

Sisyphus wird nicht zur Ruhe kommen: ethische Herausforderungen der Künstlichen Intelligenz



Volker Thiel

(Free University of Berlin, vtuberatung@gmail.com)

ORCID: 0000-0002-6434-5505

1. Einleitung

In den vergangenen Jahren fand ein rasend schneller wissenschaftlicher Fortschritt zur Entwicklung der Künstlichen Intelligenz (KI) mit schwacher und starker Ausprägung statt. Unter KI mit schwacher Ausprägung werden alle programmierbare Maschinen, wie z.B. Industrieroboter, die vorgegebene Arbeiten ausführen, zusammengefasst. KI Maschinen mit starker Ausprägung sind die Maschinen, die durch neuronale Netze (Deep Learning) in die Lage versetzt werden, „eigenständige“ Entscheidungen zu treffen, wie es beispielsweise bei Kriegsdrohnen bereits der Fall ist. Diese Entwicklung stellt nicht nur den Einzelnen, sondern ganze Gesellschaften und letztlich die gesamte Menschheit vor neue ethische Orientierungsproblematiken: woran können wir uns orientieren (persönlich, als Gesellschaft, als Menschheit) um unser Verhalten untereinander und in Bezug zur KI auszurichten? Welche ethischen Herausforderungen ergeben sich bei der Entwicklung von KI?

Ethische Orientierungen, als ein Ergebnis philosophischer Überlegungen, sind eingebettet in eine Vielzahl von externen Bestimmungsfaktoren wie z.B. kulturelle Traditionen, Entwicklung des Wissenschaftsbetriebes, Sprache und Zeitereignisse. Diese externen Faktoren beeinflussen das Zustandekommen von ethischen Überlegungen, die räumlich und zeitlich zu unterschiedlichen Schwerpunktsetzungen führten und führen. Deshalb konnte sich eine universelle Ethik, die von allen Menschen gleichermaßen akzeptiert wird, bis heute nicht herausbilden. Trotzdem gibt es aber in den räumlich und zeitlich sich unterschiedlich entwickelnden Ethiken normative Gemeinsamkeiten, wie z.B. das Tötungsverbot eines anderen Menschen. Diese Gemeinsamkeiten und der seit Jahrhunderten andauernde abendländische philosophische Diskurs über die unveräußerlichen Rechte eines jeden einzelnen Menschen führte letztlich zur allgemeinen

Erklärung der Menschenrechte. Aber allein die Forderung, die Würde des Menschen (d.h. jedes einzelnen Menschen!) ist unantastbar (Grundgesetz für die Bundesrepublik Deutschland, 1967, Artikel 1,1), scheint nicht Konsens aller Menschen zu sein. Und wenn bereits diese Basisforderung zu Interpretationsvarianzen führt, um wieviel größer ist die Herausforderung, allgemeinverbindliche ethische Orientierungen im Umgang mit KI zu formulieren und durchzusetzen. Um den Diskursrahmen einzugrenzen werden sich die weiteren Ausführungen kurz auf richtungsweisende abendländische Traditionen beschränken. Die Beziehung von Ethik und KI ist dahingehend zu differenzieren: geht es um die grundsätzliche Frage, ob KI ethisch sein kann, oder eher um die technologische Herausforderung (Entwicklung von entsprechender Software), wie KI ethisches Handeln implementiert, oder gar selbstreferentiell entwickeln kann (Sloane 2019, 9).

2. Von den Anfängen regelgeleiteten Verhaltens und moralischer Normen

Die Geschichte der Menschheit zeigt, dass Menschen Gattungswesen sind, d.h. Menschen haben immer schon in Familienverbänden, Gruppen, Sippen und später auch größeren Verbänden wie Stämmen oder Gesellschaften zusammengelebt. Um dieses Miteinander zu organisieren waren und sind eindeutige Regeln notwendig. Dabei spielte es zunächst keine Rolle, wie diese Regeln entstanden sind und durchgesetzt wurden. Sei es durch bloße Gewalt – Unterordnung unter dem Mächtigsten (Faustrecht) – oder, sehr viel später in der neueren Entwicklung von Gesellschaften, durch konsensuale Übereinkommen und Verträge.

Die ältesten erhaltenen schriftlichen Regelwerke sind der Kodex des Hammurabi – auf den hier nicht weiter eingegangen wird - und die Zehn Gebote. Die Zehn Gebote sind nach biblischer Aussage den Israeliten durch die göttliche Offenbarung des Moses verkündet worden (Die Bibel, Exodus 20,1–20,21 und Deuteronomium 5,1–5,22) sind also nicht durch Übereinkunft (Vertrag) zustande gekommen. Sie definieren grundlegende Forderungen, an die sich jeder Israelit (Jude) zu halten hat und tragen zugleich dazu bei, das Zusammenleben der Israeliten als Volk regelorientiert zu ermöglichen. Wie begrenzt ihre gesellschaftliche Wirkung ist, zeigt sich in den aktuellen politischen Auseinandersetzungen in Israel gerade in Hinblick auf das Zusammenleben mit Menschen anderer Orientierungen.

Die jüdische Sekte der Christen übernahm (u.a.) die Zehn Gebote und machte sie zu einer ihrer Glaubensgrundlagen, die bis in die Gegenwart – in allen christlichen Religionen - Bestand haben. Aber auch hier gilt, die Zehn Gebote haben nur eine verbindliche Gültigkeit für alle Christen, nicht aber für alle Menschen. Und woran orientieren sich andere, nicht christliche Menschen? Viele orientieren sich an ihren persönlichen, anerzogenen (familiären) moralischen Maßstäben, andere an Forderungen, wie sie durch weise Menschen (Philosophen) oder durch (göttliche) Offenbarungen (Propheten)

verkündet wurden. Wieder andere an den Gewohnheiten und Gesetzen (Sitten) ihrer Länder. Diesen Orientierungen ist gemein, dass sie keinen selbstgesetzten Regeln folgen, wohl aber die fremdgesetzten akzeptieren, die so unterschiedlich sein können, wie es die Kulturen sind. Dabei kann es passieren, dass diese moralischen Orientierungen im Wettbewerb, teilweise sogar im Widerspruch, zueinanderstehen können. Eine Moral kann z.B. fordern, dass alle Menschen, die nicht dieser Moral folgen, zu missionieren sind. Gleichzeitig können demgegenüber die Gesetze eines Staates entgegenstehen, weil sie fordern, dass unterschiedliche Orientierungen gleichberechtigt nebeneinander friedlich koexistieren sollen. Auch wenn es keine universale Moral gibt, bleibt die Frage offen, ob es dennoch so etwas wie verallgemeinerbare Verhaltensnormen und Regeln geben kann, die Allgemeingültigkeit beanspruchen, akzeptiert und durchgesetzt werden können. Um eine Antwort zu finden, werden im Folgenden exemplarisch Gedanken einiger abendländischer Philosophen erwähnt, die sich mit ethischen Fragen beschäftigten und deren Erkenntnisse für die weitere Diskussion wegweisend wurden. Dass es sich dabei um eine verkürzte subjektive Auswahl und Interpretation handelt, sei vorausgeschickt, denn hier ist nicht der Ort für eine ausführliche, gebührende Würdigung dieser großen Denker.

3. Ethik als Disziplin abendländischer Philosophie

Die Nikomachische Ethik (NE) des Aristoteles (1969) ist das erste in sich geschlossene Werk eines abendländischen Philosophen, das sich systematisch mit Fragen der Ethik, genauer mit der Frage nach einem geglückten Leben (Eudaimonia), auseinandersetzt. Für Aristoteles war es selbstverständlich, dass alles Handeln und Leben einem Ziel zustrebt und der vernünftige Mensch sein Handeln an den Tugenden orientiert. Deshalb beschreibt und bewertet er die unterschiedlichen Tugenden auf ihre Tauglichkeit hin, ein geglücktes Leben durch Handeln/Tätigsein zu erreichen. Jede dieser Tugenden hat ihre Berechtigung, ist aber allein für ihn zur Erreichung des Zieles nicht ausreichend. Deshalb kommt er am Ende seines Werkes für sich zu dem Schluss, dass für ihn das höchste Gut das ist, was man um seiner selbst willen erkennt.¹ Bekannt wurde die NE im abendländischen Kulturkreis erst im Mittelalter über die Rezeption islamischer Gelehrter (Schweizer 2017, Position 4325), die sie bereits seit Jahrhunderten kannten und diskutierten. Die Universitätsgründungen mit der Vormachtstellung der Scholastischen Theologie bzw. Philosophie verhalf ihr zu einem endgültigen Durchbruch und bildet die Grundlage für philosophische Diskurse bis in die Gegenwart. Im weiteren Verlauf der Philosophiegeschichte, ausgehendes Mittelalter, Beginn der Renaissance, stand nicht mehr die Frage nach einer persönlichen Vervollkommnung (Eudaimonia)

¹ „(...) was dem einzelnen wesenseigen ist, das stellt für den einzelnen von Natur das Höchste und das Lustvollste dar. Für den Menschen ist dies also das Leben des Geistes, nachdem dieser vor allem das wahre Selbst des Menschen darstellt, und dieses Leben ist dann also auch das glücklichste“ (Aristoteles 1969, 290).

im Fokus des Interesses, sondern die Frage, wie ein gemeinschaftliches (friedliches) Zusammenleben auf ethischen Grundsätzen möglich sei. Das Paradigma der Tugendlehre wurde in Britannien durch den Utilitarismus und in Preußen durch Kants Philosophie (des Kategorischen Imperativs) abgelöst. Beide Philosophien beanspruchen für ihre Ethik eine universelle Geltung, die die einzelnen Menschen in Bezug auf ihre Nächsten (und alle anderen Menschen) zu berücksichtigen haben (Pflichtenethik).

Diese philosophischen Erkenntnisse und Forderungen fanden auch Eingang in politische Manifeste wie z.B. die Amerikanische Verfassung (Bill of Rights), Forderungen in der Französischen Revolution bis hin zur allgemeinen Erklärung der Menschenrechte der Vereinten Nationen. Die Einsicht, dass es ethische begründbare universelle, unveräußerbare Menschenrechte gibt und der Wille, diese universal durchsetzen zu wollen, scheinen gegeben zu sein. Nur zwischen der „Erkenntnis“ und der Realität klaffte und klafft eine erhebliche Lücke. Nicht nur einzelne Personen, sondern ganze Staaten verweigern den universalen Geltungsbereich von der Würde des einzelnen Menschen und erklären ihre Interpretation für die richtige, auch wenn sie offensichtlich im Widerspruch zur Allgemeinen Erklärung der Menschenrechte steht.² Aus diesem Verhalten ergeben sich unterschiedliche Fragestellungen: Sind die Menschenrechte tatsächlich je nach kultureller Verbundenheit oder politischer Lage interpretierbar oder haben sie fortwährende Allgemeingültigkeit? Und wenn sie eine fortwährende Allgemeingültigkeit haben, was dürfen und können andere Staaten unternehmen, um die Staaten, die sich offensichtlich gegen z.B. die Allgemeinen Menschenrechte verhalten, wieder einzuhegen? Welche Maßnahmen und Mittel sind legitim?

Nehmen wir als Beispiel zur Illustration den Atomwaffen Sperrvertrag. Alle den Vertrag unterzeichneten Staaten erklären, dass sie den Bau, Einsatz und die Verbreitung von Atomwaffen ablehnen. Es gibt Staaten, die nicht dem Abkommen beigetreten sind und sich trotzdem daranhalten und es gibt Staaten, die gerade weil sie nicht beigetreten sind für sich beanspruchen, dass dieser Vertrag für sie keine Gültigkeit besitzt (z.B. Nordkorea). Welche Reaktionsmöglichkeiten haben die Vertragstreuen Staaten auf die Abtrünnigen und Gegner einzuwirken? Der Appell an *pacta sunt servanda* reicht schon lange nicht mehr aus. Aber auch Sanktionsandrohungen und -umsetzungen haben nur eine begrenzte Wirkung. Festzustellen ist, nicht alle Staaten sind bereit ihre nationalen Interessen internationalen Abkommen zu unterstellen. Sollte deswegen ganz auf solche Verträge verzichtet werden (mehr dazu weiter unten)? Aktuell betreffen ethische Grundeinstellungen und ihre Geltung nicht mehr nur und ausschließlich das menschliche Zusammenleben, sondern zunehmend auch das Interagieren von Menschen mit KI Maschinen (Misselhorn 2018, 8).

² Dazu zählen u.a. Staaten wie der Iran, China, die Vereinigten Arabischen Emirate, Russland, die Philippinen, Kuba, Venezuela, Brasilien. Die Liste lässt sich fortsetzen.

4. Künstliche Intelligenz (KI) als ein Ergebnis menschlicher Intelligenz

Ein entscheidender Schritt auf dem Weg zur Entwicklung der KI Maschinen war die Konstruktion von Industrierobotern, die Arbeiter von monotonen, immer wiederkehrenden Arbeitsabläufen befreiten. Die Industrieroboter ersetzen und verdrängen zwar Menschen von ihren angestammten Arbeitsplätzen, was aber häufig durch einen anderen, neuen Arbeitsplatz kompensiert werden konnte. Viele der unmittelbar Betroffenen teilten noch mit den Wissenschaftlern, Tüftlern und Unternehmen die Vision einer besseren, menschenwürdigeren Arbeitswelt.

Diese Entwicklung, die erst vor ca. 80 Jahren einsetzte, scheint sich überschlagen zu haben: die Verarbeitung und Auswertung unvorstellbar großer Datenmengen in Bruchteilen von Sekunden ist ebenso Realität wie die Weiterentwicklung von KI als selbstlernende Systeme (Deep Learning). Auch wenn diese durch Algorithmen gesteuerten Maschinen — nach Meinung von Experten — erst maximal die Intelligenz von Embryonen haben (Malsburg 2019, 1), also diese Entwicklungsstufe sich gerade erst am Anfang befindet, ist das Ziel vorhersehbar: eine selbstlernende und selbst Entscheidungen treffende Maschine auf Grundlage von KI zu entwickeln. Erste Erfolge auf diesem Weg sehen neurowissenschaftliche Forscher in der Entwicklung künstlich neuronaler Netze (Deep Learning), die Abwägungsentscheidungen selbständig treffen können und bereits Eingang in die Entwicklung autonomer Waffensysteme (Kriegsroboter) gefunden haben. In dem Moment, wo diese Maschinen untereinander und erst Recht mit Menschen interagieren, stellen sich Fragen wie, wer ist für die Entscheidungsparameter der Maschinen letztlich verantwortlich oder verfügen diese KI-Maschinen über eine genuine, eigene Verantwortung über ihre Entscheidungen? Welche Verantwortung tragen dabei die Entwickler und Programmierer und welche Rolle soll der Gesetzgeber dabei spielen (mehr dazu weiter unten)?

Alle Maschinen, vom Industrieroboter bis hin zur selbstlernenden KI, sind von Menschen entwickelt und programmiert worden, d.h. sie sind von diesen (Menschen) abhängig und zu verantworten. Keine Maschine besitzt ein Gehirn und verfügt über Intelligenz, so dass keine auch noch so intelligenten Algorithmen kombinierbare KI Maschine diese Defizite ausgleichen kann (Boos 2019). Trotzdem bleibt die Frage zu beantworten, gibt es in der Entwicklung von KI Maschinen so etwas wie „Rote Linien“, die nicht überschritten werden dürfen? Und wenn Ja, wer legt sie fest, überwacht sie und was geschieht, wenn sie dennoch überschritten werden?

5. KI und ethische Herausforderungen

Wie bereits in der Einleitung erwähnt, ist die Frage nach Ethik und KI sowohl eine grundsätzliche als auch eine nach dem technologisch Machbaren (also eher eine Frage nach der Softwareprogrammierung). Gesetzt den Fall, Maschinen mit KI werden irgendwann

in der Lage sein, komplexe Zusammenhänge selbstreferentiell zu erfassen, auszuwerten und eigenständige Entscheidungen zu treffen, sind sie dann Maschinen-Menschen oder anders gefragt, was unterscheidet den Menschen noch von einer KI Maschine, was ist das spezifisch Menschliche?

Von der Antike bis zum ausgehenden Mittelalter bestimmte das Verhältnis von Körper und Geist (und später auch von Körper, Geist und Seele) die philosophische Diskussion, wenn es darum ging, die Personalität des Menschen zu verstehen. Die Antwortrichtung auf die Jahrhunderte lange (abendländische) Beschäftigung mit der Frage, was den Menschen als Menschen (in seiner Personalität) auszeichnet, würde heutzutage eher darauf verwiesen, dass der Mensch in der Lage ist, durch Selbstbewusstsein, Willensfreiheit und seine Fähigkeit zur Selbstreflexion zu handeln (Misselhorn 2018, 205). Dabei bedeutet Selbstbewusstsein, sich seiner selbst bewusst zu sein, ich weiß um meine Einmaligkeit als Person, ich bin ein Individuum. Willensfreiheit ist Voraussetzung, Verantwortung für eigene Handlungen (oder Unterlassungen) zu übernehmen, und die Selbstreflexion ist die Fähigkeit, sich reflexiv mit seiner eigenen Persönlichkeit, seinen Ideen und Handlungen kritisch auseinander zu setzen. Dass bei der Persönlichkeitsentwicklung Herkunft, soziale Verhältnisse, Erziehung, Sozialisation und Lebenserfahrungen eine nicht unerhebliche Rolle spielen, ändert nichts an der persönlichen Verantwortung jedes einzelnen Menschen für sich und seine Handlungen. Diese Vermögen besitzen alle Menschen, aber können auch Maschinen mit KI diese Vermögen besitzen?

Selbst lernende Maschinen (Deep Learning) können Eindrücke nicht nur registrieren, sondern in Beziehung zueinander setzen, um dann Entscheidungen zu treffen (s. Kriegsdrohnen). Aber auch diese Entscheidungsfindungen basieren auf vorprogrammierten Algorithmen, sind also kein Ergebnis einer eigenständigen, selbstverantworteten, freien Willensentscheidung. Und ebenso verfügen KI Maschinen (noch nicht) über das Vermögen der Selbstreflexion, eine grundlegende Voraussetzung zur Entwicklung persönlicher Intelligenz. Daraus folgt, dass Maschinen mit KI keine ethischen Handlungen verantworten können, denn Verantwortung zu übernehmen setzt Personalität voraus.

6. KI und ethische Grenzen

Die Entwicklung von Maschinen mit KI ist grundsätzlich dort eine Bereicherung, wo sie z.B. monotone Arbeitsabläufe ersetzen, permanente Überwachungen und Kontrollen durchführen können oder als Assistenten Menschen unterstützen. Zu Problemen kann es immer im konkreten Einzelfall kommen. So können z.B. Pflegeroboter, ergänzend zur menschlichen Pflege, bestimmte Dienste übernehmen, die sowohl den Pflegebedürftigen helfen, als auch den Pflegenden entlasten. Problematisch wird es in dem Moment, wo es um den Ersatz von zwischenmenschlichen (personalen) Beziehungen geht. Ist es mit der Würde des zu pflegenden Menschen verantwortbar, wenn Maschinen menschliche

Zuwendung (Empathie) vorspielen? Aufgrund ihrer Konstruktion sind sie nicht in der Lage Emotionen (wie z.B. Empathie) selber zu empfinden, wohl aber auf emotionale Bedürfnisse reagieren zu können. Sind solche Vortäuschungen ethisch zu rechtfertigen?

Vergleichbares gilt für andere Einsatzmöglichkeiten von Maschinen mit KI, wie z.B. bei Telefonhotlines (mehr dazu weiter unten). Die grundsätzliche zentrale Frage aber ist, darf eine Maschine mit KI stellvertretend für den Menschen Entscheidungen treffen? Um sich einer Antwort zu nähern, sei als Beispiel ein nicht abwendbarer Unfall mit einem selbstfahrenden Auto angenommen. Die Steuerung erfasst gleichzeitig, dass ein Kind auf die Straße läuft und dass sich auf dem Gehweg Fußgänger befinden. Eine sofortige Entscheidung muss getroffen werden. Eine Vollbremsung würde die Mitfahrenden Personen verletzen und ggf. zu Auffahrunfällen des nachfolgenden Verkehrs führen und scheidet deshalb aus. Wie soll die autonome Steuerung entscheiden? Nach dem Prinzip des geringsten Schadens? Das würde bedeuten, ein Menschenleben gegen ein anderes vorzuziehen. Ist das gerecht bzw. rechtens?³ Auch das Argument, dass ggf. durch autonomes Fahren die Zahl der Unfalltoten reduziert werden könnte, löst das ethische Problem nicht, sondern versucht durch Quantifizierung eine neue Dimension einzuführen. Das offensichtliche Dilemma ist, dass eine konstruierte Maschine, selbst wenn sie selbstreferentiell dazu lernen würde, ihre Entscheidungen ausschließlich nach logarithmischen Wahrscheinlichkeiten trifft und keine persönliche Verantwortung übernehmen kann. Die einzige Lösung besteht darin, dass der Fahrer (Mensch) die finale Entscheidung treffen muss, auch wenn sie unweigerlich einen Kollateralschaden zur Folge haben wird.

Wenn für (ethische) Entscheidungen vor allem vorprogrammierte Nützlichkeitsargumentationen (Utilitarismus) ausschlaggebend sind, dann verabschiedet sich der Mensch von seiner individuellen Verantwortlichkeit und gibt damit einen Teil seiner Persönlichkeit auf.

7. Exkurs zur Entwicklung des autonomen Fahrens

Es scheint förmlich ein Wettbewerb unter Internetanbietern, Automobilherstellern und selbsternannten Experten zur Entwicklung des autonomen Fahrens ausgebrochen zu sein. Jeder Teilnehmer möchte seine Lösung als erster nicht nur auf der Straße, sondern als Prototyp für eine Massenherstellung der Öffentlichkeit präsentieren, so als wenn die potentiellen Käufer bereits Schlange stehen würden. Doch wie groß die Nachfrage tatsächlich ist, kann gegenwärtig niemand verlässlich sagen. Sicherlich gibt es Technikbegeisterte, die sich das selbstfahrende Auto wünschen. Aber zu welchem Preis? Und hier ist nicht nur der Preis für ein Auto gemeint, sondern der Preis, der sich aus gesellschaftlichen Forschungs- und Entwicklungskosten, die Ausstattung der

³ Vgl. Grundgesetz (a.a.O., Artikel 1 und 2).

Infrastruktur mit Datenübertragungseinrichtungen, ökologische Folgekosten und nicht zuletzt mit der Datensicherheit (Informationelle Selbstbestimmung⁴) ergibt. Ungeklärt ist auch, wie lassen sich selbstfahrende Autos in den bestehenden Verkehr integrieren und führt deren Einführung zu einer messbaren Verhaltensänderung anderer Verkehrsteilnehmer? Solange es auf diese offenen Fragen keine validen Antworten gibt, ist die Entwicklung, autonomes Fahren zu ermöglichen, eine technische Herausforderung, ohne dass der praktische Nutzen nachgewiesen werden kann. Vielversprechender wäre es, die Praktikabilität bereits vorhandener Fahrassistenten zu erhöhen und weitere, neue Assistentenentwicklungen auf ihre Sinnhaftigkeit (Aufwand, Ertrag, Kosten, Infrastrukturbelastungen, Datensicherheit, ökologische Auswirkungen etc.) zu überprüfen und in bestehende Fahrzeugtypen zu integrieren. Nicht die Substitutionen traditioneller Automobile durch selbstfahrende Autos lösen irgendwelche Probleme in den Innenstädten, Ballungsgebieten und auf den Fernstraßen. Sie führen zu einer neuen entscheidenden Fragestellung, die nach dem Verhältnis von KI Anwendungen und Ethik.

8. Konsequenzen für den Einsatz von Maschinen mit KI

Der Einsatz von Maschinen mit KI muss differenziert nach deren Anwendungsbereichen betrachtet und bewertet werden. Maschinen als Ersatz für den Menschen, z.B. in der Produktion von Gütern, sind heute Standard und weitestgehend akzeptiert (sofern weitere Automatisierungen nicht Massenarbeitslosigkeit zur Folge haben). Problematischer wird es, wenn KI Maschinen als Assistenten oder gar Vertreter für den Menschen eingesetzt werden. Assistentenleistungen in Form von Begleitung, Überwachung oder Kontrolle, die den Menschen (Arzt, Pfleger, Pilot, Autofahrer etc.) unterstützen, sind hilfreich und werden bereits eingesetzt. Bei ihrem Einsatz ist aber darauf zu achten, dass der verantwortliche Mensch auch tatsächlich jederzeit in den Prozess eingreifen kann, um die Situation (positiv) zu retten. Tragische Vorfälle belegen aber, dass z.B. durch Gewöhnung, der Autopilot wird es schon richten, Blitzentscheidungen und Handlungen tendenziell verlernt werden⁵. Noch problematischer ist es, wenn KI Maschinen den Menschen ersetzen, z.B. in der (Alten-) Pflege, im telefonischen Servicebereich, durch den Einsatz von Drohnen oder beim autonomen Fahren. Zum autonomen Fahren wurde bereits ausführlich Stellung bezogen (s.o.) und festgestellt, der Mensch (Fahrer) muss nach seinem eigenen Selbstverständnis verantwortlich die finale Entscheidung treffen. Das gleiche muss umso mehr bei der Entwicklung und dem Einsatz von autonomen Waffensystemen (Drohnen) gelten. Für ihr Ziel, Anlagen oder gar Menschen zu vernichten, sind sie bereits jetzt in der Lage, neue Informationen für ihren Einsatz auszuwerten und ggf. eigenständig Flugbahnänderungen vorzunehmen. Das

4 Vgl. Bundesbeauftragter für den Datenschutz und die Informationsfreiheit (Hrsg.): DSGVO – BDSG. Texte und Erläuterungen (= BfDI-Info. Nr. 1, Bonn 2019).

5 Vgl. Fry (2019, Position 2367 ff).

autonome Töten findet bereits statt. Die Frage aber, wer letztlich die Verantwortung dafür trägt, ist nicht beantwortet. Nur weil etwas technisch möglich ist, ist es nicht automatisch ethisch vertretbar. Um diese zentrale Fragestellung geht es: muss nicht bei jedem Einsatz von Maschinen mit KI immer die letzte Entscheidung durch einen Menschen (oder eine Gruppe von Experten) zu verantworten sein?

Auch ist es ethisch problematisch, Menschlichkeit durch „Maschinenlichkeit“, z.B. in der Pflege, ersetzen zu wollen. Hier liegt zumindest eine geduldete Täuschung (vielleicht sogar Betrug?) vor (s.o.). Vergleichbar, wenn auch weniger brisant, verhält es sich bei telefonischen Servicediensten. Der menschliche Anrufer kommuniziert unbewusst mit einer Maschine. Würde er auch bewusst eine Maschine anrufen, anstelle mit einem Menschen zu sprechen? Eine besondere Bewertung verlangt der Einsatz von KI bei schienengebundenen Fahrzeugen wie Bahn, S-Bahn oder U-Bahn. Diesen Transportmitteln ist gemein, dass sie nicht ausweichen können, da sie schienengebunden sind. Zur Verhinderung eines möglichen Crashes sind sie auf schnelles Abbremsen angewiesen. Hier ist KI dem menschlichen Reaktionsvermögen bei weitem überlegen. Und langjährige Modellversuche belegen, dass der Fahrer überflüssig ist. Also ein eindeutig positiver Einsatzort für KI Maschinen. Wenn wir zu dem Schluss gelangen, dass ein Mensch (eine Gruppe) grundsätzlich die letzte Entscheidung zu treffen hat, dann ist es notwendig, sich darüber national und multinational zu einigen, um internationale Standards zu definieren und auf deren Einhaltung zu achten.

9. Möglichkeiten und Grenzen zur Durchsetzung von nationalen und multinationalen Regelungen

Jeder Staat kann per Gesetz festlegen, welche Forschungen er nicht zulässt (z.B. Forschung an und mit Embryonen oder das Anreichern von Uran zu Herstellung von Atombomben). Zudem kann er für Verstöße gegen seine Gesetze Strafen androhen, die aber, wie ein Blick in die Geschichte zeigt, trotzdem Vergehen nicht grundsätzlich verhindern werden. So war in vielen europäischen Ländern über Jahrhunderte das Sezieren von Leichen bei strengen Strafen (weltlicher- und kirchlicherseits) verboten. Trotzdem konnten diese Verbote und Strafandrohungen den Forscherdrang nicht gänzlich verhindern. Denn es gab immer wieder Einzelne, die diese Verbote nicht akzeptierten und sich bewusst möglicher staatlicher Repressalien aussetzten oder konsequent gleich in ein Land auswanderten, das diese Verbote nicht verfolgte und sanktionierte. Auch wenn solche Verbote nur eine eingeschränkte Regulierungsmacht haben, sind sie unverzichtbar, definieren sie doch sittliche Standards eines Staates. Niemand käme auf die Idee, das Tötungsverbot deshalb aus der Strafprozessordnung ersatzlos zu streichen, nur weil es in Einzelfällen nicht beachtet wird. Der Staat muss nachvollziehbare, klare Grenzen setzen und deren Einhaltung garantieren.

Noch viel komplizierter wird es, wenn das Ziel ist, Bi- bzw. Multinationale Abkommen

zu schließen. Gerade die Verträge zur nicht Herstellung und Verbreitung von Atomwaffen (Atomwaffen Sperrvertrag⁶) zeigen, wie schwierig es ist, internationale Abkommen zu vereinbaren und deren Durchsetzung zu gewährleisten. Wie können Staaten, die solchen Vereinbarungen erst gar nicht beitreten wollen, zu einem Beitritt bewegt werden und wie können Staaten, die sich an die Vereinbarungen nicht weiter halten, wieder zur Vertragstreue zurückgewonnen werden? Alle bekannten Maßnahmen zeigen ihre Wirkungsgrenzen auf. Das betrifft Strafmaßnahmen ebenso wie Gewaltandrohungen oder gar Gewaltanwendungen (Krieg). Sie tragen langfristig nicht zur Lösung eines Problems bei (s. dazu aktuell die Atomanreicherungen in Nordkorea oder im Iran), bestenfalls zu einer Verzögerung aktueller negativer Auswirkungen. Es bleibt am Ende nur das Bemühen, internationale Verträge zustande zu bringen und deren Einhaltung zu fordern. Eine wahrhaftige Sisyphusarbeit.

Literaturangaben

- Aristoteles. 1996. *Nikomachische Ethik*. Übersetzung und Nachwort von F. Dirlmeier. Anmerkungen von E. A. Schmidt. Stuttgart: Reclam.
- Boos C. 2019. *Was künstliche Intelligenz NICHT kann?* Beitrag vom 24. März 2019. Link: <https://www.youtube.com/watch?v=er1J4QMW1c4>.
- Bundesbeauftragter für den Datenschutz und die Informationsfreiheit (Hrsg.) 2019. DSGVO – BDSG. Texte und Erläuterungen (= BfDI-Info. Nr. 1), Bonn. <http://www.bfdi.bund.de/SharedDocs/Publikationen/Infobroschueren/INFO01.pdf>. (abgerufen am 12. März 2019).
2005. *Die Bibel. Die Heilige Schrift des Alten und Neuen Bundes*. Freiburg im Breisgau: Herder.
- Fry H. 2019. *Hallo World, Was Algorithmen können und wie sie unser Leben verändern*, aus dem Englischen von S. Schmid. München: C.H. Beck.
- Merker R. 1967. *Grundgesetz für die Bundesrepublik Deutschland*. Stuttgart: Reclam.
- Malsburg C. v. d. 2019. „Gesucht: Vorbilder für kluge Automaten,“ *Frankfurter Allgemeine Zeitung FAZ* vom 13. März 2019.
- Misselhorn C. 2018. *Grundfragen der Maschinenethik*. Stuttgart: Reclam.
- Schmid Noerr G. 2006. *Geschichte der Ethik*. Stuttgart: Reclam.
- Schwartzmann R. 2018. „Das Recht der Maschinen,“ *Frankfurter Allgemeine Zeitung FAZ* vom 25. Oktober 2018.
- Schweizer G. 2017. *Iran verstehen: Geschichte, Gesellschaft, Religion*. Stuttgart: Klett-Cotta.

⁶ Vgl. Der Atomwaffensperrvertrag ist seit März 1970 in Kraft. Ihm gehören 191 Staaten an.

Sloane M. 2019. „Intelligente Maschinen hinter ethischen Rauchwänden,“ *Frankfurter Allgemeine Zeitung FAZ* vom 24. Juni 2019.

Volker Thiel (Berlin)

Sisyphus Cannot Rest. Ethical Challenges of Artificial Intelligence

Abstract: The use of machines with artificial intelligence in different areas of life and the trend to develop a „homunculus“ makes ethical rating necessary. Starting with early rules of behavior and moral norms, traditions in the occidental ethic history will be presented and faced with the question, whether they can be of any help to understand artificial intelligence, and which conclusions should be considered while developing and using machines with (strong) artificial intelligence by individuals, states, and finally, the whole world.

Keywords: artificial intelligence; machine ethics; automotive driving; emotional elusions; personal responsibility.

Ethics in Progress (ISSN 2084-9257). Vol. 10 (2019). No. 2, Art. 4, pp. 33-44.

Creative Commons BY-SA 4.0

DOI:10.14746/eip.2019.2.4