- 37 Auf der computationalen Ebene ließe sich die Fähigkeit, autonom interne Zustände zu ändern, bspw. in Form eines algorithmischen Strukturschemas näher beschreiben,<sup>74</sup> wodurch es nicht nur möglich wäre, die funktionale Äquivalenz von Autonomie zu erklären, sondern auch ein Licht auf die Unterscheidung von operationaler und funktionaler Verantwortungszuschreibung zu werfen. Schließlich geht es im Rahmen dieses

<sup>68</sup> Alan Turing, *Computing Machinery and Intelligence*, in: *Mind*, 59/236/1950, 433-460.

<sup>69</sup> Thomas Nagel, *What is it like to be a bat?*, in: *The Philosophical Review*, LXXXIII, 1974, 435-450.

<sup>70</sup> Mark Coeckelbergh, *The Moral Standing of Machines, Towards a Relational and Non-Cartesian Moral Hermeneutics*, in: *Philosophy and Technology*, 27/2014, 61-77, 63.

<sup>71</sup> Harry Frankfurt, *Freedom of the Will and the Concept of a Person*, in: *Journal of Philosophy*, 68/1/1971, 5-20.

<sup>72</sup> John Rawls, *Justice as fairness, A restatement*, 2001, 23.

<sup>73</sup> Stephen Darwall, *The Value of Autonomy and Autonomy of the Will*, in: *Ethics* 116/2006, 263-284, 265. Ich danke Wulf Loh für diese Idee der Ergänzung des Wallch-Allen'schen Modells um Darwalls Autonomie-Typen.

<sup>74</sup> Vgl. Janina Sombetzki, *Roboterethik – ein kritischer Überblick*, in: Matthias Maring (Hrsg.), *Zur Zukunft der Bereichsethiken, Herausforderungen durch die Ökonomisierung der Welt*, 2016 (im Erscheinen).

Abschnittes darum, zu erklären, inwiefern Roboter als moralische Handlungssubjekte (im Sinne des ersten Arbeitsfeldes der Roboterethik) identifiziert, sodass sie ebenso als Verantwortungssubjekte verstanden werden können. Dafür muss es möglich sein, ihnen in einem funktional äquivalenten Sinne die Bedingungen für die Möglichkeit von Verantwortungsübernahme zuzuschreiben (Kommunikations- und Handlungsfähigkeit sowie Urteilskraft). Während „determinierte Algorithmen“ bei gleichem Input zu demselben Output gelangen, gelangen „deterministische Algorithmen“ bei gleichem Input über dieselben Zwischenschritte zum selben Output. Vielleicht wäre es vorstellbar, Maschinen, die vornehmlich auf der Grundlage deterministischer Algorithmen funktionieren, in weder der funktionalen noch in der operationalen Sphäre zu verorten – sie immer noch als Maschinen zu sehen, allerdings fast den nicht-mechanischen Werkzeugen näher als der operationalen Sphäre. Diese wird eindeutig mit artifiziellen Systemen betreten, die vornehmlich durch determinierte (aber nicht-deterministische) Algorithmen strukturiert sind. Und schließlich könnten solche seltenen Fälle artifizieller Systeme, die vornehmlich auf der Grundlage nicht-determinierter (und nicht-deterministischer) Algorithmen operieren, in der funktionalen Sphäre lokalisiert werden.

- 38 Schauen wir uns ein paar Beispiele an: Das artifizielle System Kismet, das Wallach und Allen als operationalen moralischen Akteur interpretieren, verfügt in einem äußerst rudimentären Sinne über Kommunikationsfähigkeit, insofern es nur wenige Laute und Worte hervorbringen kann. Urteilskraft – sollte man gewillt sein, Kismets Verhalten überhaupt vernünftig zu nennen – ist minimal darin zu sehen, dass er auf sehr einfache Fragen reagiert. Allerdings (es sei an die intersubjektiven Subkompetenzen der Urteilskraft erinnert, vgl. B.VI) wäre es möglich, Kismet in seiner Begrenztheit doch Vertrauen entgegenzubringen, da dieser Roboter gar nicht die Möglichkeit hat, aus seinem vorgegebenen Handlungsrahmen auszubrechen. Die größte Herausforderung, Kismet als einen operational verantwortlichen Akteur zu verstehen, liegt wohl in der Zuschreibung von Handlungsfähigkeit bzw. Autonomie in ihrer ganzen Komplexität, da Kismets Folgenbewusstsein (Wissen), Kontextwahrnehmung als Geschichtlichkeit, seine Personalität sowie Einflussmöglichkeit doch sehr beschränkt sind. Kismet kann seine Ohren, Augen, Lippen sowie seinen Kopf bewegen und reagiert auf externe Stimuli wie die menschliche Stimme – das ist aber auch schon alles. Zusammengefasst steht Kismet – in Übereinstimmung mit Wallachs und Allens Urteil – immer noch vollständig unter der Kontrolle seiner Nutzer, seine Algorithmen gelangen ausschließlich zu deterministischen Ergebnissen. Kismet verantwortungsbefähigt zu nennen, ist vielleicht damit vergleichbar, einem Säugling oder einigen Tieren Verantwortung zuzuschreiben. Und dennoch – verglichen mit dem Regen, der für das Nass-Sein der Straße nur in einem metaphorischen Sinne verantwortlich sein kann – eröffnet die Verantwortungszuschreibung bzgl. Kismet einen Diskussionsraum darüber, ob es sinnvoll ist, einige artifizielle Systeme als potenzielle Verantwortungssubjekte zu identifizieren, auch wenn dieser Raum verständlicherweise recht klein ist.

**S. 26**

- HFR 03/2016 S. 17 -

- 39 Der Roboter Cog ist vielleicht als Beispiel für einen in sehr schwachem Maße funktional verantwortlichen Akteur zu sehen, insofern seine Kommunikationsfähigkeit und seine Urteilskraft gegenüber Kismet doch deutlich gesteigert sind. Und was noch wesentlicher erscheinen mag – seine Handlungsfähigkeit bzw. Autonomie ist aufgrund eines „unsupervised learning algorithm“<sup>75</sup> deutlich komplexer als die Kismets. So beginnt Cog, ohne, dass er zuvor in dieser Weise programmiert worden wäre, ein Spielzeugauto nur noch von vorne oder hinten anzustoßen, um es in Bewegung zu versetzen, nachdem er durch mehrere Versuche feststellen konnte, dass es sich nicht bewegt, wenn es von der Seite angestoßen wird. Cog lernt durch Erfahrung, und vielleicht ist es gerade diese (in seinem Fall in der Tat begrenzte) Fähigkeit zu lernen, die es uns erlaubt, ihn als einen schwachen funktionalen Akteur zu verstehen oder aber als immer-

<sup>75</sup> Rodney A. Brooks/Cynthia Breazeal/Matthew Marjanović/Brian Scasselatti/Matthew M. Williamson, The Cog Project, Building a Humanoid Robot, in: Christopher Nehaniv (Hrsg.), *Computation for Metaphors, Analogy, and Agents*, 1999, 52-87, 70.



hin stark operational verantwortlich. Cog verantwortungsbefähigt zu nennen, ist wohl vergleichbar damit, einem jungen Kind Verantwortung zuzuschreiben. Vor dem Hintergrund dieser Überlegungen lässt sich Wallachs und Allens Intuition, das oben bereits erwähnte medizinische Expertensystem MedEthEx als funktionalen moralischen Akteur zu begreifen, nun bestätigen, denn sowohl dessen Urteilskraft, als auch sein Handlungswissen und seine Autonomie sind deutlich gesteigert im Vergleich zu den zuvor genannten Beispielen.

- 40 Und schlussendlich lassen sich autonome Fahrassistenzsysteme als ein weiteres Beispiel für operational verantwortliche Akteure anführen, denn zwar mögen ihre Kommunikationsfähigkeiten und Urteilskraft ähnlich entwickelt sein wie die von Cog oder sogar weiter. Allerdings ist die Handlungsfähigkeit bzw. Autonomie autonomer Fahrassistenzsysteme aus guten Gründen in strengen Grenzen gehalten; sie können nicht lernen und verfügen nicht über nicht-determinierte (nicht-deterministische) Algorithmen. Im folgenden Abschnitt werde ich darauf zurückkommen, was zu tun ist, wenn Roboter nicht oder nur zu einem sehr geringen Ausmaß als Verantwortungsakteure einzustufen sind.
- 41 Zusammengefasst kann mit Hilfe von Darwalls Differenzierung zwischen vier Formen der Autonomie in Kombination mit Wallachs und Allens Ansatz funktionaler Äquivalenz eine klare Grenze zwischen genuiner (menschlicher) Akteursfähigkeit im vollen Sinne und artifizierlicher (operationaler und funktionaler) Akteursfähigkeit gezogen werden. Während menschliche Handlungssubjekte über alle vier Autonomietypen verfügen, kann Maschinen zumindest auf absehbare Zeit nur rationale und „agential“ Autonomie in einer funktional äquivalenten Weise zugeschrieben werden. Insofern es um verschiedene Bereiche der Verantwortungszuschreibung geht (moralische, rechtliche, politische Verantwortung), ist ein artifizielles System dann autonom (bzw. handlungsfähig) zu nennen, wenn es die Kriterien funktionaler Moralzuschreibung erfüllt. Eine generelle Modifikation der implementierten algorithmischen Strukturen ist wohl bei keinem artifiziellem System – selbst bei rein nicht-determinierten (nicht-deterministischen) Sets an Algorithmen – im selben Ausmaß wie im Rahmen der menschlichen Entwicklung vorstellbar, von der Wünschbarkeit ganz zu schweigen. Trotzdem mutet vor dem Hintergrund des gerade Gesagten der Einwand trivial an, dass doch auch im Falle nicht-determinierter (nicht-deterministischer) Algorithmen nicht alle vorstellbaren Ergebnisse möglich sind, dass also artifizielle Systeme letztlich immer programmiert sind. Denn auch Menschen sind für gewöhnlich nicht zu allem in der Lage, sondern bleiben in ihren Möglichkeiten ebenfalls beschränkt, selbst wenn man ihren adaptiven Spielraum sehr viel größer einschätzt als der eines noch so komplexen Roboters jemals sein könnte.

**S. 27**

- HFR 03/2016 S. 18 -

- 42 Artifizielle Verantwortungszuschreibung ist bislang nur in begrenztem Maße möglich. Die besten Aussichten gewähren evolutionäre Lernmodelle; maschinelles Lernen wird hier äquivalent zum kindlichen Lernen untersucht. Diese Ansätze beruhen auf einer meta-ethischen Annahme über die Kontextsensitivität von Moral. Moralisches und verantwortliches Handeln bedarf der Erfahrung und eines situativen Urteilsvermögens. Beides kann sich ein artifizielles System nur verkörpert aneignen. In den 1990er Jahren war es u.a. Brooks, der als einer der ersten das Zusammenwirken von artifiziellem System und Umwelt als Bedingung für die Entwicklung von Kompetenzen betrachtete und von dieser Annahme ausgehend das Feld der „behavior-based robotics“ begründete.<sup>76</sup> Zahlreiche Projekte, die sich an dem Ansatz verkörperten menschlichen Lernens orientieren – wie bspw. die Lernplattformen iCub, Myon, Cb<sup>2</sup>, Curi, Roboy (die im Detail unterschiedlichen evolutionsbasierten Ansätzen folgen) –, entwickeln Systeme, die sich ähnlich Kindern Kompetenzen aneignen, aus denen sie dann in spezifischen Kontexten konkrete Handlungsprinzipien ableiten. Eine Vermutung geht dahin, dass wir uns hier

---

<sup>76</sup> Vgl. Rodney A. Brooks/Cynthia Breazeal/Matthew Marjanović/Brian Scasselatti/Matthew M. Williamson, The Cog Project, Building a Humanoid Robot, in: Christopher Nehaniv (Hrsg.), *Computation for Metaphors, Analogy, and Agents*, 1999, 52-87; Rodney A. Brooks, *Intelligence Without Reason*, in: *Proc. IJCAI-91*, 1991, 569-595.

im Bereich nicht-determinierter (nicht-deterministischer) Algorithmen aufhalten, die eigenständig lernen können (evolutionäre, genetische Algorithmen).

Bislang ist maschinelles Lernen jedoch nur in nicht moralischen bzw. schwach moralischen Kontexten möglich. Von Verantwortungsübernahme ließe sich also bislang nur in einem nicht (rein) moralischen (vielleicht juristischen) Sinne sprechen, artifizielle Systeme sind (noch) keine voll verantwortungsbefähigten Akteure.

#### 43 **D.II. Roboter als Wertträger\*innen und die Relationselemente der Verantwortung – Verantwortungsnetzwerke**

Im vorherigen Abschnitt wurde über Wallachs und Allens Ansatz funktionaler Äquivalenz festgestellt, dass – mit Blick auf das Arbeitsfeld der Roboterethik, in dem über Roboter als moralische Handlungssubjekte nachgedacht wird – artifizielle Systeme bislang und in absehbarer Zukunft nicht als Verantwortungsakteure zu identifizieren sind, insofern die zur Verantwortungszuschreibung nötigen Kompetenzen (Kommunikations- und Handlungsfähigkeit, sowie Urteilskraft) nur in einem schwach funktionalen oder gar nur in einem operationalen Sinne äquivalent simuliert werden können. Da wir Roboter gegenwärtig nicht im exklusiven Kreis der moral agents finden, möchte ich nun zum Abschluss dieses Artikels ein paar Überlegungen dazu anstellen, welche Rolle dem Verantwortungsphänomen innerhalb des Arbeitsfeldes der Roboterethik zu artifiziellen Systemen als moral patients – als Wertträger\*innen – zuzugestehen ist. Hierzu wird eine Rückbesinnung auf die in den Abschnitten B.I bis B.V vorgestellten fünf Relationselemente der Verantwortung hilfreich sein.

S. 28

- HFR 03/2016 S. 19 -

44 Wie bereits festgestellt, ist unser traditionelles Verständnis von Verantwortung (etymologisch durch die drei Komponenten wiedergegeben, dass es sich dabei normativ um die Fähigkeit des Rede-und-Antwort-Stehens handelt, die mit einer spezifischen psychomotivationalen Verfasstheit des fraglichen Verantwortungssubjekts einhergeht; vgl. die Einleitung zu B) insofern ein stark individualistisches, als wir immer ein Subjekt benötigen, das als Verantwortungsträger\*in fungiert (vgl. B.I). Das ist auch bei der Zuschreibung kollektiver Verantwortung der Fall und schließlich auch dann, wenn wir – obwohl mir an dieser Stelle der Raum für die nötigen Ausführungen fehlt – anderen Subjekten als natürlichen Personen Verantwortung zuschreiben. Nur dann ist die Zuschreibung von Verantwortung nicht oder zumindest nur metaphorisch möglich, wenn die potenziellen Subjekte die nötigen Kompetenzen nicht oder nicht hinreichend ausgeprägt mitbringen – wie im Fall von Pflanzen, Tieren, Kindern, Menschen mit einer körperlichen oder geistigen Beeinträchtigung oder Behinderung oder Maschinen. In Fällen, in denen wir Verantwortung zuschreiben wollen, aber die Subjektposition der fraglichen Verantwortlichkeit nicht besetzbar erscheint, haben einige Verantwortungstheoretiker\*innen in den vergangenen Jahren behelfsmäßige Begrifflichkeiten zu entwickeln versucht, die ohne eine Bestimmung dieses Relationselements auskommen.<sup>77</sup> Ich bin skeptisch, dass damit bzgl. der eigentlichen Aufgabe, die das Verantwortungskonzept hat, nämlich – wie in der Einleitung (Abschnitt A) ausgeführt – in intransparenten Kontexten, die durch komplexe Hierarchien und vielfach vermittelte Handlungsabläufe gekennzeichnet sind, für mehr Struktur, mehr Transparenz und Handlungsorientierung zu sorgen, geholfen ist. Schließlich suchen wir de facto immer nach einer Trägerin oder einem Träger (Singular oder Plural), die in der Lage sind, die eingeforderte Verantwortung zu schultern. Traditionell und etymologisch funktioniert unser Verständnis von Verantwortung in dieser Weise, weshalb es in meinen Augen nicht praktikabel erscheint, dieses grundlegende Relationselement einfach nicht zu besetzen bzw. es mit einer Wortblase wie dem „System“ nur oberflächlich zu bestimmen.

<sup>77</sup> Vgl. Günter Wilhelms, Systemverantwortung, in: Ludger Heidbrink/Claus Langbehn/Janina Sombetzki (Hrsg.), Handbuch Verantwortung, 2016 (im Erscheinen).

- 45 Jetzt scheint es allerdings in der Tat so zu sein, dass wir uns in Situationen wiederfinden, in denen einige der in das fragliche Geschehen involvierten Parteien die zur Verantwortungszuschreibung notwendigen Kompetenzen nicht oder nur in einem geringen Ausmaß mitbringen, wir aber dennoch die deutliche Intuition haben, dass es hier um Verantwortung geht, ohne doch zu wissen, wer nun in welchem Ausmaß zur Verantwortungsübernahme angesprochen sein kann. Nehmen wir das Beispiel autonomer Fahrassistenzsysteme, die in Abschnitt D.I als operational verantwortliche künstliche Akteure eingestuft wurden, vergleichbar mit der Verantwortungsbefähigung eines Säuglings, Tieres oder eines sehr jungen Kindes. Das autonome Fahrassistenzsystem kann zwar einen Wertträger insofern abgeben, als es Teil unseres moralischen Universums und moralisch bedenkenswert ist, dass ihm ein instrumenteller Wert zugeschrieben wird – aber als moralischen Akteur in einem signifikanten (d. h. zumindest in einem funktionalen Sinne) lässt es sich nicht begreifen. Und dennoch wissen wir nicht so recht, ob wir es aus der Verantwortung gänzlich entlassen können.
- 46 Für solche und vergleichbare Kontexte möchte ich den Begriff des Verantwortungszusammenhangs von Christian Neuhäuser<sup>78</sup> übernehmen und spezifizieren.<sup>79</sup> Die diesen Überlegungen zugrundeliegende These lautet, dass wir all denjenigen Parteien in einer gegebenen Situation Verantwortung zuschreiben, die an dem fraglichen Geschehen beteiligt sind, in dem Maße, in dem sie die nötigen Kompetenzen zur Verantwortungszuschreibung mitbringen. Ein Verantwortungszusammenhang trägt der Tatsache explizit Rechnung, dass sich innerhalb einer Verantwortungszusammenhangskonstellation in manchen Fällen Relationselemente überlagern können wie in dem Fall der Verantwortung der Eltern für ihre Kinder, in dem die Kinder (bzw. deren Wohlergehen) einerseits das Objekt besagter Verantwortlichkeit darstellen, andererseits auch die Adressat\*innen (also den Grund des Vorhandenseins dieser Verantwortlichkeit).<sup>80</sup>

**S. 29**

- HFR 03/2016 S. 20 -

- 47 Innerhalb des Verantwortungszusammenhangs „Verantwortung im Straßenverkehr“ gehören die autonomen Fahrassistenzsysteme ebenso dazu wie die menschliche Fahrerin (selbst dann, wenn sie nicht aktiv am Fahrprozess beteiligt ist), die Besitzer\*innen, die Vertreter\*innen, die Programmierer\*innen und die Designer\*innen, aber auch die Öffentlichkeit, Jurist\*innen, Fahrlehrer\*innen und alle am Straßenverkehr Beteiligten. Verantwortungszusammenhänge haben häufig ungewöhnliche Ausmaße – als weitere Beispiele wären die Verantwortungszusammenhänge „Klimaverantwortung“, „Verantwortung in internationalen Beziehungen“ und auch „deutsche Verantwortung“ zu nennen – und bündeln in sich unterschiedliche Verantwortungsobjekte. Von Verantwortungszusammenhängen kann man dann sprechen, wenn man eigentlich – sehr schön zu veranschaulichen am Fall der Klimaverantwortung<sup>81</sup> – gar nicht mehr weiß, ob hier in einem gehaltvollen Sinn Verantwortung definiert werden kann, gerade weil bspw. die Bestimmung eines Subjekts schwierig erscheint oder aber sich keine eindeutige Instanz ausmachen lässt oder aber die normativen Kriterien nicht benannt werden können. In einem Verantwortungszusammenhang erfüllen die involvierten Parteien unterschiedliche Funktionen bzw. be-

<sup>78</sup> Vgl. Christian Neuhäuser, Roboter und moralische Verantwortung, in: Eric Hilgendorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, 269-286.

<sup>79</sup> Den ich erstmals mit Wulf Loh in dem gemeinsamen Artikel *Autonomy and Responsibility in Hybrid Systems – The Example of Autonomous Cars*, in: Patrick Lin/Keith Abney/George Bekey (Hrsg.), *Robot Ethics*, 2017 (im Erscheinen) auf Wulf Lohs Idee hin zum Einsatz bringe.

<sup>80</sup> Zu diesem Beispiel vgl. ausführlich Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, 117 f. Mit einem Blick auf Normen wie bspw. § 254 BGB lässt sich vielleicht bereits eine ansatzweise zivilrechtliche Realisierung des Verantwortungszusammenhangs vermuten. Danach wird der zu ersetzende Schaden je nach Ausmaß des Verschuldens gequotelt; in diese Richtung geht auch die Möglichkeit der in Anspruch Genommenen, andere Schuldner\*innen in Regress zu nehmen, nachdem bereits Schadensersatz geleistet wurde (Bspw. kann bei einem Unfall, bei dem ein autonomes Fahrassistenzsystem beteiligt ist, die bzw. der Geschädigte zunächst die Halterin oder den Halter des Fahrzeugs in Anspruch nehmen. Diese\*r kann ggf. die Herstellerin des Fahrzeugs in Regress nehmen, welche sich wiederum an die Herstellerin der Software wenden kann.). Ich danke Frau Rebecca Sieber für diesen Hinweis.

<sup>81</sup> Zur Klimaverantwortung und den mit dieser Verantwortlichkeit einhergehenden Herausforderungen vgl. ausführlich Janina Sombetzki, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, Kapitel 13.

setzen manchmal mehrere Relationspositionen zugleich, sind einmal die Verantwortungssubjekte, in einem anderen Fall die Instanzen und wieder in einem anderen Fall das Objekt und vielleicht zugleich Adressat\*innen einer Verantwortlichkeit.

48 Es wäre äußerst schwierig, ein oder mehrere konkrete Verantwortungssubjekte für die Verantwortung im Straßenverkehr auszumachen, da diese viel zu „groß“ und umfassend ist, als dass eine Person oder eine geringe Anzahl Einzelner dafür Rede und Antwort stehen könnte. Als Verantwortungsnetzwerk „Verantwortung im Straßenverkehr“ werden hier jedoch mehrere Verantwortungsbereiche – bspw. moralische, juristische und politische Verantwortlichkeiten (definiert über moralische, juristische und politische Normen) – umfasst. Der Straßenverkehr stellt nur das übergeordnete Verantwortungsobjekt dar, für das nicht eine oder mehrere Personen gehaltvoll „die Verantwortung“ tragen, das sich jedoch in unterschiedliche „kleinere“ und weniger komplexe Verantwortungsgegenstände ausdifferenziert, für die dann die unterschiedlichen Parteien jeweils eine spezifische Verantwortung übernehmen. Verantwortung für den Straßenverkehr kann in einem Fall die Sicherheit der am Straßenverkehr beteiligten Menschen bedeuten, in einem anderen Verständnis die Verantwortung dafür, schnell und effizient von A nach B zu gelangen und in noch einem anderen Fall die Verantwortung dafür, dass die moralischen und ethischen Herausforderungen, die mit einer Beteiligung am Straßenverkehr einhergehen, hinreichend diskutiert bzw. denjenigen, die sich am Straßenverkehr beteiligen, mit hinreichender Ausführlichkeit zuvor deutlich gemacht wurden. All diese und zahlreiche weitere Verantwortlichkeiten sind Teil des Verantwortungsnetzwerkes „Verantwortung im Straßenverkehr“, und über die beschriebenen Teilverantwortungsgegenstände wird bereits nachvollziehbar, dass wir jeweils ganz unterschiedliche Subjekte in unterschiedlichem Ausmaß dafür zur Verantwortungsübernahme ansprechen würden, dass es jeweils unterschiedliche Instanzen, Adressat\*innen und Normen sind, die zur Konkretisierung der jeweiligen Verantwortlichkeit zu definiert werden verlangen.

49 Gegenwärtig wird ein autonomes Fahrassistenzsystem, das nur in einem operationalen Sinne als sehr schwacher Verantwortungsakteur identifizierbar ist, die Subjektposition einer Verantwortlichkeit innerhalb des Verantwortungsnetzwerkes „Verantwortung im Straßenverkehr“ nicht besetzen können, da es immer potenziell qualifiziertere Verantwortungssubjekte gibt. Allerdings ist denkbar, es als Verantwortungsobjekt und als Adressat in eine oder mehrere der Verantwortlichkeiten dieses Verantwortungsnetzwerkes einzubinden. In dieser Weise kann Verantwortung im Bereich der Roboterethik, die sich mit artifiziellen Systemen als Wertträger\*innen befasst, letztlich alle denkbaren Maschinen in etwaige Verantwortungskonstellationen integrieren. Interessanter – aus Sicht der Verantwortungstheoretikerin – wird das Ganze noch, sobald wir nicht nur operationale, sondern auch eindeutig funktional verantwortliche artifizielle Systeme entwickelt haben, die ggf. auch die Subjektposition innerhalb einer Verantwortlichkeit zu besetzen befähigt sind und in einer fernen posthumanistischen Zukunft vielleicht sogar die Rolle der Instanz übernehmen.

S. 30

- HFR 03/2016 S. 21 -

## 50 E. Fazit und Ausblick

Vor dem Hintergrund von Wallachs und Allens Ansatz einer funktionalen Äquivalenz von Kompetenzen und kombiniert mit einem algorithmischen Strukturschema ließe sich nun den Positionen eines Anthro-, Patho-, Bio- und Physiozentrismus eine weitere Sicht zur Lokalisierung von Phänomenen im moralischen Universum hinzufügen, die all die Wesen mit einem Eigenwert bemisst, die lernfähig sind – vielleicht einen Mathenzentrismus (vgl. Fußnote 48). Lernfähigkeit bedeutet mindestens eine Programmierung durch nicht-determinierte (nicht-deterministische) Algorithmen. Solche Wesen befänden sich im oberen Bereich der Wallach-Allen'schen funktionalen Moralzuschreibung und hätten unter dieser Perspektive einen Eigenwert. Robotern, die vor allem auf der Grundlage determinierter (nicht-deterministischer) Algorithmen arbeiten und sich eher im Bereich operationaler Moralzuschreibung bewegen, wäre aus dieser Perspektive immerhin ein hoher instrumenteller Wert zuzuschreiben.

- 51 Liegen die nötigen Bedingungen für Verantwortungszuschreibung im klassischen Sinne nicht vor, lässt sich nach dem Ansatz funktionaler Äquivalenz Verantwortung bislang nur als juristische, soziale, politische und wirtschaftliche operationale (und ggf. auch funktionale) Verantwortung denken, nicht aber als moralische operationale oder funktionale Verantwortung. Es tragen immer noch die Designer\*innen und Nutzer\*innen – bzw. alle beteiligten Menschen – die erste Verantwortung, die sie wohl auch nie gänzlich abgeben – zumindest so lange nicht, wie wir in Menschen die einzigen genuinen moralischen Akteur\*innen sehen, die über alle vier Autonomietypen nach Darwall verfügen. Sollten wir irgendwann von funktionaler Verantwortung bei einigen sehr komplexen Maschinen sprechen können, wäre denkbar, dass Menschen zu diesen immer noch in einem ähnlichen Verhältnis stehen, wie Eltern zu ihren fast erwachsenen Kindern. In Unfallsituationen könnten solche außergewöhnlichen funktionalen Verantwortungssubjekte ihre menschlichen ‚Eltern‘ in Sachen Verantwortungszuschreibung zum Teil entlasten – wenn ihnen auch nicht die Verantwortung gänzlich abnehmen. Gegenwärtig lassen sich artifizielle Systeme immerhin bereits eindeutig als Verantwortungsobjekte und -adressat\*innen in Verantwortungnetzwerke einbinden.

*Zitierempfehlung:* Janina Sombetzki, HFR 2016, S. 10 ff.