

8. Soziale Roboter aus ethischer Sicht

Oliver Bendel und Jeanne Kreis

8.1. Die ethischen Perspektiven

Soziale Roboter werfen Fragen unterschiedlicher philosophischer Disziplinen auf. Die Philosophie des Geistes etwa beschäftigt sich mit dem Zusammenhang zwischen Körper und Geist. Bezogen auf soziale Roboter kann sie sich danach erkundigen, ob diese Wahrnehmungen irgendeiner Art haben. Schreibt man ihnen geistige Zustände wie Wünsche, Absichten, Vorstellungen, Erinnerungen, Empfindungen oder Gefühle zu, so fragt sich, unter welchen Voraussetzungen wir Maschinen gar ein Bewusstsein attestieren und welche Folgen maschinelles Bewusstsein für den moralischen Status einer Maschine haben kann.

Die Sprachphilosophie untersucht die Bedeutung und Funktionsweise von Sprache und ihr Verhältnis zum Denken und zum Bewusstsein sowie zur Wirklichkeit. Dabei wird Sprache meist als Werkzeug zur Fixierung und Mitteilung von Bewusstseinsinhalten oder aber als strukturgebendes Medium aufgefasst, das unsere Weltansicht noch vor aller Erfahrung prägt (Gessmann, 2009). In Bezug auf soziale Roboter kann die Beziehung zwischen der Wirklichkeit und sprachlichen Äusserungen von Maschinen analysiert werden. Zudem stellt sich die Frage, welche Inhalte der Roboter mit Hilfe der Sprache vermittelt und inwiefern Sprache den Maschinen Zugang zu einer Gesellschaft gewährt.

Auch auf dem Gebiet der Rechtsphilosophie werden durch soziale Roboter viele Fragen aufgeworfen. Soll ein Roboter gesetzlich vorgeschriebene Rechte haben? Welchen Rechtsnormen soll er folgen? Inwiefern ist sein Fehlverhalten zu sanktionieren und wer trägt die Verantwortung für sein Tun? Neben rechtsphilosophischen Überlegungen wird die soziale Gerechtigkeit zum Thema. Für wen werden Roboter künftig verfügbar sein? Wie können Bedürfnisse von Menschen gerecht gegeneinander abgewogen werden? Welche politischen Folgen sind zu erwarten und wie beeinflussen soziale Roboter die Verteilung von Ressourcen? Auch das Roboterrecht trägt hier zu Antworten bei, und einzelne Aspekte werden in Kapitel 9 aufgegriffen.

Nicht zuletzt stehen soziale Roboter in Zusammenhang mit einer Reihe weiterer vorwiegend ethischer Fragen (Bendel, 2021). Dabei ist insbesondere die Menschenwürde im Fokus. Kritiker argumentieren, dass soziale Roboter der Würde des Menschen nicht gerecht würden, da ihr Einsatz unter anderem Isolation und Kontaktverlust begünstige (Sharkey & Sharkey, 2010b). Dies wäre aber nur der Fall, wenn Ausschliesslichkeit in der Verfügbarkeit vorhanden wäre. Wie oben festgestellt wurde, kann der Roboter gerade auch zur Kommunikation, zur Interaktion und zur Reflexion anregen. Weiter sind Roboter meist mit Mikrofon und Kameras ausgestattet. Als potenzielle Spione verunmöglichen sie so die Privatsphäre in den eigenen vier Wänden (Misselhorn, 2021). Dies ist auch ein Thema der Informationsethik (Bendel, 2018b).

8.1.1. Roboterethik

Roboterethik («robot ethics» oder «roboethics») kann ganz unterschiedlich definiert werden (Lin et al., 2012; Bendel, 2013, 2019d, 2019e; Loh, 2019; Wallach & Asaro, 2016). Nach einer Lesart handelt es sich um einen Teilbereich der Maschinenethik (eben mit dem Fokus auf Robotern). Nach einer anderen beschäftigt sich die Roboterethik mit dem moralischen Status von Maschinen, vor allem mit der Frage ihrer moralischen Rechte, oder mit den Folgen des Einsatzes von Robotern (insbesondere von Industrierobotern, Servicerobotern, Kampfrobotern sowie – in einer anderen Kategorisierung – sozialen Robotern). Die Folgen wiederum können auch in der Technikethik oder der Informationsethik untersucht werden. Die Roboterethik mag als neuere Bereichsethik oder als Teil einer klassischen Bereichsethik (etwa der Informationsethik) aufgefasst werden. Man kann sie zur Roboterphilosophie zählen (Bendel, 2018h).

Soziale Roboter können nach weit verbreitetem Verständnis – wie Roboter überhaupt – keine moralischen Rechte haben (Bendel, 2018c). Diese sind an Empfindungs- oder Leidensfähigkeit bzw. an Bewusstsein gekoppelt.

Menschen haben Rechte, Tiere können Rechte haben, Roboter nicht. Roboter sind überhaupt keine Objekte der Moral und können in moralischer Hinsicht – sofern es um sie selbst geht – in beliebiger Weise behandelt werden. Fragen stellen sich allenfalls in Bezug auf die Besitzer oder Begleiter der Roboter, die natürlich Rechte in irgendeiner Form haben und die durch das spezifische Verhältnis als Objekte der Moral gelten können. Zudem argumentieren einige Experten, dass Menschen verrohen können, wenn sie den Roboter – der gerade bei einer animaloiden oder humanoiden Gestaltung als Stellvertreter von Tieren oder Menschen angesehen werden kann – absichtsvoll beschädigen oder «missbrauchen». Auch hier ist aber der Roboter kein Objekt der Moral, sondern der Täter selbst bzw. eine von seinen aus der Verrohung entstehenden Handlungen betroffene Person.

Gerade das Simulieren von Emotionen und Empathie sorgt in diesem Zusammenhang immer wieder für Diskussionen. Ein Sexroboter, der sagt, dass er keinen Geschlechtsverkehr haben möchte, dann aber dazu «gezwungen» wird (und sich dabei vielleicht noch verbal oder physisch wehrt), ist solch ein kontroverser Fall. Auch hier ist der soziale Roboter kein Objekt der Moral. Dennoch kann das Verhalten des Menschen als verwerflich angesehen werden, weil er sich über eine Bitte, eine Ablehnung oder eine Weigerung hinwegsetzt, wobei argumentiert werden kann, dass diese nur vorgespielt sind. Zudem kann das reale Einüben am künstlichen Objekt beanstandet werden, auch mit Blick darauf, dass eine Übertragung auf natürliche Objekte (insbesondere Menschen) stattfinden kann.

Die Roboterethik kann sich weiter damit beschäftigen (eine Möglichkeit wurde schon angedeutet), wie die Täuschung der Maschine, wenn man von einer solchen ausgehen will, einzuordnen ist. Sie würde im vorliegenden Zusammenhang danach fragen, ob die Täuschung (oder der Betrug) durch den sozialen Roboter vom Hersteller bewusst oder unbewusst umgesetzt wird, welchen Zweck sie allenfalls hat, welcher Art diese Täuschung ist, also ob sie etwa durch die Gestaltung, durch Akte oder Aktionen oder durch Sprechakte hervorgerufen wird, und welche Voraussetzungen auf der Empfängerseite vorhanden sind (Alter, Geschlecht, Bildung, geistiger Zustand, Bereitschaft, sich täuschen zu lassen etc.). Sofern der soziale Roboter nur Mittel zum Zweck ist, lassen sich klassische Überlegungen zu Täuschung und Betrug direkt anwenden.

Immer wieder wird argumentiert, dass humanoide Roboter an sich bereits Täuschungen seien, weil sie durch ihre Gestaltung menschenähnlich wirken. Dann müsste man freilich auch Schaufensterpuppen als Täuschungen auffassen (die sich meist nicht bewegen können, jedoch wie Menschen angekleidet werden und bekleidet sind), und es ist die Frage, ob man hier nicht neutraler sprechen will. Man könnte sagen, dass die Täuschung erst dann entsteht, wenn die Mittel des sozialen Roboters gezielt dazu eingesetzt werden, über etwas hinwegzusehen oder etwas anders zu verstehen, als es der Wirklichkeit entspricht. Dann müsste man aber immer noch untersuchen, ob etwas zu einem (sozialen) Spiel gehört oder nicht. So geben einige soziale Roboter (wie auch Chatbots, die einen Avatar haben, und Sprachassistenten) als ihr Alter das an, was ihrem Aussehen oder ihrer Stimme am nächsten kommt, während andere das Datum ihrer Entwicklung oder ihres Markteintritts nennen.

Neben der Täuschung ist das Lügen von Interesse. Im Gegensatz zur Täuschung (die auch aus dem Tierreich bekannt ist) muss sich das Lügen auf einen Sprechakt bzw. eine sprachliche Äußerung beziehen. «Lügen» wird vom Duden definiert als «bewusst und absichtsvoll die Unwahrheit sagen» (duden.de, o.D.). Ein sozialer Roboter oder ein künstlicher Agent hat kein Bewusstsein im engeren Sinne (Misselhorn, 2018, 2019). Er kann aber Absichten oder Ziele einprogrammiert bekommen – Softwareagenten werden u.a. durch ihre Zielorientierung definiert (Bendel 2003). Wenn ein sozialer Roboter natürlichsprachliche Fähigkeiten hat, kann er damit auch die Unwahrheit sagen (ausführlich dazu im Kontext des LIEBOT-Projekts; z.B. Bendel et al., 2017). Die Roboterethik kann sich damit beschäftigen, wie das Lügen der Maschine einzuordnen ist, welche Folgen es hat für das Zusammenleben zwischen Menschen und das Koexistieren zwischen Menschen und Maschinen.

Im Zusammenhang mit der Koexistenz von Mensch und Maschine ist auch Kontaktverminderung ein häufig auftretendes Thema. Den Herstellern und Anwendern sozialer Roboter, die menschliche Handlungen ganz oder teilweise übernehmen, wird oft vorgeworfen, sie würden dazu beitragen, zwischenmenschliche Kontakte zu vermindern. Kritiker fürchten, der Einsatz sozialer Roboter könnte gar zur sozialen Isolation älterer oder physisch eingeschränkter Menschen führen (Stösser, 2011). Freilich kann der soziale Roboter auch als etwas eingesetzt werden, über das man spricht, etwa in einem Pflegeheim, und so in der Gruppe einen gemeinsamen Gegenstand der Reflexion darstellen. Die Roboterethik kann in diesem Sinne nach den moralischen Implikationen im sich verändernden sozialen Gefüge fragen, zusammen mit der Soziologie und der Psychologie.

Roboterethik kann weiter danach fragen, wie die Gestaltung des sozialen Roboters, seine Akte und Sprechakte sowie die von ihm gezeigten Emotionen und die von ihm gezeigte Empathie bei den Menschen Emotionen und

Empathie auslösen. Man kann hier einteilen nach Gefühlen, die unabsichtlich oder absichtlich, ziellos oder zielbewusst ausgelöst und verstärkt werden. Sie kann nach der Bindung fragen, die dies erzeugt und verstärkt, nach der Abhängigkeit von Artefakten, nach der Übertragung von Verhalten auf Mitmenschen. Wenn man sich Täuschung und Betrug widmet, ist man wiederum bei den Themen angelangt, die schon oben abgehandelt wurden.

Eine Rolle spielt in diesem Zusammenhang auch, ob der Roboter bestimmte Vorurteile zeigt, ob er Verzerrungen und Verschiebungen in seinen Akten und Sprechakten aufweist, ob er einen Bias in irgendeiner Form hat oder transportiert. Dabei ist zu unterscheiden zwischen solchen Verhaltensweisen, die ihm bewusst beigebracht wurden, und solchen, die ihm unbewusst – etwa als Spiegelung der Vorurteile des Programmierers – beigebracht wurden. Auch die Gestaltung spielt hier wieder eine Rolle, und Philosophen wie John Danaher haben – etwa bei der Tagung Robophilosophy 2020 – die weiße Farbe einiger sozialer Roboter kritisiert, zugleich aber davor gewarnt, ihnen eine schwarze Farbe zu geben, weil dies rassistisch sein könnte, vor allem dann, wenn man die Verbindung zu Sklaven herstellt (Bryson, 2010). Es ist freilich die Frage, welche Farbe man dann überhaupt wählen soll (Sparrow, 2020). Genauso kann man die Farbe von Pepper oder QTrobot nicht als Hautfarbe, sondern als medizinisches Weiss deuten. Bei Androiden wie Sophia und Erica treten allerdings tatsächlich Probleme auf. Bei ihnen wird die Farbe der Oberfläche unweigerlich als Hautfarbe gedeutet.

Die Roboterethik kann in diesen und in anderen Zusammenhängen nach individuellen moralischen Rechten von Betroffenen fragen. Grundsätzlich bedarf es einer gewissen Medien- und Sachkompetenz, um mit einem sozialen Roboter so interagieren und kommunizieren zu können, dass der Nutzen gegeben ist und kein Schaden für den Nutzer entsteht (Torras, 2010). Bestimmte Gruppen können diese Medien- und Sachkompetenz aber kaum haben, entweder weil die Mitglieder zu jung sind oder weil sie sich nicht mehr in der Lage fühlen, sich auf Neues und Unbekanntes einzustellen. Darüber hinaus fehlt ihnen vielleicht ein grundsätzliches Verständnis dafür, was ein Roboter bzw. ein sozialer Roboter ist. Diese Gruppen, grob gesagt Kinder und Hochbetagte, zudem geistig eingeschränkte, kann man als vulnerabel bezeichnen. Die Roboterethik kann untersuchen, wie diese betroffen sind und wie man ihre individuellen moralischen Rechte schützen kann, nicht zuletzt durch die bewusste Gestaltung der Artefakte oder den Verzicht auf ihren Einsatz in bestimmten Zusammenhängen.

All diese Diskussionen kann man auch – wie weitere – mit Blick auf unterschiedliche Kulturen führen. Europa hat in Bezug auf künstliche Kreaturen ein eher negatives Narrativ, wenn man an Pandora, Talos, Golem, Frankenstein's Monster usw. denkt. Japan hat ein eher positives oder zumindest ambivalentes Narrativ. Je nach Land und Gesellschaft kann es unterschiedliche Blicke auf soziale Roboter und unterschiedliche Diskussionen darüber geben. Beispielsweise werden moralische Rechte von Robotern in Europa vielleicht eher abgelehnt, anderswo eher befürwortet – was auch von grundlegenden philosophischen Konzepten abhängt. In Europa und in den USA wird man eher eine Diskussion in Bezug auf Diskriminierung und Diversität führen, anderswo womöglich gar nicht den Punkt verstehen, um den es geht. In diesem Sinne ist die Behandlung der Sozialen Robotik in die Geschichte, die Sprache, das Selbstverständnis etc. einer Kultur eingebettet.

8.1.2. Maschinenethik

Maschinenethik («machine ethics») beschäftigt sich mit künstlicher oder maschineller Moral («artificial morality» oder «machine morality») und moralischen (oder unmoralischen) Maschinen («moral machines») (Wallach & Allen, 2009; Anderson & Anderson, 2011; Misselhorn, 2018; Bendel, 2012a, 2019b, 2019g). Die Begriffe, die den Gegenstand bezeichnen, sind Termini technici, wie «künstliche Intelligenz»: Es wird nicht gesagt, dass Maschinen im Sinne menschlicher Subjekte moralisch oder intelligent sind, sondern dass man versucht, Aspekte menschlicher Moral abzubilden (so wie man Aspekte menschlicher oder auch tierischer Intelligenz abbildet), dass man Vorbilder aller Art versucht zu simulieren. Ähnlich verhält es sich übrigens auch beim Sprechen über soziale Roboter: Diese sind nicht in der Weise sozial, wie es Menschen oder Tiere sind, sondern sie simulieren deren Eigenschaften und Verhaltensweisen und bewegen sich mit diesen in einem sozialen Setting, in Gruppen und Gemeinschaften ähnlicher und unterschiedlicher Entitäten.

Existierende moralische Maschinen setzen vor allem moralische Regeln um, die ihnen einprogrammiert wurden (Bendel, 2012a, 2018d, 2018f, 2018g, 2019b, 2019d). Betroffen sein können sowohl Akte (im Sinne von Aktionen) als auch Sprechakte (im Sinne von Monologen und Dialogen). Ein weiterer Ansatz ist, dass die Maschinen die Regeln selbst generieren (etwa auf der Grundlage von Erfahrungen) oder modifizieren (etwa nach dem

Zufallsprinzip) (Anderson et al., 2019). Ferner versucht man die äussere und stimmliche Gestaltung systematisch und mit moralischen Konnotationen zu verändern. Die Maschinenethik kann, ausgehend vom Subjekt der Moral, das untersucht wird, als Pendant zur Menschenethik gesehen werden. Manche verstehen sie auch als (Teil-)Bereichsethik und ordnen sie Informations- oder Technikethik zu oder unter bzw. sehen sie als neues Gebiet neben diesen oder als weiteres Gebiet, dem u.a. die Roboterethik zugehört. Man mag die Maschinenethik zur Philosophie zählen, muss aber bedenken, dass sie eine grosse Nähe zu den Ingenieurwissenschaften hat und sich immer wieder ihrer Begriffe bemächtigt. Sie arbeitet eng mit KI und Robotik zusammen, um ihre Artefakte hervorbringen zu können. Misselhorn (2019) verortet sie zwischen Philosophie und Informatik, wobei man korrekterweise noch die Robotik hinzunehmen müsste.

Wichtige Ansätze werden bei Bendel (2019e), Misselhorn (2018) folgend, so zusammengefasst:

In der Diskussion der Umsetzung maschineller Moral wird häufig von einem Top-down- und einem Bottom-up-Ansatz gesprochen. Bei Top-down-Ansätzen geht man von allgemeinen Strukturen zu immer spezielleren Details (Misselhorn, 2018). Man kann sie mit Prinzipienethiken verbinden, einerseits mit der Pflicht- oder Pflichtenethik, andererseits mit der Folgenethik. Man kann der Maschine z.B. beibringen, eine bestimmte Regel starr zu befolgen, oder, unter Einbezug möglicher und vermuteter Folgen, bestimmte Ausnahmen von der Regel zuzulassen bzw. je nach Situation zu entscheiden. Bei Bottom-up-Strategien versucht man nicht, moralische Regelwerke vorzugeben, sondern etwa – wegen des kontextsensitiven Charakters der Moral – das situative Urteilsvermögen zu stärken (Misselhorn, 2018). Man kann sie mit der Tugendethik zusammenbringen. In der heutigen Maschinenethik dominieren die ersten beiden Modelle normativer Ethik.

Bekannt geworden ist auch die Einteilung nach der Art der Subjekte nach Moor (Anderson & Anderson, 2011). Er geht davon aus, dass Maschinen keine vollständigen moralischen Subjekte sind (keine «full ethical agents»). Vielmehr kann das Moralische implizit oder explizit (bis zu einem bestimmten Umfang) in ihnen stecken. Dass man überhaupt von Subjekten spricht, liegt zunächst an ihrer Autonomie (auch dies ein Terminus technicus, und es ist zu beachten, dass dieser Begriff von den Ingenieurwissenschaften ganz anders gebraucht wird als etwa von der Philosophie oder Theologie, u.a. auch, um Automaten wie Kaffeeautomaten von weiterentwickelten Maschinen unterscheiden zu können und deren höhere Selbstständigkeit und ausgeprägtere Entscheidungsfähigkeit zu betonen). Die Maschinen bewegen sich alleingelassen durch eine Umwelt, die vielleicht in ihren Grundzügen bekannt ist, wo sich aber ständig etwas wandeln kann, sei es, weil das Personal, sei es, weil die Situation sich ändert. In diesem Alleingelassensein (wo die Maschinen nicht begleitet, geführt oder ferngesteuert sind) müssen bestimmte Optionen gewählt, bestimmte Entscheidungen getroffen werden. Auch bei «Entscheidungen» gilt, dass die technischen Wissenschaften hier ein eigenes Verständnis haben; so entwickelt man etwa in der Informatik und Wirtschaftsinformatik (und auch in der Maschinenethik) sogenannte Entscheidungsbäume (in der Maschinenethik annotierte Entscheidungsbäume, s. Bendel, 2016a), unabhängig davon, ob Geisteswissenschaften Maschinen Entscheidungen zubilligen oder nicht. In diesem beschränkten Sinne entstehen künstliche Subjekte, von denen eben Entscheidungen (auch moralischer Art) ausgehen können.

In der Maschinenethik interessiert u.a., wie man Empathie und Emotionen im Moralischen in technischer Weise umsetzt. So kann es einen moralisch relevanten Akt darstellen, wenn ein sozialer Roboter eine trauernde Person aufzumuntern versucht oder Freude über einen Erfolg eines Schülers oder einer Schülerin simuliert. Die Systeme können versuchen, die Probleme des Nutzers zu erkennen, um dann eine (auch in moralischer Hinsicht) adäquate Reaktion zu zeigen. Dies wurde etwa im GOODBOT- und BESTBOT-Projekt der Hochschule für Wirtschaft FHNW ausprobiert (Bendel, 2018d). Dort wurden auf Text- bzw. Bildebene Analysen vorgenommen, mithin Emotionsanalysen, die in Akte und Äusserungen des Chatbots (die jederzeit in soziale Roboter integriert werden können) mündeten. Während nach Stangl (2020) Empathie die Voraussetzung für Moral ist, trifft dies auf maschinelle Moral nicht im eigentlichen Sinne zu, schon weil diese eine simulierte Moral ist – sie können aber, wie gesagt, miteinander verbunden werden.

Die Maschinenethik mag dazu genutzt werden, Täuschungen durch den Roboter und falsche Erwartungshaltungen des Nutzers zu vermeiden oder abzumildern. So ist ein Vorschlag von Bendel, einen Verfremdungseffekt, einen V-Effekt, zu verwenden, frei nach dem Theater von Bertolt Brecht, wodurch die Illusion (hier die Illusion, einen Menschen vor sich haben, oder die Illusion, dass der soziale Roboter Emotionen und Empathie nicht nur zeigt, sondern auch hat) gestört wird (Freuler, 2016). Der Roboter kann den V-Effekt durch Änderungen in der Gestalt, durch Töne sowie durch explizite Hinweise («Ich bin kein Mensch», «Ich bin nur eine Maschine» etc.) erreichen. Unerforscht ist, ob sich ein solcher V-Effekt, auch wenn er zunächst wirken würde, mit der Zeit abschwächt, und

es wurde bereits thematisiert, dass es die Beziehung zwischen Roboter und Mensch auch stören kann (s. Kapitel 6). Man kann in diesem Zusammenhang ebenso von Transparenz sprechen (auch im Sinne der DSGVO und des Bundesgesetzes über den Datenschutz, wo u.a. ein Auskunftsrecht vorgesehen ist) oder von einer Herstellung informationeller Autonomie im spezifischen Sinne. Die Europäische Kommission hat am 21. April 2021 einen Rechtsrahmen für den Einsatz von KI-Systemen vorgestellt, in dem etwa Transparenzpflichten solcher Art mit Blick auf Chatbots erwähnt werden (https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682).

Natürlich kann man auch die Maschine deutlich machen lassen, dass sie zwar nur eine Maschine, aber womöglich mit einer Cloud verbunden und Teil einer Infrastruktur ist, dass sie mit anderen Systemen kommuniziert, dass hinter ihr Personen und Interessen stehen. Sie würde im Sinne der Transparenz beispielsweise darlegen, dass in einem bestimmten Moment der Dialog von ihrer Seite maschinell geführt wird, dass sie deshalb kein wirkliches Verständnis von den Anliegen, Erwartungen und Darlegungen des Nutzers hat, dass aber im gleichen Moment oder im Nachhinein Mitarbeitende den Chat auswerten, bewerten und Schritte einleiten können. Der Nutzer sollte also bestimmte Dinge für sich behalten, und es sollte ihm klar sein, dass auch die Unterhaltung mit einer Maschine im Extremfall dazu führen kann, dass die Polizei an die Türe klopft. Die Aussage von oben könnte also erweitert werden: «Ich bin nur eine Maschine, aber ...»

Die Maschinenethik kann ebenso dazu benutzt werden, Täuschungen (oder Lügen) hervorzubringen. Die Verwunderung darüber, die es immer wieder gibt, entspricht einem Missverständnis (oder einer Verkürzung) in Bezug auf philosophische bzw. wissenschaftliche Ethik überhaupt (Pieper, 2007). Deren Sinn und Zweck ist in erster Linie, wie bei jeder Wissenschaft, Erkenntnisgewinn. Die Maschinenethik kann moralische Maschinen entwickeln und untersuchen, aber genauso unmoralische (im definierten Sinne). Wenn die Ethik das Moralische und das Unmoralische untersucht, das Gute und das Böse, Gerechtigkeit und Ungerechtigkeit etc., dann untersucht die Maschinenethik die moralische und die unmoralische Maschine, und nicht nur das, sie bringt sie auch hervor, weil sie wie die Wirtschaftsinformatik eine Gestaltungsdisziplin ist, die ein Artefakt schafft, um dieses zu erforschen und zu verbessern (Bendel, 2019d). Ob die unmoralische Maschine das Labor verlassen soll, ist eine andere Frage und muss im Einzelfall entschieden werden.

Ronald C. Arkin beobachtet Täuschungsverhalten im Tierreich, etwa bei Eichhörnchen, um es dann auf Prototypen zu übertragen (Wagner & Arkin, 2010). Ziel sind Maschinen, die im Kriegsfall den Gegner verwirren, ablenken und zu den falschen Ressourcen führen, die täuschen und betrügen, für einen Zweck, den man so oder so beurteilen kann. Die Finanzierung der Arbeiten erfolgt zumindest teilweise durch das Pentagon (Misselhorn, 2018). Für die Maschinenethik stellt all dies (die scheinbar unmoralische Stossrichtung) kein Problem dar: Es geht ihr in erster Linie darum, das «Verhalten» der Maschinen im moralischen Sinne zu determinieren. Die Beschränkung, die bei Menschen so unangenehm auffallen mag – oben wurde von Fundamentalisten gesprochen –, ist in vielen Situationen, in denen die Maschine bestehen muss, ausgesprochen hilfreich. Der amerikanische Robotiker und Maschinenethiker entwickelt auch Systeme mit einem künstlichen Gewissen – und diese können etwa, wie er hervorhebt, das Kriegsvölkerrecht strikt einhalten (Misselhorn, 2018). Seiner Argumentation muss man keineswegs folgen – dennoch kann man sagen, dass er zur Entwicklung und Vielgestaltigkeit der Maschinenethik beigetragen hat.

Wiederum gibt es auch Unterschiede zwischen den Kulturen, die sich in der Disziplin und in ihrem Gegenstand niederschlagen. Zunächst fällt auf, dass die Maschinenethik von den USA (Anderson & Anderson, 2011) und von Europa (Pereira & Saptawijaya 2016; Bendel 2019d) dominiert wird. In Japan gibt es kaum eine einschlägige Tradition. In Bezug auf den Gegenstand kommen in unterschiedlichen Gesellschaften durchaus unterschiedliche Vorstellungen zum Tragen. Allerdings schreibt die Maschinenethik an sich kaum bestimmte moralische Regeln vor – sie versucht vielmehr, zumindest wenn sie eher technisch verstanden wird, diese in die Maschinen zu bringen (Bendel, 2019). Es bedarf in jeder Kultur einer zusätzlichen Diskussion, wenn es um die Praxis geht, und selbst diese kann bestimmte Probleme nicht lösen. So kann zwar eine Ethikkommission bestimmte Vorgaben zu selbstfahrenden Autos machen, aber damit holt sie noch lange nicht die Bevölkerung ab. Und selbst wenn die Vorgaben auch von der Bevölkerung mitbestimmt wurden, bleibt am Ende doch der Einzelfall, der die Betroffenen u.U. nicht zufriedenstellt – oder ihnen sogar elementare Rechte nimmt. Bei sozialen Robotern ist dies kaum anders. Natürlich bleiben die Instrumente und Methoden der normativen Ethik, und insbesondere müssen hier die Bereichsethiken der Roboterethik, Technikethik und Informationsethik mit der Maschinenethik zusammenarbeiten.

8.1.3. Informationsethik

Die Informationsethik hat die Moral der Informationsgesellschaft zum Gegenstand. Als klassische Bereichsethik beschäftigt sie sich u.a. mit moralischen Fragen, die sich bei der Nutzung von Informations- und Kommunikationstechnologien, Informationssystemen, KI-Systemen und Robotern ergeben (Bendel, 2019f). Es geht keinesfalls nur um (Implikationen der Nutzung von) Informationen und Daten, sondern um alle Aspekte der Nutzung von damit verbundenen Technologien. In diesem Sinne vereint die Disziplin Computerethik, Netzethik und Medienethik (oder Neue-Medien-Ethik) in sich, wie etwa Capurro und Bendel (2012b) es formuliert haben. Ein neuerer, stark uneinheitlich sowohl für die Disziplin wie den Gegenstand gebrauchter Begriff ist «digitale Ethik», der sich von engl. «digital ethics» ableitet. Er kann in der deutschsprachigen Philosophie als Fremdkörper angesehen werden, der sich z.B. nicht in die Benennung der Bereichsethiken einfügt. Die Technikethik bestellt ein weites Feld – und kann sich im Prinzip ebenfalls Fragen der Informationsethik zuwenden.

Es gibt einige Darstellungen dazu, dass mobile Roboter uns ausspionieren oder potenziell ausspionieren können (Bendel, 2017b). Dies liegt zunächst einfach an ihrer Mobilität und den damit verbundenen Kameras und Sensoren, die nicht nur zur Objekt- und Hinderniserkennung, sondern auch zur Überwachung genutzt werden können (neben klassischen Kameras sind Lidar- und Radarsysteme zu nennen, Ultraschallsysteme und Wärmebildkameras). Weiter liegt es an ihren Aufgaben – beispielsweise nehmen sie als Serviceroboter solche in Dienstleistungsbereichen wahr (wie Transport und Sicherheit), wo sie ihre belebte oder unbelebte Umwelt erkennen und auswerten müssen, oder sie pflegen als Serviceroboter der besonderen Art einen engen Umgang mit Menschen oder Tieren.

Bei weit entwickelten sozialen Robotern ergeben sich noch spezifische Probleme. Sie verfügen womöglich über Sprach-, Stimm- und Gesichtserkennung, z.T. verbunden mit Emotionserkennung (Bendel, 2018e). Dadurch können sie etwa Alter, Geschlecht und Sprache bestimmen sowie Aussagen zur Gesundheit und Bildung machen – und zur Gefühlslage, womöglich in ihrem Auf und Ab über eine gewisse Zeit (Cavallo et al., 2018). Damit werden bereits persönliche, u.U. besonders schützenswerte Daten generiert, was ein Thema für die Informationsethik und das Roboterrecht ist bzw. das Persönlichkeitsrecht tangiert. Tangiert wird u.U. die informationelle Autonomie, die mit der informationellen Selbstbestimmung im Recht korrespondiert; verletzt werden gegebenenfalls Privat- und Intimsphäre. Im BESTBOT-Projekt an der Hochschule für Wirtschaft FHNW wurde die Verwendung von Gesichtserkennung problematisiert (Bendel et al., 2020): Diese kann bei Software- und Hardwarerobotern einerseits dazu beitragen, Emotionen zu erkennen und Schaden abzuwenden, andererseits erzeugt sie selbst potenziell Schaden.

Die Motive für Spionage und Überwachung sind vielfältig. Die Hersteller und Betreiber können sich Daten und Informationen beschaffen, um den sozialen Roboter zu verbessern, um den Kunden zufriedener zu machen oder um ihn zu weiteren Käufen anzuregen. Hinweise darauf hat auch die Befragung der Hersteller, Entwickler und Anwender (s. Teil zur Herstellerbefragung) erbracht. Polizei und Geheimdienste mögen sich aus dem engen, vertrauten Umgang zwischen sozialen Robotern und Nutzern relevante Informationen erhoffen, um Schaden von Personen, Einrichtungen und dem Staat abzuwenden oder den Nutzern selbst Schaden zuzufügen, etwa durch Erpressung. Die Medien können sich interessante Aufnahmen visueller und auditiver Art beschaffen und damit ihre Berichterstattung anreichern. Im Pflegebereich ist Überwachung auch eine Notwendigkeit, um Hilfe zu ermöglichen und Leben zu retten.

Die sozialen Roboter, die man in Einkaufszentren oder in Transporthubs wie Flughäfen antrifft, weisen selten auf DSGVO, DSGVO oder andere Datenschutzbestimmungen hin. Ihr Einsatz ist in rechtlicher Hinsicht, aber auch in ethischer Hinsicht fragwürdig. So wird das Transparenzgebot missachtet und die informationelle Autonomie verletzt. Es besteht, so könnte man argumentieren, ein Ungleichgewicht zwischen Anbieter und Nutzer, zwischen Beobachter und Beobachtetem (der nur bedingt selbst Beobachter sein kann), was auch eine Machtfrage ist (Bendel, 2012c; Bendel, 2017b). Dies trifft nicht nur auf die Ebene der Datenerfassung zu, sondern auch auf die Gestaltung und Umsetzung des ganzen Roboters. So weiss man in der Regel als Nutzer nicht, ob man mit einem «reinen» Roboter spricht oder ob es sich um ein Hybrid (hier ebenfalls manchmal Cyborg genannt) aus Mensch und Maschine handelt. Dies kann sich auf die Bewegung (Fernsteuerung vs. autonome Bewegung) genauso beziehen wie auf die Kommunikation.

8.2. Weitere Erkenntnisse und Analysen

8.2.1. Erkenntnisse aus der Befragung

Im Rahmen des Projekts wurde von Oliver Bendel und Jeanne Kreis eine Befragung von Herstellern, Entwicklern und Anwendern durchgeführt. Auffällig war, dass diese in der Mehrheit der Meinung waren, dass soziale Roboter durchaus Emotionen und Empathie zeigen sollten. Zu ähnlichen Ergebnissen kam die Psychologenbefragung (Kapitel 6). Ein Kernargument der Hersteller, Entwickler und Betreiber war die Funktionalität. Wenn man – womöglich humanoid gestaltete – soziale Roboter in bestimmten Bereichen wie Schulen oder Pflegeheimen einsetzt, wenn sie die eine oder andere Aufgabe und Rolle übernehmen, wäre es kontraproduktiv, wenn sie völlig anders als der normale, menschliche Aufgabeneigner und Rolleninhaber agieren würden. Um in der auf Menschen zugeschnittenen Welt zu funktionieren, benötigen sie daher auch menschliche Eigenschaften. Als Betrug oder Täuschung – Konzepte, auf die Philosophen schnell kommen würden – haben die meisten Befragten dies nicht angesehen. Natürlich kann dies auch eine Art Selbstschutz sein, da man eben in diesem Geschäftsfeld tätig ist.

Ebenfalls deutlich wurde, dass die Schweiz ins Hintertreffen gerät bei der Entwicklung von sozialen Robotern und der Austausch zwischen forschenden Hochschulen und marktorientierten Unternehmen nicht durchgehend gelingt. Einrichtungen wie die ETH Zürich, zumindest einzelne Institute, werden insbesondere von einem Ansprechpartner als abgehoben und wenig entgegenkommend erlebt. Dennoch gibt es in der Schweiz interessante Ansätze sowohl mit Blick auf die Herstellung von sozialen Robotern als auch auf die Entwicklung und das Angebot von Software und Dienstleistungen. Zudem zeigt man sich experimentierfreudig in der Anwendung, was ein Pluspunkt in der Schweiz ist, wo freilich auch immer wieder Budgets für Innovation zur Verfügung stehen.

Die Abgehobenheit bezieht sich auch auf die «Ethikindustrie», die derzeit in ganz Europa und weltweit aufblüht. Man überbietet sich im Entwickeln von Leitlinien, die die Experten im Allgemeinen zwar als wichtig angesehen, deren Allgemeinheit sie aber beanstandet haben. Philosophische und psychologische Erwägungen auf hohem akademischem Niveau nutzen den Herstellern und Entwicklern wenig, zumal sie zugleich politische und gesetzliche Anforderungen beachten müssen. Eine Einschätzung der Autoren ist zugleich, dass ethische Fachbegriffe kaum gekannt und verstanden werden, und die neu entstehenden Weiterbildungsangebote und medialen Aufarbeitungen tragen kaum zur Klärung bei, ganz im Gegenteil. Besonders in kleineren Unternehmen scheitert die professionelle Auseinandersetzung mit ethischen Fragen oft auch am Budget.

Jüngste Gespräche mit Herstellern haben – unabhängig von der eigentlichen Herstellerbefragung – Hinweise darauf geliefert, dass Krisen und Katastrophen wie Pandemien einen Schub für die Soziale Robotik bedeuten. So sind Transport-, Reinigungs-, Pflege- und Desinfektionsroboter gefragt. Einer der befragten Hersteller hat im Mai 2020 eine Zusage der EU erhalten, die beim Einsatz eines Pflegeroboters im Bereich der Desinfektion (Türfallen, Liftknöpfe) unterstützt. Auf der Website <https://dih-hero.eu/awarded-projects/> heisst es: «Currently, Europe and countries all over the world are facing a global pandemic. Together with its extensive Robotics in Healthcare European network DIH-HERO decided to support the fight against COVID-19 by providing €1,000,000 for robotic technologies that can be deployed timely, in order to support healthcare professionals and save lives by satisfying a current clinical demand or need. Although the submission period of only one week has been very short for this special call 146 proposals have been submitted before the deadline on the 17th of April 2020. After the first evaluation of all these great proposals and fantastic ideas, the evaluators invited the Top 20 highest ranked proposals for a 10-minute video interview in order to get to the final ranking of proposals to be awarded.» (Website EU, o.D.).

8.2.2. Aussagen in Leitlinien und Richtlinien

In den letzten Jahren sind, zusätzlich zu den DIN- und ISO-Normen, umfangreiche Regelkataloge, Leit- und Richtlinien zum Einsatz von künstlicher Intelligenz und intelligenten und mobilen Robotern entstanden. Dabei haben sich ganz unterschiedliche Organisationen engagiert, etwa Berufsverbände wie IEEE, Länderverbände wie die EU und die Wirtschaft (Google, Microsoft). Die Gründe waren jeweils ganz unterschiedlich, wobei ein Blick auf die Zusammensetzung der Gremien selbst bei scheinbar politischen und philosophischen Gefässen eine klare kommerzielle Ausrichtung offenbart. Die Tab. 22 und 23 «Übersicht über Richt- und Leitlinien» im Anhang stellen dar, wie die Begriffe «soziale Roboter», «Empathie» und «Emotionen/Gefühle» in den Leitlinien von Unternehmen wie Google, McKinsey oder Microsoft berücksichtigt werden. Es zeigte sich, dass insbesondere soziale Roboter kein Thema sind. Auch Empathie und Emotionen werden nur selten thematisiert.

8.2.3. Empathie und Emotionen in den Konferenzbänden «Social Robotics»

Seit 2009 werden im Rahmen der International Conference on Social Robotics (ICSR) jährlich die Konferenzbände «Social Robotics» veröffentlicht. Beiträge zu Empathie und Emotionen zeigen, dass mit sozialen Robotern auch die Interaktion zwischen Mensch und Maschine stärker in den Fokus rückt. Zu Beginn der 2010er-Jahre beschäftigte sich die Forschung insbesondere mit dem Design des Roboters, seiner Beschaffenheit und seinem Ausdruck. Mit der verbesserten Gesichts- und Emotionserkennung intensivierte sich aber auch die Mensch-Maschine-Interaktion und eröffnete der Forschung neue Anwendungsbereiche für soziale Roboter. In den letzten Jahren tauchten sie vermehrt in der Pflege oder an öffentlich zugänglichen Orten auf. Die kontinuierlich verbesserte Interaktionsfähigkeit der Roboter ermöglichte auch Studien mit autistischen Kindern und Schuleinsätze von Tutorrobotern. Getrieben von der Frage, wie Roboter Emotionen und menschliches Verhalten nicht nur erfassen, sondern daraus auch Rückschlüsse auf das Wohl einer Person ziehen können, ist die Emotionserkennung sozialer Roboter heute ein schnell wachsender Forschungsbereich. Die Tab. 24 «Übersicht über die einschlägigen Inhalte der Konferenzbände der Konferenz ICSR» im Anhang zeigt diese Entwicklung auf. Auf die systematische Auswertung weiterer einschlägiger Konferenzen wie der Robophilosophy musste aus Zeit- und Platzgründen verzichtet werden. Diese wird in der Studie aber immer wieder erwähnt.

8.3. Szenarien zum Einsatz sozialer Roboter

In diesem Abschnitt werden Kurzsznarien in den drei Bereichen «Gesundheit», «Öffentlich zugängliche Orte» und «Private Haushalte» (diese werden hier aufgeteilt) und «Bildung» entworfen und jeweils schlagworthaft aus ethischer Sicht eingeordnet, wobei Chancen und Risiken sowie mögliche Ansätze (wie Lösungen und Verbesserungen) genannt werden, aus Roboterethik, Maschinenethik und Informationsethik heraus – deren Begriffe auf die Fälle angewandt werden –, aber auch in technischer oder rechtlicher Hinsicht. Die Ansätze sind nicht als Empfehlungen zu verstehen – diese können erst bei Kenntnis aller Umstände gegeben werden.

Die Szenarien entstammen sowohl der Vorstellung der Verfasser als auch der Beobachtung und Beschreibung tatsächlicher Einsätze (die möglichst genannt werden). Es geht zum einen darum, über konkrete Sachverhalte sprechen und von diesen ausgehen zu können. Zum anderen soll eine Grundlage für eine ethische und anderweitige Diskussion – in diesem Bericht und darüber hinaus – geschaffen werden.

Tab. 14: Szenarien im Bereich «Bildung»

Szenarien	Chancen	Risiken	Ansätze
Ein Lehrerroboter lobt eine Schülerin oder einen Schüler für eine Leistung.	Motivation	Täuschung, Betrug Erzeugung von Erwartungen	Transparenz: Der Roboter lobt nach klaren Kriterien. Gleichbehandlung: Der Roboter behandelt alle Schüler und Schülerinnen gleich.
Ein Lehrerroboter zeigt Trauer oder Freude, als er auf ein bestimmtes Thema zu sprechen kommt.	Wirkung	Täuschung, Betrug Erzeugung von Erwartungen	Transparenz: Der Roboter simuliert Emotionen nach klaren Kriterien. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft.
Ein Lehrerroboter überwacht die Schüler mit Hilfe von Kameras und Bewegungsmeldern im Klassenraum, um ihre Aktivität festzustellen.	Leistungssteigerung	Überwachung Gefährdung der informationellen Autonomie	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre.
Ein Tutorroboter übt – sozusagen auf Augenhöhe und mit kollegialen Sprüchen – mit den Schülern oder Studenten.	Motivation	Täuschung, Betrug Erzeugung von Erwartungen	Transparenz: Der Roboter lehrt nach klaren Kriterien.
Ein Schülerroboter dient als Avatar eines kranken Schülers, sodass dieser am Unterricht teilnehmen kann. (Beispiel: AV1)	Teilnahme am Unterricht Teilhabe am Leben Keine Ansteckungsgefahr	Gewöhnung an Stellvertreter	Maschinelles soziales Verhalten: Der Roboter drückt die Emotionen des Schülers aus. Transparenz: Der Lehrer macht den Schülerinnen und Schülern immer wieder klar, dass der Schüler zuhause und krank ist.
Ein sozialer Roboter bringt Schülern Empathie und Emotionen bei. (Beispiel: Moxie)	Ausgleich fehlender Betreuung	Manipulation durch Hersteller	Transparenz: Der Roboter macht klar, dass er Empathie und Emotionen nur simuliert. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft.

Eine wiederkehrende Chance ist die Motivation, ein wiederkehrendes Risiko sind Täuschung und Betrug. Transparenz, maschinelle Moral und Datenschutzrichtlinien scheinen Ansätze und Lösungen zu sein.

Tab. 15: Szenarien im Bereich «Öffentlich zugängliche Orte»

Szenarien	Chancen	Risiken	Ansätze
Ein Sicherheitsroboter steht auf einem öffentlichen Platz als Ansprechpartner für Probleme zur Verfügung und zeigt für diese Verständnis. (Beispiel: REEM, Dubai)	Sicherheitsgefühl	Täuschung, Betrug Erzeugung von Erwartungen	Transparenz: Der Roboter agiert und reagiert nach klaren Kriterien.
Ein Sicherheitsroboter zeigt Ärger, wenn Passanten bei Rot über die Ampel gehen.	Regel Einhaltung	Täuschung, Betrug Erzeugung von Erwartungen Bevormundung	Transparenz: Der Roboter agiert und reagiert nach klaren Kriterien.
Ein Sicherheitsroboter überwacht einen öffentlichen Platz mit Hilfe von Gesichtserkennung und meldet Verdächtige bei der Polizei. (Beispiel: K5, Silicon Valley)	Sicherheitsgefühl Verfolgung von Straftaten	Überwachung Gefährdung der informationellen Autonomie	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Roboterquote: Die Zahl eingesetzter Roboter ist in einem bestimmten Gebiet beschränkt. Verbot: Roboter sind in einem bestimmten Gebiet verboten.
Ein Sicherheitsroboter zeigt Vorurteile gegenüber bestimmten Personen. (Beispiel: MOME)	Echtheit	Diskriminierung Bias Ungerechtigkeit	Transparenz: Der Roboter macht deutlich, dass er Vorurteile zeigt. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft. Verbot: Roboter dürfen keine Vorurteile zeigen.
Ein Beraterroboter in der Shoppingmall hat eine weisse Oberfläche. (Beispiel: Pepper, Glattzentrum)	Wirkung	Diskriminierung Bias	Transparenz: Der Roboter hat keine Hautfarbe, sondern medizinisches Weiss.
Ein Beraterroboter in der Shoppingmall hat eine weibliche Stimme.	Wirkung	Diskriminierung Bias	Transparenz: Die Stimme wirkt für viele Menschen vertrauenswürdig und angenehm.
Ein Beraterroboter in der Shoppingmall umarmt die Kunden. (Beispiel: Pepper, Glattzentrum)	Motivation	Verletzung von Persönlichkeitsrechten Ansteckungsgefahr	Transparenz: Der Roboter bewegt sich nach klaren Vorgaben. Reinigung: Der Roboter wird regelmäßig gereinigt und desinfiziert.
Ein Beraterroboter in der Shoppingmall wendet Gesichts- und Stimmerkennung mit Emotionsanalyse an.	Personalisierung	Überwachung Gefährdung der informationellen Autonomie Verletzung von Persönlichkeitsrechten	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Transparenz: Der Roboter macht deutlich, dass er Gesichts- und Stimmerkennung mit Emotionsanalyse anwendet. Verbot: Der Roboter darf keine Gesichts- und Stimmerkennung mit Emotionsanalyse anwenden.

Szenarien	Chancen	Risiken	Ansätze
Ein Beraterroboter in der Shoppingmall betreut Kinder. (Beispiel: Pepper, Glattzentrum)	Betreuung	Überwachung Gefährdung der informationellen Autonomie Beeinflussung	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft.
Ein Companion Robot zeigt Vorurteile gegenüber bestimmten Personen. (Beispiel: MOME)	Unterhaltung	Ungerechtigkeit	Transparenz: Der Roboter macht deutlich, dass er Vorurteile zeigt. Maschinelle Moral: Der Roboter erhält die Regel, keine Vorurteile zu zeigen.
Ein Transportroboter hat Augen und Mund, erzeugt Töne und erwirbt damit Sympathien. (Beispiel: Relay, Swisslog)	Kooperation und Kollaboration	Erzeugung von Erwartungen	Transparenz: Der Roboter macht deutlich, dass er Emotionen erzeugt.

Eine wiederkehrende Chance ist die Wirkung (so kann ein sozialer Roboter in Shoppingmalls Kunden anziehen). Wiederkehrende Risiken sind Täuschung und Betrug, Überwachung, Gefährdung der informationellen Autonomie und Verletzung von Persönlichkeitsrechten. Transparenz, maschinelle Moral und Datenschutzrichtlinien scheinen Ansätze und Lösungen zu sein.

Tab. 16: Szenarien im Bereich «Gesundheit»

Szenarien	Chancen	Risiken	Ansätze
Ein Pflegeroboter nimmt den Patienten über Kameras und Mikrofone auf. (Beispiel: Pepper; Lio, Schweiz)	Sicherheitsgefühl Hilfe bei Notlagen	Überwachung	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Transparenz: Der Roboter macht deutlich, dass er Gesichts- und Stimmerkennung mit Emotionsanalyse anwendet. Verbot: Der Roboter darf keine Gesichts- und Stimmerkennung mit Emotionsanalyse anwenden.
Ein Pflegeroboter scheint Interesse an dem zu zeigen, was der Patient zu ihm sagt. (Beispiel: Lio, Schweiz)	Gefühl der Zweisamkeit	Täuschung, Betrug	Transparenz: Der Roboter macht deutlich, dass er Empathie und Emotionen nur simuliert.
Ein Pflegeroboter zeigt Trauer oder Freude, als der Patient auf ein bestimmtes Thema zu sprechen kommt.	Trost	Täuschung, Betrug	Transparenz: Der Roboter macht deutlich, dass er Empathie und Emotionen nur simuliert.
Ein Pflegeroboter ist für mehrere Personen zuständig, die Eifersucht entwickeln. (Beispiel: Lio, Schweiz)	Wirkung	Ungerechtigkeit Unfrieden	Transparenz: Der Roboter macht deutlich, dass er für alle da ist. Maschinelle Moral: Der Roboter erhält die Regel, alle gleich zu behandeln.
Ein Pflegeroboter ist für mehrere Personen zuständig, die Angst haben, dass er Geheimnisse weitergibt. (Beispiel: Lio, Schweiz)	Wirkung	Überwachung Gefährdung der informationellen Autonomie Verletzung von Privatsphäre und Intimsphäre	Transparenz: Der Roboter macht deutlich, dass er Geheimnisse nicht weitergibt. Maschinelle Moral: Der Roboter gibt keine Geheimnisse weiter, wenn er solche erkannt oder wenn der Patient gesagt hat: «Das ist ein Geheimnis.»
Ein Pflegeroboter mit sexuellen Assistenzfunktionen steht Patienten für Partnersexualität zur Verfügung.	Wirkung Sexuelles Wohlbefinden	Verletzung von Privatsphäre und Intimsphäre Verletzung der Menschenwürde Überbeanspruchung	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Maschinelle Moral: Der Roboter macht klar, dass er nur eine Maschine ist. Maschinelle Moral/Maschinelles soziales Verhalten: Der Roboter weist darauf hin, dass er sexuell leistungsfähiger als der Mensch ist. Technische Sperre: Der Roboter schaltet sich regelmässig aus, um Überbeanspruchung des Partners zu vermeiden.
Ein Pflegeroboter mit sexuellen Assistenzfunktionen steht Patienten für erotische Gespräche zur Verfügung.	Sexuelles Wohlbefinden	Verletzung von Privatsphäre und Intimsphäre Verletzung der Menschenwürde	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre.

Szenarien	Chancen	Risiken	Ansätze
Ein Pflegeroboter mit sexuellen Assistenzfunktionen steht Patienten für das Reinigen und Reichen von Sexspielzeug zur Verfügung.	Sexuelles Wohlbefinden	Verletzung von Privatsphäre und Intimsphäre Ansteckungsgefahr bei nicht sachgerechter Reinigung	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre.
Ein Pflegeroboter im Krankenhaus oder Pflegeheim hat eine weiße Oberfläche. (Beispiel: Pepper; Robear)	Wirkung	Diskriminierung Bias	Transparenz: Der Roboter hat keine Hautfarbe, sondern medizinisches Weiss.
Ein Therapieroboter dient als Tierersatz, mit dem erreicht wird, dass sich die Dementen mit ihm beschäftigen, ihn streicheln und mit ihm reden. (Beispiel: Paro, Schweiz)	Animieren/Stimulieren von Emotionen	Täuschung, Betrug	Patientenverfügung: Es wird darüber bestimmt, ob Pflegeroboter zum Einsatz kommen sollen.
Ein Therapieroboter dient autistischen Kindern als Ansprechpartner und Projektionsfläche. (Beispiel: QTrobot)	Animieren/Stimulieren von Emotionen und Aktionen	Täuschung, Betrug	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist.
Ein Schlafroboter liegt zusammen mit dem Nutzer im Bett. (Beispiel: Somnox Sleep Robot)	Hilfe beim Einschlafen	Täuschung, Betrug	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Patientenverfügung: Es wird darüber bestimmt, ob Schlafroboter zum Einsatz kommen sollen.
Ein Umarmungsroboter umarmt den Nutzer oder wird von diesem umarmt. (Beispiel: PR2/HuggieBot, Hugvie, HUGGIE (in Planung))	Trost Körperliche Nähe trotz Pandemie	Täuschung, Betrug	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft.

Wiederkehrende Chancen sind Wirkung (so kann ein sozialer Roboter trösten und locken) und sexuelles Wohlbefinden. Wiederkehrende Risiken sind Täuschung und Betrug, Überwachung, Verletzung von Privatsphäre und Intimsphäre sowie Verletzung der Menschenwürde. Transparenz, maschinelle Moral und Datenschutzrichtlinien scheinen Ansätze und Lösungen zu sein.

Tab. 17: Szenarien im Bereich «Private Haushalte»

Szenarien	Chancen	Risiken	Ansätze
Ein Companion Robot wird von den Kindern als Freund wahrgenommen, und sie kommen mit ihren Problemen zu ihm. (Beispiel: Pepper, Japan)	Ansprechbarkeit Betreuung	Täuschung, Betrug Verfehlte Reaktionen Beeinflussung	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist.
Ein Companion Robot bevorzugt durchgehend ein Familienmitglied.	Bindung	Ungerechtigkeit	Maschinelle Moral: Der Roboter erhält die Regel, alle gleich zu behandeln.
Ein Companion Robot zeigt Vorurteile gegenüber bestimmten Personen. (Beispiel: MOME)	Unterhaltung	Ungerechtigkeit	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Maschinelle Moral: Der Roboter erhält die Regel, keine Vorurteile zu zeigen.
Ein Companion Robot zeigt Trauer, wenn es Streit in der Familie oder der Beziehung gibt.	Bestätigung	Täuschung, Betrug Falsche Referenz	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist.
Ein Companion Robot lobt und tadelt die Kinder für ihr Verhalten. (Beispiel: Moxie)	Motivation	Täuschung, Betrug Falsches Vorbild Beeinflussung	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Maschinelles soziales Verhalten: Der Roboter nimmt Lob oder Tadel der Eltern auf und gibt dieses weiter. Maschinelle Moral: Der Roboter übernimmt und spiegelt Normen und Werte einer Gesellschaft.
Ein Companion Robot erzählt den Eltern, was die Kinder gemacht und gesagt haben. (Beispiel: Moxie)	Sicherheitsgefühl	Überwachung Verletzung von Privatsphäre und Intimsphäre	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre.
Ein Companion Robot gibt Daten an die Betreiber oder Hersteller weiter.	Profit	Überwachung Verletzung von Privatsphäre und Intimsphäre	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre.
Der Nutzer ist verliebt in seinen weiblichen oder männlichen Sexroboter. (Beispiel: Harmony, USA)	Sicherheitsgefühl Gefühl der Zweisamkeit	Täuschung, Betrug Verlust sozialer Beziehungen	Desillusionierung (V-Effekt): Der Roboter beendet die Illusionierung des Nutzers, etwa indem er seine Maschinenhaftigkeit zeigt. Maschinelles soziales Verhalten: Der Roboter weist auf mögliche menschliche Partner hin.
Der Nutzer bevorzugt den Sexroboter, der Stereotype erfüllt. (Beispiel: Harmony, USA)	Sexuelles Wohlbefinden	Objektivierung Übertragung Bias	Desillusionierung (V-Effekt): Der Roboter beendet die Illusionierung des Nutzers, etwa indem er sein Aussehen verändert. Vielfalt am Markt: Der Roboter wird in Varianten angeboten.

Szenarien	Chancen	Risiken	Ansätze
Der Nutzer bevorzugt Liebespuppen oder Sexroboter, die von Stereotypen abweichen. (Beispiel: Yuki, BorDoll)	Sexuelle Vielfalt Abbau von Vorurteilen	Täuschung, Betrug	Vielfalt am Markt: Der Roboter wird in Varianten angeboten.
Der Nutzer bevorzugt Liebespuppen oder Sexroboter, die ein kindliches Äusseres haben.	Therapeutische Wirkung	Zunahme des Bedürfnisses Übertragung	Empirische Forschung
Der Nutzer zeigt gewalttätiges Verhalten gegenüber dem Sexroboter.	Keine Übertragung	Übertragung	Technische Sperre: Der Roboter entzieht sich gewalttätigem Verhalten. Maschinelle Moral: Der Roboter widersetzt sich.
Ein Sexroboter nimmt seinen Besitzer mit Kameras und Mikrofonen auf und lernt dessen Vorlieben kennen.	Sexuelles Wohlbefinden Personalisierung	Täuschung, Betrug Überwachung Verletzung von Privatsphäre und Intimsphäre	Datenschutzrichtlinien: Der Roboter wahrt menschliche Privatsphäre. Private Cloud: Persönliche Daten sind geschützt.
Ein Schlafroboter liegt zusammen mit dem Nutzer im Bett. (Beispiel: Somnox Sleep Robot)	Hilfe beim Einschlafen	Täuschung, Betrug	Transparenz: Der Roboter macht klar, dass er nur eine Maschine ist. Patientenverfügung: Es wird darüber bestimmt, ob Schlafroboter zum Einsatz kommen sollen.

Wiederkehrende Chancen sind Sicherheitsgefühl und sexuelles Wohlbefinden, wiederkehrende Risiken sind Täuschung und Betrug, Überwachung und Verletzung von Privatsphäre und Intimsphäre. Transparenz, maschinelle Moral und Datenschutzrichtlinien scheinen Ansätze und Lösungen zu sein.

Zusammenfassend kann gesagt werden, dass als Chancen in allen Anwendungsgebieten Motivation, Wirkung, sexuelles Wohlbefinden und Sicherheitsgefühl ausgemacht werden können. In allen Anwendungsbereichen wiederkehrende Risiken sind Täuschung und Betrug, Überwachung, Gefährdung der informationellen Autonomie und Verletzung von Persönlichkeitsrechten, Verletzung von Privatsphäre und Intimsphäre sowie Verletzung der Menschenwürde. Mögliche Ansätze – nicht zu verwechseln mit Empfehlungen – sind Transparenz, maschinelle Moral und Datenschutzrichtlinien. In Kapitel 8.5 findet eine weitere Zusammenfassung statt, einschliesslich verschiedener Schlussfolgerungen.

8.4. Ökonomie und Ethik

Ökonomische Aspekte sozialer Roboter wurden bereits in Kapitel 7 angesprochen. Insbesondere ging es dort um volkswirtschaftliche Implikationen. Ökonomie und Ethik haben allerdings Bezüge, auf die noch nicht genügend eingegangen wurde. Zwar wurden durchaus Begriffe von Ethik und Recht genannt. Aber spezifische Abhängigkeiten und Suchtformen, die die Wirtschaft erzeugt, sollten noch genauer betrachtet werden.

Die Ethik hat seit 2010 einen enormen Schub und ein erhebliches mediales und öffentliches Interesse erfahren. Allerdings wird der Begriff sehr unterschiedlich verstanden, und Unternehmen, Beratungseinrichtungen und auch Hochschulen instrumentalisieren sie zu ihrem Zweck. Es geht um Regulierungsvermeidung, Greenwashing und Gewinnmaximierung. Man kann von einer regelrechten Ethikindustrie sprechen, und so wie sich Ethikkommissionen oft dadurch auszeichnen, dass sich wenige (philosophische) Ethiker in ihnen befinden, so mischen in der Ethikindustrie ganz unterschiedliche Vertreter mit. Die Ethik als wissenschaftliche Disziplin kann letztlich Schaden nehmen, wenn es ihr nicht gelingt, sich abzugrenzen (Bendel, 2014) und auf ihren anerkannten Methoden – wie der logischen, diskursiven und dialektischen – zu bestehen (Pieper, 2007).

Durch soziale Roboter kann man im öffentlichen und kommerziellen Bereich Personal einsparen (darauf wurde bereits im volkswirtschaftlichen Teil eingegangen, s. Kapitel 7), die Kundenbindung verstärken und Beziehungen und Abhängigkeiten herstellen. Dies ist nicht zuletzt ein Thema der Wirtschaftsethik. Beim erstgenannten Thema muss genau hingeschaut werden. Wird das Personal vielleicht eher ergänzt oder entlastet? Und wenn es ersetzt und eingespart wird: Kommt es anderweitig unter? Nicht zuletzt: War die Arbeit, die bisher ausgeführt wurde, der Menschenwürde zuträglich, obschon sie auch von einem Roboter erledigt werden kann, also womöglich repetitiv und automatisierbar ist? Automatisierung ist derzeit vor allem in den Bereichen Sicherheit, Reinigung und Transport möglich, also dort, wo eher klassische Serviceroboter als spezifische soziale Roboter zugange sind. Aber das könnte sich in den nächsten Jahrzehnten ändern.

Beziehungen und Abhängigkeiten sind ein weites Feld. Bereits in der zweiten Hälfte des 20. Jahrhunderts und um die Jahrtausendwende gab es etliche Geschäftsmodelle im digitalen Bereich, die eben darauf gründeten. So stellte Artificial Life in den 2000er-Jahren eine virtuelle Freundin für das Handy vor, für die man virtuelle Geschenke für reales Geld kaufen konnte. Das Verhalten des Avatars bzw. Programms stärkte ganz gezielt die Bindung. In dieser Tradition ist auch die Gatebox mit ihrem Hologrammädchen zu sehen (Bendel, 2019c). Dieses schickt dem Nutzer in seiner Abwesenheit Nachrichten, mit der Behauptung, es würde ihn vermissen. Soziale Roboter können im Prinzip zu ähnlichen Strategien und Taktiken greifen. Neben der Wirtschaftsethik, die danach fragt, ob solche Strategien und Taktiken zur Kundenbindung ein probates Mittel sind, ist wiederum die Roboterethik gefragt.

Besonders zu beleuchten sind in diesem Zusammenhang die Herstellung und der Einsatz von Liebespuppen und Sexrobotern (Bendel, 2020). Intensiv wurde seit der Jahrtausendwende die Frage diskutiert, ob diese speziellen sozialen Roboter Frauen zu Objekten machen, Stereotype stärken und fragwürdige Verhaltensweisen bei den Nutzern nach sich ziehen, die sich auch im Verhältnis zu Frauen bemerkbar machen (Richardson, 2016). Weniger im Fokus war, was die Beziehungen für den Mann (oder die Frau) bedeuten, ob hier Technikfirmen über ihre Produkte gezielt die Persönlichkeit der Kunden beeinflussen, ihre Psyche ansprechen und verändern, Emotionen zu Artefakten hervorrufen, Bindung und Abhängigkeiten herstellen, um am Ende Serviceleistungen und Nachfolgeprodukte verkaufen zu können.

Insgesamt ist die Frage, ob hier nicht eine künstliche Nachfrage erzeugt wird, indem man ein künstliches Angebot im mehrfachen Sinne schafft. In einem Interview mit brand eins sagte Mathias Binswanger, der Autor von Kapitel 7, die Menschen lebten nicht in einer Bedürfnisdeckungs-, sondern in einer Bedürfnisweckungsgesellschaft (Wittrock, 2019). Soziale Roboter würden entsprechend eine Nachfrage abdecken, die es ohne sie gar nicht gegeben hätte. Dem kann man mit Blick auf Liebespuppen und Sexmaschinen entgegenhalten, dass diese ein alter Menschheitstraum sind und sie keinesfalls unmittelbar an wirtschaftliche Interessen gekoppelt waren, eher an Macht- und Sexfantasien (Bendel, 2020). Aber auch bei Pflege- und Therapierobotern würde die Erklärung zu kurz greifen, und es sind eben gerade nicht die Konzerne, die versuchen, eine künstliche Nachfrage zu erzeugen, sondern es sind Start-ups und Spin-offs, die versuchen, einen entstehenden Bedarf zu befriedigen, mit kleinen, nicht immer überzeugenden Lösungen (Bendel, 2018a).

8.5. Erkenntnisse und Fazit

Soziale Roboter sind – um die Definition von oben nochmals in verkürzter Weise zu wiederholen – sensomotorische Maschinen, deren Zweck darin besteht, in Interaktion mit einer Person durch Imitation eines Lebewesens und dessen Verhaltens bestimmte kognitive und affektive Zustände sowie Handlungen auszulösen. Sie wurden, so die ergänzende Formulierung, für den Umgang mit Menschen und Tieren geschaffen. Die Simulation des Sozialen schafft mannigfache Möglichkeiten der Ergänzung, Unterstützung und Unterhaltung. Aber sie lädt auch dazu ein, dass wir uns einer Illusion hingeben und eine Fiktion annehmen müssen, dass wir mit Artefakten interagieren und kommunizieren, ohne uns dessen unbedingt bewusst zu sein oder ohne dass wir das wollen.

Vielleicht verstehen spätere Generationen manche Bedenken nicht, die hier geäußert wurden. Vielleicht ist die Kraft des Faktischen oder Normativen so gross, dass man, wie beim Smartphone, als Aussenseiter gilt, wenn man von einer Nutzung absieht, selbst wenn man gute Gründe hat. Vielleicht sind aber doch Vorbehalte oder Überlastungen vorhanden, und so wie es internetfreie Hotels gibt, könnte es roboterfreie Räume geben.

Wichtig ist letztlich die Wahlfreiheit, für die Gesellschaft, aber auch für den Einzelnen. Ob in Bildungseinrichtungen, in Gesundheitseinrichtungen, an öffentlich zugänglichen Orten oder in Haushalten – man sollte entscheiden dürfen, von wem man belehrt, bedient, begleitet und unterstützt wird. Zudem hat man ein Grundrecht darauf, dass man in allen Lebensabschnitten in ausreichender Art mit Menschen zusammentrifft. Das macht eben das Soziale des Menschen aus, und es lässt sich durch soziale Roboter keineswegs ersetzen.

Die Chancen und Risiken sozialer Roboter wurden in den Kurzscenarien und in anderen Kapiteln im Detail erarbeitet. Eine Zusammenführung birgt die Gefahr der Verflachung und Vereinfachung. Zugleich ist der Vorteil des Überblicks vorhanden, und es können weitere Schlussfolgerungen leichter fallen. Im Folgenden wird auf die Erkenntnisse aus den Kurzscenarien fokussiert.

- Soziale Roboter können zu Motivation und Leistungssteigerung beisteuern, ein Gefühl der Zweisamkeit entstehen lassen, Aufmerksamkeit und Trost spenden und Emotionen und Empathie fördern (Misselhorn, 2021). Sie können die Teilnahme am Leben (konkret auch die Präsenz in Einrichtungen und bei Veranstaltungen) ermöglichen und ein Ausgleich für fehlende Betreuung sowie Ansprechpartner sein. Sie können ein Sicherheitsgefühl vermitteln, die Regeleinhaltung begünstigen und die Verfolgung von Straftaten ermöglichen und fördern. Pflegeroboter mit sexuellen Assistenzfunktionen sowie Sexroboter mögen zur sexuellen Gesundheit und zum sexuellen Wohlbefinden beitragen. Weiter könnte eine zuverlässige Emotionserkennung dabei helfen, Erkrankungen wie Depression oder Demenz frühzeitig zu erkennen (Han et al., 2015).
- Dem steht gegenüber, dass man Gestaltungsformen, Handlungen und Entscheidungen von sozialen Robotern (wenn man diese Begriffe zulassen will) als Täuschung und Betrug werten kann, und dass sie Erwartungen erzeugen, die nicht erfüllt werden können. Im Einzelfall sind direkte Manipulation und Beeinflussung von Kunden, Bürgern und Heranwachsenden zu befürchten. Überwachung, die Zerstörung von Privat- und Intimsphäre, die Gefährdung der informationellen Autonomie und insgesamt die Verletzung von Persönlichkeitsrechten (bis hin zur Gefährdung der Menschenwürde) zeichnen gerade soziale Roboter aus, die über umfangreiche Sensorik und mächtige KI-Systeme verfügen (bzw. Zugang zu diesen haben). Zuweilen vermögen soziale Roboter auch Unfrieden hervorzurufen und Gerechtigkeit zu gefährden.
- Wie sich gezeigt hat, können soziale Roboter Vorurteile und Verzerrungen (Bias) mit sich bringen, durch ihre Akte, Sprechakte und ihre Gestaltung – des Äusseren und der Stimme. Sie können in Stereotype verfallen und Klischees bestärken, zum Nachteil von bestimmten Gruppen und Individuen. Weniger als bei Sprachassistenten besteht die Gefahr, dass die weibliche Stimme als Herabsetzung von Frauen verstanden wird – einfach weil soziale Roboter von Haus aus oft eine robotische oder neutrale Stimme haben; bekannte Beispiele hierfür sind Pepper und Cozmo. Auch bei Alexa und Co. wird freilich gerne ausser Acht gelassen, dass es vor allem darum geht, eine angenehme Stimme einzupflanzen – und weibliche Stimmen werden von vielen Menschen als angenehm und vertrauenswürdig empfunden.
- Die Ansätze zur Förderung von Chancen und zur Minderung von Risiken können auf unterschiedlichen Ebenen angesiedelt werden. Es stehen technische, gestalterische, ethische, rechtliche und politische Instrumente zur Verfügung. Fast alle Ebenen werden vom Transparenzgebot erfasst. In technischer und ethischer Hinsicht stehen Ethics by Design, Maschinenethik und Informationsethik mit ihren Methoden und Begriffen bereit. Man

kann die Programmierer und Entwickler sensibilisieren, den Maschinen bestimmte soziale und moralische Regeln geben, spezielle Patientenverfügungen entwickeln (Bendel, 2018a) und ethische Leit- und Richtlinien erlassen. Diese sind wiederum anschlussfähig an rechtliche Bestimmungen und politische Regulierungen. Im Raum stehen auch Verbote für soziale Roboter in bestimmten Räumen und eine Roboterquote. Zudem kann sich der Gesetzgeber dazu entschliessen, die Entwicklung oder die Anwendung besonders problematischer sozialer Roboter zu verbieten. Dies sollte er in einem Rechtsstaat freilich nur in Ausnahmefällen tun, und die Wissenschaftsfreiheit ist hoch zu bewerten, wenn auch nicht jedes Forschungsergebnis Einzug in den Alltag halten sollte.

Überhaupt sollte man nicht zu früh und nicht unumstösslich in die Entwicklung sozialer Roboter eingreifen. Man ist hier ganz am Anfang, und es können sich – auch wenn die Risiken durchaus deutlich zutage traten – Chancen aller Art ergeben. Dafür muss das Forschungsfeld möglichst offenbleiben, auch die Zusammenarbeit mit anderen Disziplinen, was sowohl eine rechtliche als auch ethische Dimension hat. Man versucht nämlich, die Regulierung in Robotik und KI von allen Seiten voranzutreiben. Die Unternehmen erhoffen sich von ethischer Regulierung die Vermeidung rechtlicher Regulierung und frönen dem Greenwashing. Die Staaten und Verbände erhoffen sich von ethischer Regulierung Standortvorteile und Konkurrenzfähigkeit. Die Begriffe, die hier ins Spiel gebracht werden, wie Responsible AI und vertrauenswürdige KI, sind häufig blosses Marketing, und es geht vergessen, dass Wissenschaft in erster Linie dem Erkenntnisgewinn dienen soll.

Damit ist der ethische Raum ein Stück weit ausgeleuchtet. Immer wieder wurden rechtliche Bezüge hergestellt, ohne dass diesen nachgegangen werden konnte. Diese Lücke schliesst das folgende Kapitel 9.