

## Themenessay

# Robotik in allen Lebenslagen? Zur zunehmenden Relevanz und Problematik von Robotern in verschiedenen gesellschaftlichen Bereichen

**Jannis Hergesell / Arne Maibaum / Martin Meister** (Hrsg.), Genese und Folgen der Pflegerobotik. Die Konstitution eines interdisziplinären Forschungsfeldes. Weinheim: Beltz 2020, 279 S., br., 39,95 €

**Janina Loh**, Roboterethik. Eine Einführung. Berlin: Suhrkamp 2019, 241 S., kt., 18,00 €

**Christiane Woopen / Marc Jannes** (Hrsg.) Roboter in der Gesellschaft. Technische Möglichkeiten und menschliche Verantwortung. Berlin: Springer 2019, 114 S., gb., 84,99 €

Besprochen von **Professor Dr. rer. nat. Michael Decker**: Karlsruher Institut für Technologie (KIT), Institut für Technikfolgenabschätzung und Systemanalyse (ITAS), E-Mail: [decker@kit.edu](mailto:decker@kit.edu)

<https://doi.org/10.1515/srsr-2022-0026>

**Schlüsselwörter:** Robotik, Autonome Systeme, Künstliche Intelligenz, Lehre

Roboter sind faszinierende Maschinen. Durch Science-Fiction Literatur und -Filme sind sie uns vertraut und die Robotikforschung muss mit der Herausforderung leben, dass die Performanz realer Robotersysteme gemeinhin noch weit hinter Commander Data aus Star Trek oder Marvin aus Per Anhalter durch die Galaxis zurückstehen. Unabhängig davon hat die Robotikforschung in den letzten Jahrzehnten große Fortschritte gemacht. Die Errungenschaften der Künstlichen-Intelligenz-Forschung (KI), die insbesondere für die Steuerung der Roboter und deren Lernfähigkeit relevant sind, haben hieran einen nicht geringen Anteil. Im digitalen Wandel, dessen Gestaltung als eine der gesellschaftlichen Herausforderungen unserer Zeit gilt, sind Roboter ein zentraler Bestandteil, weil sie über Manipulation Veränderungen in der Umwelt vornehmen können, und weil sie sich selbst fortbewegen und dabei in spezieller Weise Daten sammeln können. Insofern kann man sie als ein Front End der Digitalisierung betrachten.

Diese generelle Einschätzung teilen alle Schriften, die seitens der Soziologischen Revue für diese Rezension ausgesucht wurden. Das schließt die Hinweise zur Science-Fiction ein, wobei damit auch Erwartungsmanagement betrieben wird, um den Blick frei zu bekommen für aktuell realistische Anwendungen. Als Antwort auf die Frage, wo die Robotik heute steht, schreibt die Industrierobotik eine Erfolgsgeschichte. Die International Federation of Robotics (IfR) gibt jährlich Berichte heraus, die sich durch fortlaufend steigende Umsätze über mehrere Jahre auszeichnen, wobei 2019 die Anzahl der neu installierten Roboter um 12% auf 373.240 Einheiten im Wert von 13,8 Mrd. USD (ohne Software und Peripheriegeräte) sank. Dieser Rückgang, nach sechs Jahren Wachstum und Spitzenwerten in Folge, wird mit Marktherausforderungen der beiden Hauptbranchen Automobil und Elektrotechnik/Elektronik erklärt. Diese bilden zusammen die Hälfte des Marktes, während Metall und Maschinen, Kunststoffe und chemische Produkte, Lebensmittel und Getränke kleinere Anteile haben. Die IfR berichtet auch über Servicerobotik. Dieser Markt wuchs 2019 stark um 32% auf 11,2 Mrd. USD. Medizinische Robotik (hauptsächlich Chirurgie) macht 47% des gesamten Umsatzes aus, Logistik etwa 43%. Sowohl für Medizin- als auch für Logistikrobotik werden auch zukünftig deutliche Zuwächse erwartet. Geringer fallen die Zuwächse im Agrarwesen aus. Private Service- und Haushaltsroboter erzielen die größten Verkaufszahlen mit 23,2 Mio Robotern (+34% in 2019) hierzu zählen auch Staubsaugroboter.

Kritisch könnte man anmerken, dass es sich bei Chirurgie und Logistik letztlich auch um Produktion handelt, die allerdings nicht in Fabrikhallen stattfindet. Bei der Logistik ist das offensichtlich, in der Chirurgierobotik ist das daVinci System prägend: Es ermöglicht minimalinvasive Operationen und ist im Einsatz durch Chirurg:innen ferngesteuert. Beide Umgebungen, das Logistikzentrum und der Operationssaal, lassen sich dabei für den Einsatz und die optimale Funktion der Robotik gestalten. Anders stellt sich das bei Assistenzrobotern für ältere oder behinderte Menschen dar, auf die sich nationale Forschungsprojekte in vielen Ländern konzentrieren, weil ein zukünftiger Wachstumsmarkt gesehen wird. Damit ändern sich die Anforderungen an Serviceroboter immens. Zum einen müssen sie in völlig unterschiedlichen Umgebungen einsetzbar sein, denn sowohl Krankenhäuser als auch Pflegeheime sind alles andere als standardisiert. Vor allem sind damit aber Tätigkeiten verbunden, die durch Menschen an Menschen verrichtet werden. Pflege ist soziale Interaktion, die besondere Expertise mit soft skills wie Freundlichkeit, Zuvorkommenheit, Hilfsbereitschaft und Gesellschaftleisten kombiniert. Das Paradigma ist die Kooperation zwischen Mensch und Roboter, weil sich nicht alle Teil-Tätigkeiten technisch umsetzen lassen, Roboter somit nicht in der Lage sind, allein eine umfassende Pflege sicherzustellen. Der Mensch ist dabei manchmal der Pflegebedürftige, wenn man sich ein Einsatzszenario in der Privat-

wohnung vorstellt, manchmal ist es eine Mensch-Roboter-Mensch-Konstellation, wenn auch professionell Pflegende oder Angehörige mitwirken.

Sind Roboter dazu in der Lage? Und sollen wir wollen, dass sie in diesen Bereichen tätig werden? Diese und ähnliche Fragen stellt die Technikfolgenforschung, die hier als Referenzpunkt für die Perspektive herangezogen wird, mit der die drei Bücher gelesen wurden. Sie richtet das Augenmerk auf (nicht)intendierte, (un)erwünschte zukünftige Folgen, auf Chancen und Risiken, die mit dem Einsatz von Robotern verbunden sind, und auch auf mögliche Empfehlungen seitens der Autor:innen, wie zukünftig mit den gesellschaftlichen Veränderungen durch Robotik umgegangen werden könnte.

Dabei sind Roboter zunächst Maschinen. Technische Werkzeuge, die für Zwecke von menschlichen Akteuren eingesetzt werden. Oben genannte Einsatzbereiche spiegeln das wider, wenn Produktions- und allgemeiner Industrierobotik angesprochen sind und auch im Servicebereich mit Logistik- und Chirurgierobotik jedenfalls Kontexte adressiert sind, in denen sich die gesamte Umgebung auf den Einsatz von Robotern anpassen lässt. Das soll die technischen Errungenschaften in diesen Bereichen nicht schmälern. Gerade in der Industrierobotik ist unter dem Stichwort „Industrie 4.0“ ein wahrer Innovationsschub hin zu flexiblen Robotersystemen entstanden, die nicht mehr in Käfigen in hohen Taktraten Wiederholungstaten des Schweißpunktesetzens absolvieren, sondern in Kooperation mit den Werker:innen gemeinsam Fertigungstätigkeiten ausführen. Und dennoch sind diese Umgebungen für den Robotereinsatz optimiert: Alle Menschen in den Montagehallen sind über den Robotereinsatz aufgeklärt, haben Schulungen absolviert, Sicherheitsunterweisungen erfolgten, Bodenbelag, Beleuchtung usw., alles ist optimiert, sowohl für die Menschen als auch für die Roboter. Der nächste Schritt ist daher, Roboter auch in nicht für sie optimierten Umgebungen einzusetzen, etwa in häuslicher Umgebung, wo der individuellen Ausgestaltung fast keine Grenzen gesetzt sind, oder in der Öffentlichkeit, wenn wir etwa an selbst fahrende Autos denken, die in allen Städten und Dörfern, auf Landstraßen und Autobahnen einsetzbar sein müssen. Während also im ersten Fall eine gegenseitige Anpassung des Roboters mit seiner kompletten Arbeitsumgebung erfolgt, muss im zweiten Fall die Adaptivität an den konkreten Handlungskontext und die vor Ort zur Verfügung stehende Autonomie des Roboters beim Ausführen von Tätigkeiten ein ganz anderes Niveau erreichen. Autonomes Handeln und Lernfähigkeit von technischen Systemen lassen sich aber nur in konkreten Handlungszusammenhängen beurteilen. Hier gilt der Pflegesektor schon seit längerer Zeit als ein besonders prominentes und auch kontrovers diskutiertes Einsatzgebiet. Die Frage ist also vereinfacht gestellt, was können Roboter heute schon und wie stellen wir sicher, dass sie in den neuen Anwendungsbereichen auch so eingesetzt werden können, dass die damit verbundenen Vorteile die möglicherweise auch auftretenden Nachteile übertreffen.

## Satus Quo im interdisziplinären Diskurs

Hier ist offensichtlich eine multidisziplinäre Betrachtung gefragt, die den disziplinübergreifenden Austausch sucht, wie in „Roboter in der Gesellschaft“ (Woo-pen/Jannes). Die ersten drei Texte sind aus informationstechnischer Sicht geschrieben und beschreiben den Stand der Technik. Wenn *Alin Albu Schäffer* als einer der herausragenden Robotikforscher in Deutschland die Robotik vorstellt, wird die technische Herausforderung sofort offensichtlich. Einerseits geht es um Achsen, Getriebe und deren feinstskalige Ansteuerung. Hierbei klingt immer die klassische Definition à la VDI-Richtlinie mit, nach der „[e]in Roboter ein frei und wieder programmierbarer, multifunktionaler Manipulator mit mindestens drei unabhängigen Achsen, um Materialien, Teile, Werkzeuge oder spezielle Geräte auf programmierten, variablen Bahnen zu bewegen zur Erfüllung der verschiedensten Aufgaben“ ist. Dieser nüchternen Beschreibung steht die Faszination, die von modular aufgebauten humanoiden Robotern ausgeht, gegenüber. Sie sind bereits technisch so ausgelegt, dass sie nachgiebig reagieren, wenn sie auf Widerstände stoßen. „Ziel dieser Regelung ist es, den Roboter so zu steuern, dass er sich unter Einwirkung externer Störungen in jede Raumrichtung wie ein Feder-Masse-Dämpfungs-System verhält, bei dem Steifigkeit und Dämpfung innerhalb bestimmter Grenzen beliebig zugewiesen werden können“ (Albu Schäffer in Woo-pen/Jannes: 6). Damit soll technisch quasi ausgeschlossen werden, dass der Roboter einem Menschen eine ernsthafte Verletzung zufügen kann. Das ist eine ganz entscheidende Qualität, wenn es um eine echte Kooperation mit Menschen geht, sozusagen um das Arbeiten Hand in Hand. Hier kann die Abschirmung des Roboters durch einen stabilen Käfig, wie sie sich in der Produktion an verschiedenen Stellen als zielführend erweist, wenn höchste Fertigungsgeschwindigkeiten etwa beim Schweißen von Blechteilen erreicht werden sollen, keine Lösung sein. Die neuen Möglichkeiten der Mensch-Robotik-Kooperation „revolutioniert“ (ebd. 13) die industrielle Fertigung, wenn es um individuelle Fertigungsschritte geht.

Auch wenn die humanoiden Roboter ihren Science-Fiction-Kollegen noch hinterherhinken, so haben sie doch einen Forschungsstand erreicht, der die Biomechanik in beide Richtungen befruchtet. Die Bewegungen humanoider Roboter sind durch Biomechanik inspiriert, das wird niemand überraschen, auch wenn man durch eine Beschreibung wie Feder-Masse-Dämpfung nicht unmittelbar an Muskelspannung in einem menschlichen Körper denkt. Die humanoide Robotik trägt heute umgekehrt zur Hypothesengenerierung in der Biomechanikforschung und letztlich den Neurowissenschaften bei, wenn beispielsweise die „neuronalen Strukturen der Bewegungssteuerung des Rückenmarks und des Hirnstamms“ (ebd. 11) aufgeklärt werden. Technisch betrachtet, so resümiert *Albu-Schäffer*, sind Roboter im ferngesteuerten Modus heute schon „in fast jeder Umgebung“ (ebd. 13)

einsetzbar, mittelfristig ist der teil-autonome Einsatz das Ziel, der unmittelbar auch mit der KI-gestützten Entscheidungsfindung und dem technischen Lernen zusammenhängt.

Eine effiziente Mensch-Maschine-Kooperation in unterschiedlichen Handlungszusammenhängen ist ohne eine Anpassung des technischen Systems an den Menschen und die Umgebung schwer vorstellbar. In der individuellen Fertigung muss der Roboter mit mehreren Menschen kooperieren und sich eigenständig auf eine:n ihm unbekannte:n Werker:in anpassen können, wenn es beispielsweise um deren Körpergröße, Armreichweite, typische Bewegungsabläufe geht (Steil in Wopen/Jannes: 17). Die KI umfasst auch das maschinelle Lernen und dieses wird sowohl unter- als auch überschätzt.

Die Überschätzung liegt uns Menschen nahe: Denn wenn wir einen humanoiden Roboter sehen, der mit natürlicher Sprache spricht, und sich in der Umgebung gut bewegen kann, dann neigen wir unwillkürlich dazu, die Performanz des Roboters deutlich zu überschätzen. Wir anthropomorphisieren den Roboter und vermuten dann höhere, menschliche kognitive Leistungen. „Es fehlt dazu nicht nur an Methoden, sondern auch an grundlegendem Verständnis davon, wie solche Intelligenz in einem vollständig lernenden kognitiven System zustande kommt. Entscheidungsfähigkeit, Metakognition, generelle Intelligenz oder Bewusstsein, wie wir sie von Menschen kennen, sind in der überschaubaren Zukunft nicht in Robotern realisierbar“ (ebd. 26).

Die Unterschätzung liegt in der Reichweite dessen, was bei einer personalisierten Anpassung an einen Menschen als Kooperationspartner:in passiert. Roboter müssen möglichst viele Daten sammeln, um sich auf den Menschen einzustellen. Umgekehrt erstellen sie dabei ein sehr detailliertes datengestütztes Bild des Menschen. „Speziell die Assistenzrobotik ist janusköpfig – Hilfe und Überwachung sind ihre beiden Gesichter“ (ebd. 31). Tatsächlich ist das in der Assistenz angelegt. Wenn der Roboter idealerweise „seinem“ Mensch assistiert, lassen sich dabei optimale Handlungsabläufe entwickeln. Diese stellen einen Qualitätsgewinn gegenüber bisherigen Handlungsabläufen dar, sie werden aber umgekehrt zum neuen Benchmark für diese Handlungen, verbunden mit entsprechendem Leistungsdruck. Die Personalisierung des Roboters bringt unweigerlich die detaillierte Analyse des Menschen mit sich. Tägliche Schwankungen in der Leistungsfähigkeit, körperbezogene Daten, Pausenhäufigkeit etc. werden im zeitlichen Verlauf aufgenommen. Der Big-Data-Vorteil liegt dabei immer beim Roboter bzw. dessen Betreiber:in. Dadurch kann es in der Kooperation mit dem Menschen zu Veränderungen kommen, wenn das technische System in seinen datengestützten Analysen den menschlichen Akteur übertrifft. Beispielhaft ist das bei der medizinischen Beurteilung von Röntgenbildern der Fall. KI-gestützte Algorithmen erreichen hier bereits bessere Ergebnisse als menschliche Ärzt:innen. Sie sind aber nur als Assis-

tenz für diese vorgesehen, die verantwortliche Entscheidung bleibt beim Menschen. Das bedeutet aber eine schleichende „Beweislastumkehr“. Der Mensch muss argumentieren, wenn er der technischen Analyse widersprechen möchte. Ist es hier weiterhin gerechtfertigt, dass der Mensch die volle Verantwortung für die Entscheidung trägt? *Steil* (in Woopen/Jannes) schlägt vor, den Fokus der ethischen Reflexion hierauf zu legen.

In der Kooperation zwischen Mensch und Maschine trifft die Autonomie des Menschen auf eine technisch autonome Robotik. Beim autonomen Fahren werden verschiedene Stufen des Zusammenspiels zwischen Fahrer:in und dem autonomen Fahrzeug beschrieben. Diese decken das Spektrum von die Fahrer:in kontrolliert zu 100% bis zum vollautonomen Fahren ab, bei dem das Fahrzeug zu 100% steuert. Dazwischen sind verschiedene Grade der anteiligen Autonomie zwischen Fahrer:in und Fahrzeug definiert. Insbesondere für Pilot:innen aber auch in Bezug auf verschiedene technische Assistenzsysteme in Kraftfahrzeugen berichtet *Johannes Weyer*, dass Pilot:innen „ein sehr hohes Vertrauen in die neuartige Form der hybriden Zusammenarbeit von Mensch und autonomer Technik an Bord moderner Verkehrsflugzeuge“ (in Woopen/Jannes: 99) haben. In Bezug auf einen möglichen Kontrollverlust in dieser Kooperation ist festzustellen, dass „das Zusammenspiel von Mensch und autonomer Technik nicht zwangsläufig zu einer Wahrnehmung von Kontrollverlust auf Seiten des menschlichen Bedieners“ führt (ebd. 103).

Unterschiedliche Arbeitsteilungen zwischen Mensch und Roboter sind auch aus arbeitswissenschaftlicher Perspektive interessant, denn faktisch ist bei 100% autonomem Fahren der menschliche Fahrer komplett ersetzt. Der digitale Wandel wird Arbeitsprozesse weitreichend verändern. Frey und Osborne (2013) haben mit ihrer Studie, welche Berufe durch Automatisierung digitalisiert werden können, großes mediales Aufsehen erregt. Sie kamen auf das Ergebnis, dass knapp die Hälfte aller Berufe auf dem US-Arbeitsmarkt innerhalb von zwei Dekaden potentiell durch Automatisierung bedroht ist. Welche Auswirkungen damit mittel- und langfristig auf dem Arbeitsmarkt verbunden sind, wird kontrovers diskutiert. In der aktuellen Diskussion wird überwiegend die Auffassung vertreten, dass sich makroökonomisch gesehen die Effekte ausgleichen, so dass geringe bis keine Auswirkungen zu erwarten sind. Das heißt, dass durch den Einsatz von Robotern zwar Tätigkeiten, die bisher von Menschen ausgeführt wurden, wegfallen, dass aber gleichzeitig andere Tätigkeiten ausgeführt werden müssen, die etwa durch die Integration des Roboters in die Fertigungsabläufe oder durch die Wartung, Einrichtung, Umrüstung der Roboter notwendig werden. Diese Einschätzung wird durch Erkenntnisse aus der ersten Automatisierungswelle in den 1980er Jahren gestützt. Auf individueller Ebene ist es aber durchaus so, dass ganze Berufe ersetzt werden können, sodass Unternehmen geraten wird, entsprechende Umschu-

lungsmaßnahmen vorzusehen. In Kombination mit aktuell weiterhin bestehender Nachfrage auf dem Arbeitsmarkt sind die Erfolgchancen nach einer Umschulung gut. Wobei das für einzelne Arbeitnehmer:innen durchaus Ab- oder Aufwertungen (up/down-skill-Effekte) (Hirsch-Kreinsen in Woppen/Jannes: 81) mit sich bringen kann, die mit entsprechenden Vergütungsdifferenzen einhergehen.

Offensichtlich werden mit der Einführung von technisch autonomer Robotik in verschiedene Handlungskontexte, wie bei jeder anderen sozio-technischen Entwicklung auch, Gewinner:innen und Verlierer:innen dieser Entwicklung erzeugt, die auch Fragen der Verteilungsgerechtigkeit aufwerfen und damit ethische Reflexion auf den Plan rufen. Das wird auch explizit und zunehmend aus der Informatik/Robotikforschung heraus eingefordert, aus der technischen Perspektive meist mit Verweis auf die EURON Roboethics-Roadmap koordiniert von Gianmarco Veruggio. Der Wirtschaftsinformatiker Oliver Bendel wendete sich der Technikphilosophie zu und trägt zur interdisziplinären Diskussion zum einen eine Einordnung der Roboterethik in die Maschinenethik bzw. andere Bereichsethiken bei, und wendet sich zum anderen der Programmierung von Moral zu. Dabei wird deutlich, dass einfaches moralisches „Entscheiden“ keiner aufwendigen Programmierung bedarf, wenn beispielsweise ein Staubsaugerroboter insekten-schonend seinen Dienst verrichten soll. Er beschreibt auch die Dilemmata, die sich bei Entscheidungssituationen für Menschen und auch für Roboter ergeben. Das Trolley-Problem ist dabei sicherlich das prominenteste, wenn ein technisches System zwar einen Unfall physikalisch nicht mehr verhindern, aber den Ausgang des Unfalls noch beeinflussen kann, z. B. durch eine schnelle Richtungsänderung. Soll das technische System so programmiert werden, dass es in die Richtung lenkt, die zum Beispiel die Anzahl der betroffenen Menschen minimal hält? Also eine Gruppe aus drei Menschen zu Ungunsten einer Zweier-Gruppe verschont wird? Die Ethikkommission, die 2017 das deutsche Verkehrsministerium beraten hat, hat explizit darauf hingewiesen, dass diese Art von Abwägungen – auch zu anderen Kriterien als der Anzahl der Personen – nicht an das technische System übertragen werden sollen.

## Autonome Robotik

Inwiefern kann an ein technisches – möglicherweise selbstlernendes – Artefakt überhaupt autonomes Entscheiden übertragen werden? In einem wörtlichen Sinne bedeutet Autonomie, dass ein autonomes „Wesen“ selbständig Gesetze zu formulieren vermag, an die es sich dann aus freien Stücken hält. Dies setzt voraus, dass dieses Wesen eigene Interessen hat und selbständig Zwecke setzen kann. Diese starke Autonomie beschreibt zum Beispiel einen autonomen Staat. Bei Men-

schen geht man von einer Handlungsautonomie aus. Ein Mensch setzt sich Ziele, die er mit einer Handlung erreichen möchte, und er wählt die Mittel, mit denen er diese Ziele erreichen kann. Verbunden ist damit unmittelbar die Zuschreibung der Verantwortung für diese Handlung. Gleichzeitig ist damit die Erwartung verbunden, dass ein Mensch auch begründen kann, warum er so gehandelt hat. Am offensichtlichsten ist das vor Gericht. Wenn beispielsweise ein Rettungsring von einer Jacht entwendet wurde, dann wird das gerichtliche Urteil vermutlich unterschiedlich ausfallen, je nachdem ob als Grund dafür das Zuwerfen des Rings zu einem Ertrinkenden oder der Verkauf desselben auf einem Flohmarkt angegeben wird. Eine solche starke Autonomie ist für technische Systeme nicht anstrebenswert – im Gegenteil. Starke Autonomie würde gerade das konterkarieren, was beispielsweise eine Service-Robotik anstrebt. Wenn ein autonomer Roboter eigene Interessen verfolgt, wird er letztendlich Dienste auch verweigern. Der „Mittel zum Zweck“-Charakter eines technischen Geräts wäre dann schlicht nicht mehr gegeben.

Anders sieht es aus, wenn wir eine schwache Form der Autonomie annehmen. In ihrem Falle geht es darum, dass vorhandene Regeln aus freien Stücken befolgt werden. Es ist nun müßig über die ‚Freiheit von Artefakten‘ nachzudenken, aber dass ein „intelligentes“ Artefakt Regeln beachtet, die ihm gegeben wurden, ist eine technische Bedingung für seine Funktionsfähigkeit. Jedes autonome Fahrzeug hat quasi als oberste Regel die Vermeidung von Unfällen bzw. die Vermeidung von Personenschäden zu beachten. Alle seine Aktionen sind dieser obersten Regel ‚verpflichtet‘. Bei schwacher Autonomie findet keine Wahl oder eine Zustimmung im moralischen Sinne statt, aber eine Art ‚Anerkennung‘ der Rahmung eigener Aktionen. Das System errechnet gemäß bestimmten Präferenzen und entsprechend der zur Verfügung stehenden Daten die bestmögliche und ökonomischste Reaktion und leitet sie ein. Dies hat allerdings nichts mit starker Autonomie als Bedingung ethischen Verhaltens zu tun. Ethik setzt im Außertechnischen an, wählt, setzt, begründet und hierarchisiert Werte, die vom System verfolgt bzw. gesichert werden sollen.

Damit verbunden ist die Frage, ob autonome Systeme selbständig Entscheidungen treffen können? Für die starke Autonomie können wir das bejahen, denn sie setzt eigene Intentionen, die Zwecksetzung und die Mittelwahl voraus. Starke Autonomie ist in diesem Sinne „erfinderisch“, das heißt, sie setzt Mittel auch anders ein, als es bei der Entwicklung eines Mittels intendiert war. Das Öffnen einer Flasche mit Kronkorken mithilfe eines Zollstocks ist hierfür ein Beispiel. Von einer solchen Freiheit der Mittelwahl lässt sich bei schwacher Autonomie nur in einem eingeschränkten Sinne sprechen. Die Mittel müssen, wenn nicht klar definiert, so doch durch ein vorgegebenes Schema eingegrenzt bzw. vage bestimmt sein. Wenn ein autonomes technisches System über unterschiedliche Aktionsoptionen



verfügt, so ist das etwas anderes als die Freiheit der Mittelwahl. Autonome Technik ist nicht erfinderisch. Das heißt aber, man sollte hier auch nicht von ‚Entscheiden‘ sprechen. Wir können uns schließlich nur entscheiden, wenn wir die Wahl hatten. In einem strengen Sinne wählt ein System aber nicht, sondern ‚rechnet‘. Das geschieht nur innerhalb vorgegebener Rahmungen.

Diese grundlegende Einschätzung, dass starke Autonomie menschlichen Akteuren „vorbehalten“ ist, wird in allen drei Büchern geteilt. *Janina Loh* widmet sich den Fragestellungen technischer Autonomie ausführlich, und geht dabei der Frage nach, die sich nach den bisherigen Ausführungen fast schon aufdrängt: Wie nah können denn technische Systeme an eine starke Autonomie herankommen?

In ihrer Einführung in die Roboterethik beginnt sie mit der „Vorstellung“ von Robotersystemen in fünf verschiedenen Handlungszusammenhängen, nämlich der industriellen Produktion, dem autonomen Fahren, der medizinischen Therapie und Pflege, der Sex- und der Militärrobotik. Diese Auswahl ist deswegen geschickt, weil sich an ihr die nahezu volle Bandbreite der ethischen Fragestellungen abarbeiten lässt. Ähnliche ethische Fragen stellen sich beispielsweise bei Haushalts- oder Spielzeugrobotern. Für die vergleichende Betrachtung werden auch konkrete Robotersysteme herangezogen (Loh: 67ff.) – Kismet und Cog Blesu-2 und iCub sowie die Roboterrobbe Paro gehören sicherlich zu den Robotern mit einem höheren Bekanntheitsgrad.

*Loh* zielt auf die moralische Handlungsfähigkeit ab, die letztendlich immer relevant ist, wie man sich an obigem Beispiel „vor Gericht“ vor Augen führen kann. Hier wird zum einen der Top-Down Ansatz verfolgt, in dem moralische Regeln, ethische Prinzipien, sowie Gesetze – etwa die berühmten Asimov’schen Robotergesetze – in der Programmierung der Roboter verankert werden (Decker 2019). Das hat Vor- und Nachteile. Der Vorteil ist, dass die Regeln sich implementieren lassen, ja in ihren „wenn-dann“-Formulierungen geradezu dafür prädestiniert scheinen. Der Nachteil ist, dass die Regeln kontextsensitiv sind, also immer eine Interpretation vorausgehen muss, bevor die Regel angewendet werden kann. Damit unmittelbar verbunden ist, dass über die Interpretation auch die Anwendung unterschiedlicher Regeln plausibel begründet werden können. Konfliktfreie Antworten sind in der realen Welt selten. Der Bottom-up-Ansatz setzt umgekehrt auf das Erlernen moralischen Handelns und ethischer Regeln, indem man die Systeme an ethischen Fallentscheidungen trainiert oder sie „sozialisiert“, das heißt die Systeme „wie Kinder“ (Brooks/Stein, 1994) über Empathie und Emotion lernen lässt. Wobei das beobachtende Mitfühlen über von außen wahrnehmbare Gefühlsausdrücke unterschieden wird von einem „sich hineinversetzenden“ Mitfühlen, welches als Produkt menschlicher Sozialisation und als Basis moralischer Beurteilungen angesehen wird. Das nach dieser Hinführung naheliegende hybri-

de Modell verbindet beide Ansätze und geht von einem ethischen Regelrahmen aus, der anwendungsbezogen ausgestaltet sein sollte und innerhalb dessen dann individuelle moralische Beurteilungen, etwa an den Wertvorstellungen der Nutzenden orientiert, erlernt werden.

„Sich anpassen“ und das damit verbundene Lernen sind Aspekte der Künstlichen Intelligenzforschung für die *Loh* die grundlegende Definition von Russel und Norvig zugrunde legt (Loh: 47), nach der die schwache KI-Hypothese besagt, dass Maschinen so handeln als ob sie intelligent wären. Die starke KI-Hypothese besagt, dass Maschinen tatsächlich intelligent handeln, also nicht „nur“ simulieren. Inwieweit kann man heute schon davon reden, dass Roboter moralische Handlungssubjekte sind. Dafür beschreibt und beurteilt *Loh* vergleichend fünf Konzepte von moral agency, die alle den Raum zwischen einfacher Maschine und dem Menschen als einzigem umfassend moralisch handelndem Wesen auf unterschiedliche und abgestufte Weise füllen. Die Beurteilung des Systems von außen spielt dabei eine zentrale Rolle, wobei die funktionale Äquivalenz von Wallach und Allen (2008) in ihrem Buch *Moral Machines* das „als ob“ der schwachen KI-Hypothese aufgreift. Wenn ein Roboter so agiert, als ob er kognitiv oder emotional sei, weil er in der Lage ist, diese kognitive Leistung oder Gefühle so zu simulieren, dass er die gleichen Reaktionen in seiner Umwelt auslöst, dann ist funktionale Äquivalenz gegeben. Die graduellen Abstufungen sind dabei mannigfaltig und untergliedern entweder mit Derwall in persönliche, moralische, rationale und agentiale Autonomie, wobei heute, so *Loh*, die beiden erstgenannten den Menschen vorbehalten sind, oder in positive und negative Autonomie, wobei letztere als technische Autonomie aufgefasst wird, die aktuell als Stand der Technik in der Robotik angesehen wird. Selbstursprünglichkeit und die Fähigkeit sein Verhalten an Gründen zu orientieren sind die zentralen Begriffe für *Catrin Misselhorn*, wenn es um die Beurteilung einer grundsätzlichen Handlungsbeurteilung geht, wobei sie beide Attribute als graduell in dem Sinne versteht, als dass sie verschiedenen Akteuren in unterschiedlichem Maß zukommen (Loh: 64f; Misselhorn in Woopen/Jannes: 63). Aus der Perspektive der Verhaltensökonomie eines technischen Systems lassen sich verschiedene Zustände des Systems so beschreiben, dass eine funktionale Äquivalenz zum Handeln nach Gründen festgestellt werden kann. Technische Systeme verfügen über eine symbolische Repräsentation der Welt, in der verschiedene Pro-Einstellungen des Geräts programmiert sind. Diese können als Intentionen interpretiert werden, weil sie für den Roboter als Referenz dienen, wie seine Umwelt beschaffen sein soll. Ein Roboter verfügt auch über Repräsentationen, die angeben, wie sich Dinge in der Welt zueinander verhalten und in welchen logischen Beziehungen sie zueinander stehen. Hier sieht *Misselhorn* die funktionale Äquivalenz zu Überzeugungen. Schließlich können Roboter Pläne entwickeln, wie sie die Umwelt verändern, um von einem Zustand zu einem anderen

zu kommen. Die Software BDI (belief-desire-intention) beispielsweise ermöglicht es dem Roboter in der Welt zu manipulieren, als ob er nach Gründen handelte.

Auch das zweite Kriterium, Selbststursprünglichkeit, sieht *Misselhorn* graduell erfüllt, wenn man sie über die folgenden drei Merkmale beschreibt: i. Interaktivität, als die Fähigkeit, auf Stimuli mit einer Zustandsänderung zu reagieren, ii. Autonomie, als Fähigkeit einen inneren Zustand ohne äußeren Reiz zu ändern, und iii. Adaptivität, als Fähigkeit, die Art und Weise zu verändern, wie auf einen Reiz reagiert wird. Auf der Basis dieser graduellen Handlungsfähigkeit knüpft *Misselhorn* für eine moralische Handlungsfähigkeit an die funktionale Moralität von Wallach und Allen an. Diese liegt unterhalb einer „vollumfänglichen moralischen Handlungsfähigkeit“, welche an Bedingungen wie Bewusstsein und einen freien Willen geknüpft ist (Loh: 65). Aber für *Misselhorn* ist sie ausreichend für ein moralisch lernfähiges hybrides Altenpflegesystem, das es noch nicht gibt, für das sie aber eine entsprechende Roadmap vorstellt, wie es entwickelt werden sollte. Die entscheidenden Schritte, wie man sozusagen Moral in die Maschine bekommt, sind dabei die Identifikation der in der Altenpflege relevanten Werte aus der Perspektive der Nutzer:innen und die anschließende Operationalisierung dieser Werte durch Szenarien, wo sie die experimentelle Philosophie in einer entscheidenden Rolle sieht. Sie soll sicherstellen, dass in diesen Szenarien auch konfligierende moralische Werte durch die Nutzer:innen zu beurteilen sind. Diese realen Entscheidungen bilden dann die Datenbasis für das adaptive technische System.

## Der prominente Anwendungsfall: Pflege!

Die Pflege älterer und kranker Menschen ist für *Misselhorn* einer der Anwendungsbereiche, in dem adaptive, graduell autonom und moralisch handelnde Robotersysteme eingesetzt werden könnten. Den Bedarf dafür skizziert sie anhand der demographischen Entwicklung, nach der die Anzahl der älteren Menschen in Deutschland prozentual immer mehr zunimmt und somit auch die Anzahl der potentiell Pflegebedürftigen. Diesen steht eine tendenziell eher abnehmende Zahl von aktuellen und zukünftigen Pflegekräften gegenüber, so dass beide Trends gemeinsam den Pflegekräftemangel verschärfen. Technische Unterstützung könnte hier helfen, indem sie Anteile der Tätigkeiten der Pflegekräfte übernimmt, wobei aus dieser Angebots- und Nachfragesituation kein negativer Effekt auf den Arbeitsmarkt der Pflegekräfte zu erwarten wäre. Diese Argumentation ist vergleichsweise weit verbreitet und findet auch große Resonanz in der öffentlichen Debatte, wobei sich gerade aus der Perspektive der Pflegewissenschaften die Frage stellt, wieso? Gute Pflege, so unterstreicht *Hülsken-Giesler* (in Hergesell/Maibaum/Meister: 148), ist Interaktionsarbeit, mit den Teilbereichen Gefühlsarbeit, bezüglich der

Pflegebedürftigen und deren Angehörigen, Emotionsarbeit, bezüglich der Pflegenden selbst, und Kooperationsarbeit, im Sinne des kontinuierlichen Herstellens von Kooperationsbeziehungen. Sie ist in zweierlei Hinsicht wissenschaftlich fundiert, nämlich orientiert an der externen Evidenz des systematischen, wissenschaftlich fundierten Pflegehandelns und an der internen Evidenz der spezifischen Besonderheiten des Einzelfalls, höchst situativ und kontextgebunden. Dieses „hermeneutische Fallverstehen“ (ebd.: 148), das sowohl kognitiv-rationale als auch affektiv-emotionale Aspekte enthält, bei *Koppenburger* und *Wüller* „Regel- und Fallwissen“ (ebd.: 163), ist konstitutiv für gute Pflege. Vor dem Hintergrund dieser Beschreibung und der Darstellung des technisch Möglichen ist es nicht verwunderlich, dass sich der pflegetheoretische und der pflegephilosophische Diskurs durch skeptische Zurückhaltung gegenüber dem Einsatz von Robotik auszeichnet (ebd.: 149), letztlich sogar ein Mismatch von tatsächlichem Bedarf in der Pflege und technischem Angebot (Hergesell/Maibaum/Meister: 8) konstatiert wird.

Insofern ist es ein besonderer Verdienst des Sammelbandes „Genese und Folgen der Pflegerobotik“, dass er dieser Frage gezielt nachgeht. *Lipp* identifiziert für Roboterpflege die Verschaltung dreier Begriffe. Das ist zum einen das „Aktive Altern“ mit dem die Hoffnung verbunden ist, „dass der demographische Wandel gerade nicht zu einer Belastung der Gesellschaft, sondern zu einer silbernen Ökonomie führt“ (Lipp in Hergesell/Maibaum/Meister: 23). „Technische Innovation“ ist das Bindeglied, das zur Gestaltungsaufgabe einer alternden Gesellschaft wird und schließlich die „robotisierte Assistenz“, die ein unabhängiges Leben, möglichst lang in den eigenen vier Wänden ermöglicht. Die Verschaltung dieser drei Begriffe wurde, so *Lipp*, erst vor einem sich verändernden Modus der europäischen Innovationspolitik möglich, die er so beschreibt: Vom wissenschaftlichen Verstehen zu technowissenschaftlicher Intervention, von Pflege zu Assistenz, vom Sicherheits- zum Innovationsdispositiv (ebd.:41). Förderungspolitisch passte dann die technisch wünschenswerte Robotik zum großen gesellschaftlichen Bedarf den Pflegenotstand zu lösen, was eine seltene win-win-Situation darstellt: Nach den Erfolgen in der Industrierobotik stellt Pflege eine echte inhaltliche Herausforderung für die Robotikforschung dar und dieses Benchmark ist von höchster gesellschaftspolitischer Relevanz. In der Umsetzung wird diese vermeintliche win-win-Konstellation aber zur vorgegeben Problemlösung. Forschungsprojekte werden nicht mehr ergebnis-offen ausgeschrieben, sondern die Servicerobotik wird schon als selbsterfüllendes Lösungsversprechen in der Ausschreibung investiert. In der Umsetzung dieser Forschungsprojekte wird dann eine spezifische Teiltätigkeit, die vom Roboter ausgeführt werden soll, identifiziert und bearbeitet, ohne sie im Kontext des Pflegehandelns ausreichend zu reflektieren.

Dabei ist genau diese Reflektion in den Forschungsprojekten meist schon durch die Ausschreibung forschungspolitisch explizit erwünscht (Bruckamp in Her-

gesell/Maibaum/Meister: 211). Unter dem Stichwort der integrierten Forschung sind ethische, rechtliche, soziale Implikationen (ELSI) der technischen Entwicklungen zu erforschen und zwar nicht „begleitend“ sondern so integriert, dass sich aus den ELSI-Forschungsergebnissen ergebende Gestaltungsvorschläge noch in der technischen Entwicklung umsetzen lassen (ebd.: 199). Hierbei ist davon auszugehen, dass man sehr kontext-bezogen forscht, denn ethische und soziale Folgen sowie die rechtliche Beurteilung des Robotereinsatzes lassen sich erst im konkreten Handlungszusammenhang beurteilen. Die Beurteilung kann ihren Anfang in einer groben Klassifizierung nehmen, so schlägt *Bruckamp* die beiden Kategorien „Assistenzroboter“ und „Sozialassistentenroboter“ (Robot-Companions) vor (ebd.: 202), *Klebbe* und *Eicher* unterteilen die Assistenzroboter noch in physisch unterstützend und mobilitätsunterstützend (in Hergesell/Maibaum/Meister: 223), aber unterhalb dieser Beschreibungsebene müssen die durch Robotik veränderten Handlungszusammenhänge und deren Folgen konkret analysiert werden. Die methodische Vorgehensweise ist partizipativ und bezieht die Nutzenden einer Technologie ein, wobei der Kreis der Nutzenden unterschiedlich gezogen wird. Die Pflegebedürftigen – hier ist zu berücksichtigen, dass kognitive Einschränkungen wie Demenz eine besondere Herausforderung darstellen – und die professionell oder familiär Pflegenden werden typischerweise einbezogen, die Familienangehörigen im Falle der professionellen Pflege nur gelegentlich wie auch Ärzte u. a.. Die Intention, die mit der Nutzereinbindung methodisch häufig verfolgt wird, ist die Bedarfsanalyse und die möglichst frühe Einbeziehung der Nutzer:innen, um diese Perspektiven in den technischen Entwicklungsprozess einzubeziehen (ebd.: 211; 227). Hier muss man allerdings enge Grenzen konstatieren: Wenn das robotische System, das in dem Projekt technisch entwickelt werden soll, bereits bei der Beantragung des Projekts weitestgehend festgelegt ist, z. B. ein Roboterarm am Pflegebett, dann ist eine ergebnis-offene Bedarfsanalyse, die einen anderen Roboter oder gar eine nicht-technische Lösung vorschlägt, nicht mehr zielführend. Es können also durch partizipative Formate nur Veränderungen am bereits festgelegten technischen System vorgeschlagen werden. Das ist im Sinne eines *constructive technology assessments* durchaus nicht unüblich. Aber die Gefahr bleibt hoch, dass zwar ein erfolgreiches Technikentwicklungsprojekt durchgeführt wurde, allerdings wieder ein *mismatch* in Bezug auf das, was für die Pflege wirklich hilfreich gewesen wäre.

Entgegen der häufig vertretenen Meinung, dass Japan bei einem ähnlichen, letztlich sogar dramatischeren demographischen Wandel eine grundsätzlich positivere Einstellung zu Pflegerobotern habe, muss man feststellen, dass sich die japanische „Robot-Revolution“ (Brucksch in Hergesell/Maibaum/Meister: 101), die seit knapp 10 Jahren propagiert wird und die seit der Jahrtausendwende verfolgte Robotik-Innovationspolitik noch verstärkt, noch nicht ereignete. Die

Realität hinkt „in weiten Teilen den theoretischen Ansprüchen hinterher“ (Hofstetter in Hergesell/Maibaum/Meister: 137). Das ursprünglich stark von humanoiden und zoomorphen Robotern geprägte Robotikbild Japans zielte zunächst auf Companionship und Kommunikation, in ähnlicher Weise wie in Europa kulturell geprägt durch Science Fiction. Heute, so die These von *Brucksch*, werden andere Entwicklungsziele in der Pflegerobotik verfolgt, die einem besseren Verständnis des Pflegekontextes geschuldet sind. Mitentscheidend für diese Trendwende ist auch in Japan eine fehlende Akzeptanz der Nutzer:innen (Brucksch in Hergesell/Maibaum/Meister: 96). Aus der Pflegeperspektive heraus rückt alltäglich genutzte Technik in den Fokus, wie der Rollstuhl und das Pflegebett, welche durch Robotik verbessert werden sollen. Die damit verbundenen Herausforderungen der Robotikentwicklung sind nicht weniger groß, auch wenn das Erscheinungsbild deutlich weniger spektakulär ist, als ein humanoider Roboter. In der politischen Umsetzung wird die Ausbildung der professionell Pflegenden deutlich stärker in den Blick genommen, und damit unterstrichen, wer die Hauptnutzenden dieser technischen Geräte sind (Hofstetter in Hergesell/Maibaum/Meister: 134). Pflegerobotik wird stärker als technisches Mittel angesehen, um die Zwecke der professionellen Pflege zu erreichen. Die mögliche Einordnung als ein im solidarisch finanzierten Gesundheitssystem zu finanzierendes medizinisches Hilfsmittel (Brucksch in Hergesell/Maibaum/Meister: 101) ist dann folgerichtig. *Hofstetter* empfiehlt, in Deutschland ebenfalls in Aus-, Fort- und Weiterbildung der Pflegekräfte darauf hinzuwirken, dass pflegerobotische Geräte in den Arbeitsalltag integriert werden.

## Ausblickendes Fazit

Gerade Handlungsempfehlungen dieser Art, Robotik über die Aus- und Weiterbildung in die professionelle Pflege zu integrieren, heben den „Mittel zum Zweck“-Aspekt der Robotik hervor. Gleichzeitig stellen Roboter eine besondere Technologie dar, weil sie auch weitreichendere Ersetzungen menschlicher Tätigkeiten mit sich bringen. Gerade wenn man an durch KI ertüchtigte adaptive Robotersysteme denkt, die einerseits eine notwendige Bedingung darstellen, um überhaupt in alltäglichen Umgebungen ihre Dienste kontextbezogen absolvieren zu können, und andererseits über höchst spezifisches, umgebungs- und personenbezogenes Wissen und graduell weitreichendere technische Autonomie verfügen, die über normale Technik hinausgehen. Somit liegen auch darüber hinaus gehende Überlegungen nahe, die Dichotomie zwischen Mensch und Maschine aufzulösen und den Roboter als Quasi-Subjekt zu betrachten (Weyer in Woopen/Jannes: 98; 2, Loh: 48ff) oder zumindest, im Sinne einer moral patiency, dafür zu plädieren, dass Roboter Objekte moralischen Handelns sein können, die somit einen spezi-

fischen (moralisch angemessenen) Umgang erfordern (Loh: 75). Diesen Überlegungen müsste man eine eigene Befassung widmen. Stellvertretend kann man sich fragen, ob man in dem Film „Ich bin dein Mensch“ zu denjenigen gehörte, die ein Bedauern empfanden, als „das Gedächtnis“ des humanoiden Roboters, das gerade in einem sehr aufwändigen KI-basierten Personalisierungsprozess auf sein menschliches Gegenüber angepasst wurde, gelöscht werden sollte. Abschließend soll, aus der Perspektive der politikberatenden Technikfolgenforschung naheliegend, auf den politischen Beratungsbedarf eingegangen werden. Neben Fragen der Arbeitsmarktentwicklung stehen hier innovationspolitische Aspekte im Zentrum, die dadurch geprägt sind „dass die Politik sich zum einen nicht aus wissenschaftlich-technologischen Innovationsprozessen heraushalten kann und andererseits nicht von der Annahme ausgehen darf, fördernde, limitierende oder prohibitive Entscheidungen ein für alle Mal treffen zu können“ (Lammert in Wopen/Jannes: 112). Spezifischen Gestaltungsbedarf sieht *Norbert Lammert* bei autonomen Robotern in militärischen und medizinischen Anwendungen, sowie bei der Moral technischer Systeme. Dieser Beratungsauftrag ist noch nicht umfänglich erfüllt, aber die disziplinäre und interdisziplinäre Basis ist gelegt.

## Literatur

- Brooks, R.A.; Stein, L.A. Building brains for bodies. *Autonomous Robots* 1994, 1, 7–25.
- Decker, M. Von den three laws of robotics zur Roboterethik. In *Mensch-Maschine-Interaktion*; Liggieri, K.; Müller, O.; Hrsg.; JB Metzler: Stuttgart, 2019; pp 345–351.
- Frey, C.; Osborne, M.A. *The Future of Employment: How Susceptible are Jobs to Computerization?*; University of Oxford – Working Paper, 2013.
- Wallach, W.; Allen, C. *Moral Machines: Teaching Robots Right from Wrong*; Oxford University Press: Oxford, 2008.