

März 2018

6. Jahrg.

84364

Seite 1–56

InTeR

Zeitschrift zum Innovations- und Technikrecht

1

Herausgegeben von

Jürgen Ensthaler

Stefan Müller

Dagmar Gesmann-

Nuissl

Herausgeberbeirat

Wilhelm-Albr. Achilles

Hans-Jürgen Ahrens

Udo di Fabio

Lars Funk

Thomas Klindt

Roman Reiss

Franz Jürgen Säcker

Klaus Schülke

Christian Steinberger

Walther C. Zimmerli

Klaus J. Zink

Schriftleitung

Lehrstuhl für

Wirtschafts-,

Unternehmens- und

Technikrecht an der

Technischen

Universität Berlin

In Verbindung mit

VDI – Verein Deutscher Ingenieure e. V.

Dr. Martin Sebastian Haase

- 1 **Immaterialgüterrechtlicher Schutz von dreidimensionalen Formen**

Patente im Automatisierten Fahren:

- 2 **InTeRview mit Thorsten Haslinde**

RA Johannes Groß

- 4 **AGB 4.0: Allgemeine Geschäftsbedingungen im Rahmen autonomer Vertragsschlüsse**

RA Fritz-Ulli Pieper, LL.M.

- 9 **Künstliche Intelligenz: Im Spannungsfeld von Recht und Technik**

Dr. Tim Jülicher und Charlotte Röttgen

- 15 **Bots im Kontext von Wirtschaftsrecht und Cybercrime**

Dipl.-Jur. Univ. Gernot Kirchner

- 19 **Big Data Management: Die Haftung des Big Data-Anwenders für Datenfehler (Vertragsrecht – Teil 1)**

*Dr. Dennis-Kenji Kipker, Dipl.-Ing. Andreas Harner,
Dipl.-Ing. Sven Müller*

- 24 **Der Mensch an der Schnittstelle zur Technik**

Dr. Janina Loh

- 29 **Verantwortung und Roboterethik – ein kleiner Überblick – Teil 2**

- 36 **Rezension**

Prof. Dr. Dagmar Gesmann-Nuissl

- 36 **Rechtsprechungsreport Innovations- und Technikrecht**

- 55 **InTeRessantes**



Abbildung 4: Haben Sie noch weitere Hinweise, Ergänzungen oder möchten Sie eine neue Einmeldung vornehmen? Auch das im IT-Security Navigator möglich!

Durch die damit geleistete Forschungsarbeit ist es erstmals möglich, nicht nur eine Vielzahl rechtlicher Vorgaben zur IT-Sicherheit vollumfassend und einfach systematisiert darzustellen, sondern mit Hilfe der unmittelbar stattfindenden Konkretisierung durch einschlägige technische Normen und Spezifikationen dem Anwender eine sofortige und zuverlässige Erst-Entscheidungshilfe zur Verfügung zu stellen, um für ihn möglicherweise verpflichtende IT-Security-Maßnahmen auf angemessene Weise zu implementieren. Das neu geschaffene Werkzeug steht im Internet jedem kostenfrei zur Verfügung. Zudem wird es laufend mit neuen Features ausgestattet, wie beispielsweise zusätzlichen Such- und Filteroptionen, der Anzeige weiterer Informationen zu den Normen und Spezifikationen auf Anforderung des Nutzers und einer grafischen Darstellung der Ergebnisse (z. B. zur Aktivität in den Normungsgremien/Domänen als Tortendiagramm und als Heat Map; daneben Statistiken zur Zahl der Referenzierungen in Gesetzen und Standards). Da die Praxistauglichkeit der Plattform im Vordergrund steht, erhält auch der Anwender die Möglichkeit zur Mitwirkung, indem er selbst neue Gesetze sowie Normen und Spezifikationen vorschlagen kann, die nach einer Prüfung in die Datenbank implementiert und dort miteinander vernetzt werden. In Zukunft wird der IT-Security

Navigator um die mobile App „Wizard“ ergänzt werden, die zusätzlich zur Suche nach Rechtsvorschriften oder Normen und Standards einen komplett strukturierten Entscheidungsbaum beinhaltet, der dem Anwender konkrete Vorschläge macht, welchen IT-Compliance-Vorgaben er genügen sollte und wie diese für ihn zu realisieren sind.

B. Fazit und Ausblick

Der Mensch steht an der Schnittstelle zur Technik, indem er unmittelbar mit der Umsetzung zeitgemäßer IT-Sicherheits- und auch Datenschutzvorgaben betraut ist. Hierfür ist ein interdisziplinäres Wissen erforderlich, das sowohl die rechtliche wie auch die technische Betrachtung des Schutzes im digitalen Arbeitsraum rund um KRITIS wie auch Industrie 4.0 einbezieht. In der Praxis aber fehlt ein solches Schnittstellenwissen oftmals; insbesondere KMU sind mit der Realisierung effektiver Cybersicherheit deshalb häufig überfordert. Mit der Forschung zum IT-Security Navigator wird aus diesem Grunde das Ziel verfolgt, vor allem das in kleinen wie mittelständischen Unternehmen tätige IT-Personal in seiner täglichen Arbeitspraxis zu entlasten, indem ihm einerseits ein effektives und gleichermaßen effizientes Werkzeug zur Bestimmung möglicher rechtlicher Pflichten im Bereich IT-Sicherheit, andererseits zur Konkretisierung der damit verbundenen technischen Vorgaben, zur Seite gestellt wird. Gerade für den mit der Umsetzung gesetzlicher Bestimmungen konfrontierten Endadressaten, der möglicherweise nicht immer über den unmittelbaren Zugriff auf juristische Ressourcen sowie auf Normen und Spezifikationen verfügt, ist eine solche einfache, schnelle wie auch verlässliche Erstorientierung wichtig, um die rechtlichen und technischen Maßstäbe seiner Arbeit zeitsparend und kostengünstig definieren zu können. Verschiedene Rückmeldungen aus den entsprechenden Fachkreisen bestätigen dies. Der IT-Security Navigator leistet damit einen bedeutsamen Beitrag zur Verbesserung der flächendeckenden IT-Sicherheit in Deutschland.

Dr. Janina Loh*

Verantwortung und Roboterethik – ein kleiner Überblick – Teil 2

III. 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 moralischer Rechte, oder inwiefern ihnen ein moralischer Wert zukommt. Die anderen interessieren sich dafür, ob und gegebenenfalls inwiefern Roboter sogar „moral agents“ sein könnten, also aktiv Träger moralischer Pflichten bzw. moralische Handlungsobjekte.¹ Beide Arbeitsbereiche ergänzen einander und Verantwortung ist in beide jeweils unterschiedlich einzubinden, wie in IV.1. und IV.2. 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 etwa 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.

Einer ganzen Reihe von Wesen und Dingen wie zum Beispiel Tieren, Pflanzen, aber auch Gegenständen wie dem

* Teil 1 des Beitrages ist in der InTeR 4/2017 erschienen. Dieser Beitrag führt als Teil 2 die Gliederung aus Teil 1 fort und schließt unmittelbar an Punkt II. des 1. Teils an. Mehr über die Autorin erfahren Sie auf Seite III.

1 Floridi/Sanders, On the Morality of Artificial Agents, in: Mindsand Machines, 14/2004, 349–379, 349.

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 Eigenwert sondern nur ein hoher instrumenteller Wert beigemessen wird. Als moralisches Handlungssubjekt hat man zugleich einen Platz im Kreis der Wertträger – dies gilt allerdings nicht umgekehrt. Lebewesen und Gegenständen kann man abhängig von der Perspektive einen moralischen Wert zuschreiben; eine anthropozentrische Position argumentiert beispielsweise dafür, dass nur dem Menschen ein Eigenwert zukommt, weitere Ansätze stellen der Patho-, der Bio- und der Physiozentrismus dar.²

Innerhalb des Arbeitsbereichs zu Robotern als Wertträgern wird das menschliche Verhalten gegenüber artifiziellen Systemen in den Blick genommen. Hier geht es darum, wie mit Robotern umzugehen ist und inwiefern ihnen (gegebenenfalls analog zu Tieren und 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 etwa bei der Formulierung von Ethikkodizes in Unternehmen,³ die Frage, inwiefern Beziehungen zu und mit Robotern denkbar und wünschenswert sind,⁴ inwiefern man Roboter ‚versklaven‘ kann⁵ und wie der Einsatz von artifiziellen Systemen zu Therapie Zwecken zu beurteilen ist.⁶ Innerhalb dieses Arbeitsbereichs verbleibt die moralische Kompetenz und Kompetenzkompetenz bei den menschlichen Designerinnen und Designern (und unter Umständen auch Nutzerinnen und Nutzern) 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 zugeschrieben 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 Urteils kraft, vgl. Abschnitt II.6.) komme ich unter dem Schlagwort der Verantwortungsnetzwerke in Abschnitt IV.2. zu sprechen. Verantwortung wird in der Mensch-Maschine-Interaktion häufig, wenn auch nicht immer, als rechtliche Kategorie diskutiert.⁷

Innerhalb des Arbeitsfeldes 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 Freiheit⁸ als Bedingung für moralisches Handeln, befassen sich andere eher mit kognitiven Kompetenzen⁹ (Denken, Verstehen, Geist, Intelligenz, Bewusstsein, Wahrnehmung und Kommunikation) und wieder andere mit Empathie und Emotionen.¹⁰

Beiden Arbeitsfeldern innerhalb der Roboterethik liegt die Frage zugrunde, was Moral bzw. was Ethik ist und wie moralische Urteile gefällt werden.¹¹ Auch hier lassen sich (allerdings nicht in diesem Artikel) verschiedene Positionen unterscheiden; in einem ersten Schritt könnte man vor-

schlagen, 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 Foots klassischem Gedankenexperiment der Trolley Cases den Fall von „driverless‘ train systems“,¹² in denen in London, Paris und Kopenhagen bereits seit Mitte der 1960er Jahre Menschen nur als Fahr-

- 2 Vgl. *Krebs*, Naturethik im Überblick, in: *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).
- 3 Vgl. *May*, Robotik und Arbeitsschutzrecht, in: *Hilgendorf* (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, 99-118.
- 4 Vgl. *Levy*, The Ethics of Robot Prostitutes, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 223-231; *Scheutz*, The Inherent Danger of Unidirectional Emotional Bonds between Humans and Social Robots, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 205-221; *Whitby*, Do You Want a Robot Lover? The Ethics of Caring Technologies, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 233-247.
- 5 Vgl. *Petersen*, Designing People to Serve, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 283-298; *Petersen*, The Ethics of Robot Servitude, in: in *Journal of Experimental & Theoretical Artificial Intelligence*, 19/2007, 43-54.
- 6 Vgl. *Boeker*, Robotik in Betreuung und Gesundheitsvorsorge, in: *Technikfolgenabschätzung, Theorie und Praxis*, 22/2013, 62-64; *Misselfhorn/Pompe/Stapleton*, Ethical Considerations Regarding the Use of Social Robots in the Fourth Age, in: *Geropsych* 26/2013, 121-133; *Borenstein/Pearson*, Robot Caregivers, Ethical Issues Across the Human Lifespan, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 251-265; *Sharkey/Sahrkey*, The Rights and Wrongs of Robot Care, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 267-282; *Datteri/Tamburrini*, Ethical Reflections on Health Care Robotics, in: *Capurro/Nagenborg* (Hrsg.), *Ethics and Robotics*, 2009, 35-47.
- 7 Vgl. *Lokhorst/van den Hoven*, Responsibility for Military Robots, in: *Lin/Abney/Bekey* (Hrsg.), *Robot Ethics, The Ethical and Social Implications of Robotics*, 2012, 145-156; *Maring*, Mensch-Maschine-Interaktion, Steuerbarkeit – Verantwortbarkeit, in: *Hubig/Koslowski* (Hrsg.), *Maschinen, die unsere Brüder warden, Mensch-Maschine-Interaktion in hybriden Systemen*, 2008, 113-129; *Asaro*, Robots and Responsibility from a Legal Perspective, in: *Proceedings of the IEEE Conference in Robotics and Automation, Workshop on Roboethics*, 2007, 20-24; *Floridi/Sanders*, On the Morality of Artificial Agents, in: *Minds and Machines* 14/2004, 349-379.
- 8 Vgl. *Dorffner*, Rationalität, Emotionalität und Körperlichkeit, Können Maschinen Menschen verstehen – und umgekehrt?, in: *Schmidinger/Sedmak* (Hrsg.), *Der Mensch – ein freies Wesen? Autonomie – Persönlichkeit – Verantwortung*, 2005, 102-112.
- 9 Vgl. *Clark*, Artificial Intelligence and the Many Faces of Reason, in: *Stich/Warfield* (Hrsg.), *The Blackwell Guide to Philosophy of Mind*, 2003, 309-321; *Clark*, Towards a Cognitive Robotics, in: *Adaptive Behavior*, 7/1999, 5-16; *Dennett*, Cognitive Wheels, The Frame Problem of AI, in: *Ford/Glymour/Hayes* (Hrsg.), *Thinking About Android Epistemology*, 2006, 147-169; *Dennett*, *Brainchildren, Essays on Designing Mind*, 1998.
- 10 Vgl. *Coeckelbergh*, Moral Appearances, Emotions, Robots, and Human Morality, in: *International Journal of Social Robotics*, 1/2009, 217-221.
- 11 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.
- 12 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 14.

gäste anzutreffen sind.¹³ 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 Unfälle im Zuginnern in Kauf genommen werden müssen.

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 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 beziehungsweise 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 IV.1. ausführlich eingegangen wird – eine Version der schwachen KI-These.

IV. 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. *Wallach* und *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 (IV.1. mit Bezugnahme insbesondere auf II.6.) und Roboter als moral patients bzw. Wertträger (IV.2. unter Rekurs insbesondere auf II.1.–II.5.) – in den Blick.

1. Roboter als Handlungssubjekte und die Bedingungen für die Möglichkeit zur Verantwortungsübernahme – *Wallachs* und *Allens* Ansatz funktionaler Äquivalenz¹⁴

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“¹⁵). Menschen gelten als moralische Akteure im genuinen Sinne, allerdings sind einige Maschinen – beispielsweise 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 zum Beispiel ein Hammer, und dennoch verbleiben sie immer noch „totally within the control of [the] tool’s designers and users“.¹⁶ In diesem Sinne sind operationale moralische Akteure „direct extensions of their designers’ values“.¹⁷ Nur besondere artifizielle Systeme haben bereits den Status funktionaler moralischer Akteursfähigkeit – so wie das medizinische ethische Expertensystem *MedEthEx*.¹⁸ 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“.¹⁹ *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 understood the meaning of symbols without actually having what one would consider to be human understanding.“²⁰

13 Hieran schließt sich die Debatte um autonome Fahrassistenzsysteme; vgl. *Loh/Loh*, *Autonomy and Responsibility in Hybrid Systems – The Example of Autonomous Cars*, in: *Lin/Abney/Jenkins* (Hrsg.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, 2017, 35–50; *Hevelke/Nida-Rümelin*, *Intelligente Autos im Dilemma*, in: *Spektrum der Wissenschaft*, 10/2015, 82–85; *Maurer/Gerdes/Lenz/Winner* (Hrsg.), *Autonomes Fahren, Technische, rechtliche und gesellschaftliche Aspekte*, 2015; *Both/Weber*, *Hands-Free Driving? Automatisiertes Fahren und Mensch-Maschine Interaktion*, in: *Hilgendorf* (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, 171–187; *Hötitzsch/May*, *Rechtliche Problemfelder beim Einsatz automatisierter Systeme im Straßenverkehr*, in: *Hilgendorf* (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, 189–210; *Knoll*, *Prädikative Fahrassistenzsysteme – Bevormundung des Fahrers oder realer Kundennutzen?*, in: *Hubig/Koslowski* (Hrsg.), *Maschinen, die unsere Brüder werden. Mensch-Maschine-Interaktion in hybriden Systemen*, 2008, 159–171.

14 Die Überlegungen in diesem Abschnitt können in zahlreichen meiner Artikel nachgelesen werden – am aktuellsten vgl. *Loh/Loh*, *Autonomy and Responsibility in Hybrid Systems – The Example of Autonomous Cars*, in: *Lin/Abney/Jenkins* (Hrsg.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, 2017, 35–50.

15 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 25.

16 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 26.

17 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 30.

18 *Anderson/Anderson/Armen*, *An Approach to Computing Ethics*, in: *Intelligent Systems*, IEEE 4/2006, 2–9.

19 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 9.

20 *Wallach/Allen*, *Moral Machines, Teaching Robots Right from Wrong*, 2009, 69.

Mit ihrem Ansatz funktionaler Äquivalenz beschreiben *Wallach* und *Allen* eine Version der schwachen KI-These,²¹ der an der Simulation spezifischer Kompetenzen in künstlichen Systemen gelegen ist und nicht daran, Maschinen tatsächlich im genuinen Sinne des Wortes mit Intelligenz, Bewusstsein und Autonomie zu konstruieren (starke KI, irrtümlich auf *Turing* zurückgeführt²²). *Wallach* und *Allen* zufolge stellt eine Annahme starker KI und der daran geknüpften Kompetenzen keine notwendige Bedingung für künstliche 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.²³ 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.²⁴ Die Frage, inwiefern künstliche 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 künstlichen Systems wird nur insofern in den Blick genommen, als es als Voraussetzung dafür zu deuten ist, Verantwortung zuzuschreiben.

Obwohl *Wallach* und *Allen* den Übergang von operationaler über funktionale 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 künstliches System ein funktionales Äquivalent zu der genuin menschlichen Fähigkeit „second order volitions“²⁵ zu bilden entwickeln könnte bzw. zu der Fähigkeit, als „self-authenticating sources of valid claims“²⁶ 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.²⁷ 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 künstliche Akteure erreichbar, insofern die besagten Gründe funktional äquivalent beispielsweise 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 künstlichen Systems ohne externe Stimuli zu ändern.

Auf der computationalen Ebene ließe sich die Fähigkeit, autonom interne Zustände zu ändern, beispielsweise in Form eines algorithmischen Strukturschemas näher beschreiben,²⁸ 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 Abschnittes darum, zu erklären, inwiefern Roboter als moralische Handlungs-subjekte (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 dann eindeutig mit künstlichen Systemen betreten, die vornehmlich durch determinierte (aber nicht-deterministische) Algorithmen strukturiert sind. Und schließlich könnten solche seltenen Fälle künstlicher Systeme, die vornehmlich auf der Grundlage nicht-determinierter (und also nicht-deterministischer) Algorithmen operieren, in der funktionalen Sphäre lokalisiert werden.

Schauen wir uns ein paar Beispiele an: Das künstliche System *Kismet*, das *Wallach* und *Allen* als operationalen moralischen Akteur interpretieren, verfügt in einem äußeren rudimentären Sinne über Kommunikationsfähigkeit, insofern es nur wenige Laute und Worte hervorbringen kann. Urteilskraft – sollte man gewillt sein, *Kismet*s 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. II.6.) 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 Heraus-

21 *Searle*, *Minds, brains and programs*, in: *Behavioral and Brain Sciences*, 3/3/1980, 417-157.

22 *Turing*, *Computing Machinery and Intelligence*, in: *Mind*, 59/236/1950, 433-460.

23 *Nagel*, *What is it like to be a bat?*, in: *The Philosophical Review*, LXXXIII, 1974, 435-450.

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

25 *Frankfurt*, *Freedom of the Will and the Concept of a Person*, in: *Journal of Philosophy*, 68/1/1971, 5-20.

26 *Rawls*, *Justice as fairness, A restatement*, 2001, 23.

27 *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 *Wallach-Allen*'schen Modells um *Darwalls* Autonomie-Typen.

28 Vgl. *Sombetzki*, *Roboterethik – ein kritischer Überblick*, in: *Matthias Maring* (Hrsg.), *Zur Zukunft der Bereichsethiken, Herausforderungen durch die Ökonomisierung der Welt*, 2016, 355-379.

forderung, 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 Persönlichkeit 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 Nutzerinnen und 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.

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 oder Autonomie ist aufgrund eines „unsupervised learning algorithm“²⁹ 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 immerhin 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.

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.

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 artifizieller (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 (oder 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.

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 unter anderem *Brooks*, der als einer der ersten das Zusammenwirken von artifiziellem System und Umwelt als Bedingung für die Entwicklung von Vermögen und Kompetenzen betrachtete und von dieser Annahme ausgehend das Feld der „behavior-based robotics“ begründete.³⁰ Zahlreiche Projekte, die sich an dem Ansatz verkörperten menschlichen Lernens orientieren – wie bspw. die Lernplattformen iCub, Myon, Cb², 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 im Bereich nicht-determinierter (nicht-deterministischer) Algorithmen aufhalten, die eigenständig lernen können (evolutionäre, genetische Algorithmen).

29 *Brooks/Breazeal/Marjanović/Scasselatti/Williamson*, The Cog Project, Building a Humanoid Robot, in: Nehaniv (Hrsg.), *Computation for Metaphors, Analogy, and Agents*, 1999, 52-87, 70.

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

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.

2. Roboter als Wertträger 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 – zuzugestehen ist. Hierzu wird eine Rückbesinnung auf die in den Abschnitten II.1. bis II.5. vorgestellten fünf Relationselemente der Verantwortung hilfreich sein.

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 psycho-motivationalen Verfasstheit des fraglichen Verantwortungssubjekts einhergeht; vgl. die Einleitung zu II.) insofern ein stark individualistisches, als wir immer ein Subjekt benötigen, das als Verantwortungsträger fungiert (vgl. II.1.). 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 in den vergangenen Jahren behelfsmäßige Begrifflichkeiten zu entwickeln versucht, die ohne eine Bestimmung dieses Relationselements auskommen.³¹ Ich bin skeptisch, dass damit bezüglich der eigentlichen Aufgabe, die das Verantwortungskonzept hat, nämlich – wie in der Einleitung (Abschnitt I.) 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 beziehungsweise es mit einer Wortblase wie dem „System“ nur oberflächlich zu bestimmen.

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 IV.1. als operational verantwortliche artifizielle 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 bedenkenwert ist, dass ihm ein instrumenteller Wert zugeschrieben wird – aber als moralischen Akteur in einem signifikanten (das heißt 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.

Für solche und vergleichbare Kontexte möchte ich den Begriff des Verantwortungsnetzwerkes von *Christian Neuhäuser*³² übernehmen und spezifizieren.³³ Die diesen Überlegungen zugrundeliegende These lautet, dass wir all diejenigen 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 Verantwortungsnetzwerk trägt der Tatsache explizit Rechnung, dass sich innerhalb einer Verantwortungskonstellation 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 Adressatinnen und Adressaten (also den Grund des Vorhandenseins dieser Verantwortlichkeit).³⁴

31 Vgl. *Wilhelms*, Systemverantwortung, in: Heidbrink/Langbehn/Loh (Hrsg.), Handbuch Verantwortung, 2017, 501-524.

32 Vgl. *Neuhäuser*, Roboter und moralische Verantwortung, in: Hilgen-dorf (Hrsg.), Robotik im Kontext von Recht und Moral, 2014, 269-286.

33 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, 35-50 auf *Wulf Lohs* Idee hin zum Einsatz bringe.

34 Zu diesem Beispiel vgl. ausführlich *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 Verantwortungsnetzwerkes vermuten. Danach wird der zu ersetzende Schaden je nach Ausmaß des Verschuldens gequotelt; in diese Richtung geht auch die Möglichkeit des in Anspruch Genommenen, andere Schuldner in Regress zu nehmen, nachdem bereits Schadensersatz geleistet wurde (Bspw. kann bei einem Unfall, bei dem ein autonomes Fahrassistenzsystem beteiligt ist, der Geschädigte zunächst den Halter des Fahrzeugs in Anspruch nehmen. Dieser kann ggf. den Hersteller des Fahrzeugs in Regress nehmen, welcher sich wiederum an den Hersteller der Software wenden kann.). Ich danke *Rebecca Sieber* für diesen Hinweis.

Innerhalb des Verantwortungszusammenhangs „Verantwortung im Straßenverkehr“ gehören die autonomen Fahrerassistenzsysteme ebenso dazu wie die menschliche Fahrerin bzw. der menschliche Fahrer (selbst dann, wenn sie oder er nicht aktiv am Fahrprozess beteiligt ist), die Besitzerinnen und Besitzer, die Vertreterinnen und Vertreter, die Programmiererinnen und Programmierer sowie die Designerinnen und Designer, aber auch die Öffentlichkeit, Juristinnen und Juristen, Fahrlehrerinnen und -lehrer und alle am Straßenverkehr Beteiligten. Verantwortungszusammenhänge haben häufig ungewöhnliche Ausmaße und bündeln in sich unterschiedliche Verantwortungsobjekte. Von Verantwortungszusammenhängen kann man dann sprechen, wenn man eigentlich – sehr schön zu veranschaulichen am Fall der Klimaverantwortung³⁵ – 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. besetzen manchmal mehrere Relationspositionen zugleich, sind einmal die Verantwortungsobjekte, in einem anderen Fall die Instanzen und wieder in einem anderen Fall das Objekt und vielleicht zugleich Adressatinnen bzw. Adressaten einer Verantwortlichkeit.

Es wäre äußerst schwierig, ein oder mehrere konkrete Verantwortungsobjekte für die Verantwortung im Straßenverkehr auszumachen, da diese viel zu umfassend ist, als dass eine Person oder eine geringe Anzahl Einzelner dafür Rede und Antwort stehen könnte. Als Verantwortungszusammenhang „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 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. Über die beschriebenen (und zahlreiche weitere) Teilverantwortungsgegenstände wird bereits nachvollziehbar, dass wir jeweils ganz unterschiedliche Subjekte in unterschiedlichem Ausmaß dafür zur Verantwortungsübernahme ansprechen würden, dass es jeweils unterschiedliche Instanzen, Adressatinnen bzw. Adressaten und Normen sind, die zur Konkretisierung der jeweiligen Verantwortlichkeit zu definiert werden verlangen.

Gegenwärtig wird ein autonomes Fahrerassistenzsystem, das nur in einem operationalen Sinne als sehr schwacher Verantwortungsakteur identifizierbar ist, die Subjektposition einer Verantwortlichkeit innerhalb des Verantwortungszusammenhangs

„Verantwortung im Straßenverkehr“ nicht besetzen können, da es immer potenziell qualifiziertere Verantwortungsobjekte gibt. Allerdings ist denkbar, es als Verantwortungsobjekt und als Adressat in eine oder mehrere der Verantwortlichkeiten dieses Verantwortungszusammenhangs einzubinden. In dieser Weise kann Verantwortung im Bereich der Roboterethik, die sich mit artifiziellen Systemen als Wertträgern befasst, letztlich alle denkbaren Maschinen in etwaige Verantwortungskonstellationen integrieren.

V. 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 2). 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.

Liegen die nötigen Bedingungen für Verantwortlichkeitszuschreibung im klassischen Sinne nicht vor, lässt sich nach dem Ansatz funktionaler Äquivalenz Verantwortung bislang nur als juristische, soziale, politische und wirtschaftliche operationale (und gegebenenfalls auch funktionale) Verantwortung denken, nicht aber als moralische operationale oder funktionale Verantwortung. Es tragen immer noch die Designer und Nutzer – beziehungsweise 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 Akteure 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 Verantwortungsobjekte ihre menschlichen ‚Eltern‘ in Sachen Verantwortlichkeitszuschreibung zum Teil entlasten – wenn ihnen auch nicht die Verantwortung gänzlich abnehmen. Gegenwärtig lassen sich artifizielle Systeme immerhin bereits eindeutig als Verantwortungsobjekte und -adressaten in Verantwortungszusammenhängen einbinden.

³⁵ Zur Klimaverantwortung und den mit dieser Verantwortlichkeit einhergehenden Herausforderungen vgl. ausführlich *Sombetzki*, Verantwortung als Begriff, Fähigkeit, Aufgabe, Eine Drei-Ebenen-Analyse, 2014, Kapitel 13.