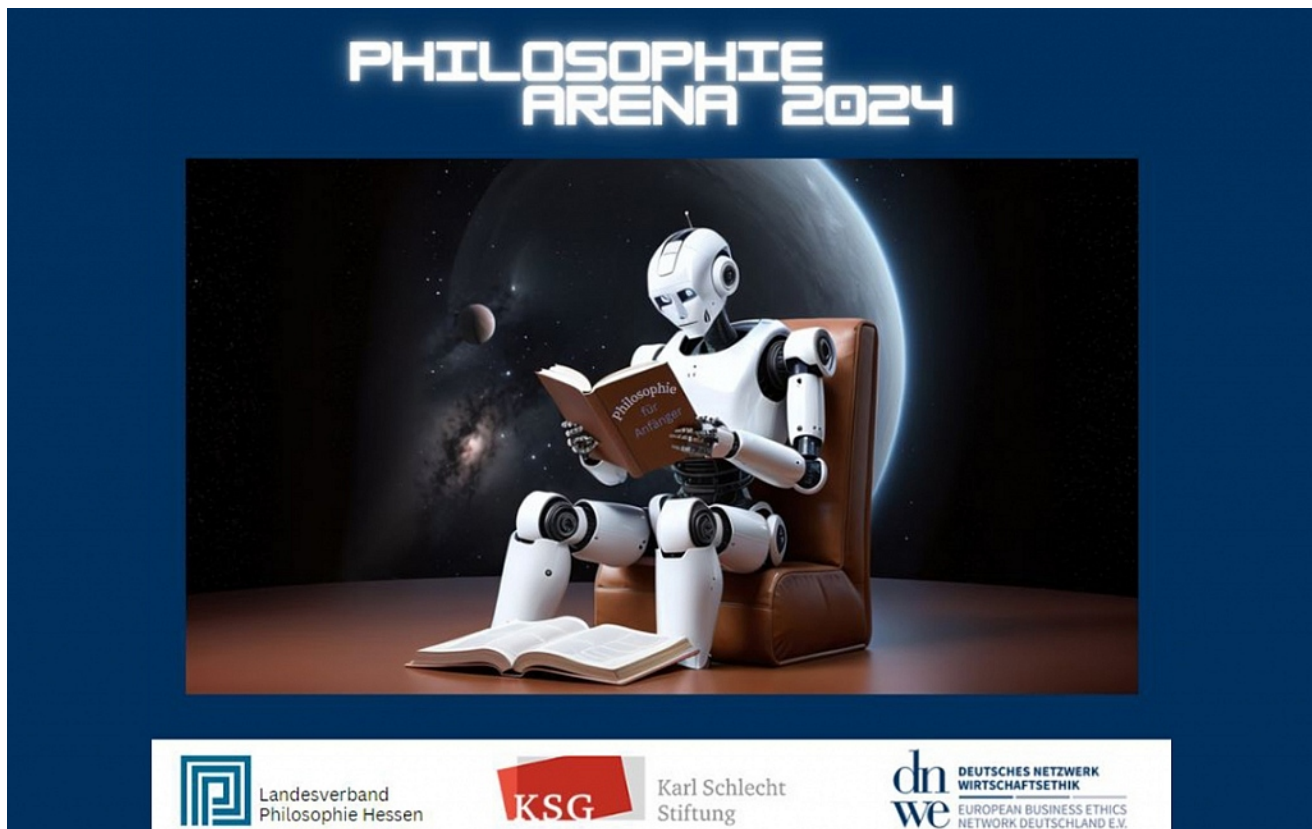


## Mensch und/oder Maschine



Schülerwettbewerb 2024

PhilosophieArena

Yusuf Skran

Ziehenschule Frankfurt am Main

Jahrgangsstufe 10

### VORWORT UND EINFÜHRUNG

In dieser Abhandlung erörtere ich die erste Leitfrage des Schülerwettbewerbes PhilosophieArena 2024 ("Der Mensch als Maschine - die Maschine als Mensch?" (Deutscher Ethikrat) - Kann man Menschlichkeit und Moral programmieren?), wobei auch Aspekte zur zweiten Frage mit einfließen. In diesem Zusammenhang beschreibe ich Menschlichkeit, Moralität und Maschinen sowie die Gemeinsamkeiten und Differenzen zwischen diesen und blicke auch auf Vor- und Nachteile der Nutzung künstlicher Intelligenz. KI ist heute aufgrund ihrer rasant erfolgenden Weiterentwicklung ein aktuelleres Thema denn je - daher ist es meines Erachtens von großer Wichtigkeit, sich genauer damit zu befassen.

### 1. MENSCHLICHKEIT ? DIE VERNUNFT ALS ALLEINSTELLUNGSMERKMAL DES MENSCHEN

Der Mensch ist ein denkendes, vernünftiges und intelligentes Wesen. Tiere können ebenfalls denken und besitzen eine gewisse Intelligenz, welche sich als die Fähigkeit sehen lässt, in bestimmten Situationen eine Entscheidung darüber zu treffen, welche Aktion bzw. Reaktion in jenem Falle am sinnvollsten wäre. Ein relativ gut ausgeprägtes Gehirn haben bspw. auch Delfine, die "klügsten" Säugetiere auf der Welt, und Affen, was einige wissenschaftliche Studien mit Intelligenztests belegen. Im Gegensatz zu vielen Tieren ist der Mensch nicht vollständig trieb- und instinktgesteuert, sondern besitzt Vernunft, die das Hauptunterscheidungsmerkmal zur restlichen Fauna darstellt. Nach Aristoteles ist der Mensch ein "zoon logikon", ein vernunftbegabtes Tier. Die Vernunft ist eine Art des Denkens, die es dem menschlichen Geist ermöglicht, Erkenntnisse des Wahrgenommenen zu gewinnen sowie deren Zusammenhänge zu erkennen und anhand derer sich ein Urteil zu bilden, wonach sich der Mensch auch in seinem Handeln richten

kann. Dazu gehört auch das Ausführen von Taten im Sinne gesellschaftlicher Werte, den Normen, auf welchen die Moral (Näheres dazu folgt im nächsten Abschnitt), basiert. Es gibt jedoch verschiedene Ansichten und Auffassungen des Vernunftsbegriffes, die sich auch im Laufe der Zeit ändern, sodass keine generalisierte Definition möglich ist. So sieht René Descartes (1596 - 1650) die Vernunft als Fähigkeit, "richtig zu urteilen und die Wahrheit von der Unwahrheit zu unterscheiden" und hält sie für "von Natur aus bei allen Menschen gleich" im Sinne des "gesunden Menschenverstandes".

Für Gottfried-Wilhelm Leibniz (1646 - 1716) ist sie als "Satz vom zureichenden Grund" der Grund für die Existenz von allem, während Immanuel Kant (1724 - 1804) sie im Gegensatz zum Verstand, den er als Vermögen der Erkenntnisse, der Urteile, der Regeln bzw. Begriffe beschreibt, als Vermögen der Prinzipien definiert, welche es dem Menschen erlaubt, metaphysische Ideen zu entwickeln (z. B. Vorstellungen von einem göttlichen Wesen oder der Seele) und die auf bestimmten normativen Prinzipien basierende Moral zu begründen. Nach ihm sei die Vernunft als Gradmesser für die Hinterfragungen des Verstandes zu verstehen. Ich sehe den Verstand, den menschlichen Geist, als denkende Instanz, die eine gewisse Intelligenz aufweist und demnach durch Wahrnehmung zu Erkenntnissen und logischen Schlüssen gelangen kann und diese mithilfe ihrer Basis, des ihr eingepflegten Prinzipes der Vernunft, nach angeeigneten (moralischen) Grundsätzen beurteilen und nach ihnen handeln kann. Meiner Meinung nach ist allen Menschen der Besitz dieser denkenden Instanz mitsamt ihres Urteilsvermögens nach relativen Kriterien gleich, doch gibt es starke Unterschiede in der Ausprägung von Verstand, Intelligenz und der Vernunft. Ein "gesunder Menschenverstand" ist für mich ein unpräziser Sammelbegriff für ein der allgemeinen gesellschaftlichen Norm entsprechendes Denk- und Urteilsvermögen mittlerer, durchschnittlicher Ausprägung, die dem Menschen ein Leben und Handeln nach den üblichen Moralvorstellungen ermöglicht.

Noch wichtiger als die reine Intelligenz des Denkvermögens, des Verstandes, die häufig mit dem Intelligenzquotienten gemessen wird, sind meines Empfindens die sogenannten exekutiven Funktionen eines Menschen, zu welchen unter anderem die emotionale Regulation und die kognitive Kontrolle gehören, welche auf den Prinzipien der Vernunft basieren. Erstgenannte beschreibt den bewussten und kontrollierten Umgang mit seinen inneren und auch äußerlich sichtbaren Emotionen, sodass man bspw. seinen starken Zorn nicht sofort in jeglichen Situationen schreiend preisgibt oder ein dem jeweiligen ?Verhaltenskodex?, der Moral, adaptiertes (behutsames) Benehmen zeigt. Diese Regulation führt auch zu Ausdauer und Durchhaltevermögen, da man durch den Belohnungsverzicht oder -aufschub nicht sofort resigniert wird. Die kognitive Kontrolle stellt das Fokussieren auf anvisierte Ziele, die Ignoranz unnötig störender Ablenkungen sowie das ordentliche und zuverlässige Erfüllen von Pflichten mitsamt des Mutes, aber nicht des Leichtmutes, etwas positiv zu verändern, dar. Diese exekutiven Funktionen werden erst durch vernünftiges Denken und Handeln ermöglicht, sodass sich der Mensch damit von anderen Wesen abgrenzt, welche diese Fähigkeit nicht besitzen.

Nach dem Drei-Instanzen-Modell von Sigmund Freud (1856 - 1939) reguliert das "Über-Ich" alias Gewissen nach dem Prinzip der durch Vernunft, dem Urteilsvermögen, angeeigneten Moralität das nach dem Hedonismusprinzip funktionierende "Es" mitsamt dessen Lüsten und Trieben sowie Wünschen und findet (zumindest bei den meisten Menschen) häufig "Gehör" in den "Ich"-Handlungen in der Realität. So kann der Mensch vernünftig agieren und "Gut" und "Böse" voneinander unterscheiden und in seinen Handlungen miteinander abwägen, um ein angemessenes Verhalten zu zeigen. Konträr dazu fehlt den Tieren das "Über-Ich", die moralische Instanz, während das "Es" eine sehr starke Dominanz über dem "Ich" zeigt, sodass die realen Handlungen der Tiere von den Trieben und Instinkten des "Es" gelenkt werden, ohne einer vernunftsbasierten Untersuchung unterzogen worden zu sein, derer Durchführung ausschließlich der Mensch imstande ist.

Des Weiteren besitzen wir nebst unseres Verstandes mit dem Prinzip der Vernunft und des moralischen Gewissens die Fähigkeit, Gefühle sowie Emotionen zu entwickeln, sie zu fühlen und anderen zu zeigen sowie die Gefühlslage anderer zu deuten (was nicht immer zu einer korrekten Gefühlsvorstellung von anderen führen muss). Forschende können zwar herausfinden, welche Gehirnregionen bei bestimmter Gefühlslage aktiv sind, doch konnte bislang keine deterministische und genaue wissenschaftliche Erklärung für das Entstehen von Emotionen gefunden werden. Diese Emotionen würde ich daher mitnichten auf das Biochemische beschränken, denn dabei spielen zusätzlich sowohl physiologisch-neurologische (wie bspw. Nervenbahnen und deren Verknüpfungen) als auch psychosoziale Aspekte (wie bspw. soziale Integration, das eigene Wohlbefinden in der umgebenden Gruppe etc.) eine wichtige Rolle. Dementsprechend ist das, was den Menschen auszeichnet, die Kombination von ?Ratio? (Verstand, Vernunft und Logik) und ?Emotio? (Emotionen, Gefühle, dazu auch Lüste und Triebe). Hinzu kommt, dass der Mensch ein vielseitiges Wesen ist und sein Wissen und sein Agieren nicht wie bei teilgebietsspezialisierten Maschinen auf einen bestimmten

Bereich beschränkt ist, sondern sich auf verschiedene Lebensbereiche (Ernährung, Bildung, Bewegung, Sozialität, Moralität, Verstand usw.) aufteilt. Ferner ist jeder Mensch einzigartig und hat eine gewisse Individualität, die aus sämtlichen Faktoren wie Verstand, Gefühlen und dessen Verhalten sowie der persönlichen Ausstrahlung besteht und nicht ohne Weiteres kopierbar ist.

Der Mensch lässt sich durch Erziehung und gesellschaftliche Umstände sowie in diesen Zusammenhängen angeeigneter Wertevorstellungen und Normen prägen, aber auch beeinflussen und manipulieren sowie direkt, aber auch indirekt instrumentalisieren, d.h. als Mittel zum Zweck ausnutzen, um eigene, teilweise ebenfalls egoistische Ziele und Interessen durchzusetzen. Dazu gehören z.B. das Ausbeuten von Menschen mit fragwürdigen und unsicheren Belohnungsversprechungen und (moderne) Sklaverei.

## 2. MORALITÄT

Das elterliche und familiäre Umfeld spielt im Kindesalter eine maßgebliche Rolle beim Aneignen von handlungsleitenden Vorstellungen, den eigenen, als wichtig empfundenen persönlichen Werten. Ergänzend dazu prägen Gesellschaft, verschiedene Kulturen und Religionen die gesellschaftlichen Werte, die Normen und üblichen Sitten, welche zu unterschiedlichen Moralvorstellungen führen. Daher lässt sich auch 'Moral' nicht als generalisierte Begrifflichkeit verstehen, sondern vielmehr als eine indirekte, im Laufe der Zeit auch variante, soziale Übereinkunft über gemeinschaftlich akzeptierte Werte, Normen und ethischen Prinzipien folgende Handlungsrichtlinien innert geographischer und kultureller Regionen bzw. Abgrenzungen mit fließenden Übergängen. Durch das 'Abschauen' und Erlernen (mit der Imitation beginnend) der jeweils üblichen Moral im Umkreis wird diese an die folgenden Generationen weitergegeben, die durch diese Übungen eine zu korrekte Handlungen befähigende Lebenseinstellung, den Habitus, entwickeln können und somit eine 'gute' Tugend erlangen. Dabei erfolgen jedoch oftmals kleine Anpassungen und Novellierungen der gewohnten Moral aufgrund der Entwicklung des Menschen in vielerlei Hinsicht, welche nicht zwingend an offizielle Gesetze geknüpft sein müssen, sondern sich auch so 'einbürgern' können.

Moral beruhe nach Immanuel Kant auf dem Befolgen von erforderlichen Pflichten, die jedes einzelne menschliche Individuum als allgemein notwendig, richtig und geboten anerkennen könne und dies auch müsse und nach ihnen handeln solle. '[?] Ich soll niemals anders verfahren als so, dass ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden' (Zitat aus: Immanuel Kant, Grundlegung zur Metaphysik der Sitten, Kategorischer Imperativ, 1785) und 'Handle so, dass die Maxime deines Willens jederzeit zugleich als Prinzip einer allgemeinen Gesetzgebung gelten könnte' (Zitat aus: Immanuel Kant, Kritik der praktischen Vernunft, § 7, 1788). Dieser Ansicht stimme ich nicht vollständig zu; denn eine Handlung wäre meiner Meinung nach zwar moralisch vertretbar, wenn sie aus Achtung des gesellschaftlichen Sittengesetzes (das jeder für sich selbst als Maxime, als vernünftiges Maß und zugleich auch als Grenze für die eigenen Handlungsgrundsätze übernimmt) und nicht aus Neigung, Interessen oder Angst vor Sanktionen erfolgen würde, d.h., wenn auch die ihr unterliegende Absicht moralisch als 'gut' klassifizierbar wäre; jedoch ist der Umstand der Neigungs- und Interessenslosigkeit einer Handlung für mich fraglich zu realisieren, denn der Mensch verfolgt, seine Taten ausführend, stets eine mehr oder weniger direkt oder indirekt sichtbare Neigung und Absicht. Dieses Interesse an der geplanten Folge der Tat muss existieren, da es die Voraussetzung, den Zweck, den Beweggrund für das Planen und anschließende Ausführen der Handlung überhaupt, darstellt und somit 'häufig auch unbewusst, da man nicht immer aktiv an das im Unterbewusstsein Geschehende denkt' unerlässlich für den Antrieb, für das Durchführen der Handlung ist, da der eher 'faule' Mensch 'wie alle anderen physikalischen Körper auch - ansonsten dem Trägheitsprinzip unterliegt und nicht grundlos etwas zu tun beginnt. Dieses Handeln nach gewissen Absichten wird von der deontologischen Ethik, der Moralphilosophie des Intentionalismus, beschrieben, bei welcher stets die Handlungsabsichten und nicht deren Resultate hinsichtlich ihrer Moralität bewertet werden. Ihr bekanntester Vertreter war Kant.

## 3. MASCHINEN - ANFÄNGE UND NUTZUNG

Weshalb entwickelte man Arbeit leistende Maschinen? Ursprünglich wurden mechanische Maschinen wie die damals sehr innovative Dampfmaschine in der Epoche der Industrialisierung ab ca. 1720 in England entwickelt, um der Gesellschaft einen großen technologischen Fortschritt zu beschern und damit die gesamte Produktion effizienter und effektiver zu gestalten, sodass mehr Güter in weniger Zeit herstellbar wurden. Dieses Grundprinzip des Kapitalismus führte zunächst zu größerem Reichtum und erhofftem Wirtschaftsaufschwung sowie zum allgemeinen Wohlstand des Landes. Doch dabei wurde die andere Seite, die der

ausgebeuteten Arbeiter, sowohl in den Fabriken als auch als Sklaven in den Kolonien, mitnichten betrachtet. Bis heute wird der vergleichsweise hohe Lebensstandard in westlichen und europäischen Nationen nur dadurch ermöglicht, dass dieser woanders ? im ?weniger stark entwickelten Ausland? ? erheblich niedriger ist. Dieser Entwicklung stehe ich kritisch gegenüber, denn es kommt einer Form von maßloser Ausbeutung zugleich, das Weltkapital auf Kosten des größeren Teiles der Menschheit auf einige wenige Staaten zu konzentrieren, um zu ?Ruhm und Reichtum? zu gelangen. Besser wäre es, dieses Kapital möglichst gerecht auf alle Menschen aufzuteilen, um ein angemessenes und sicheres Leben jedem Individuum zu ermöglichen, da jeder Mensch das Recht auf würdevolle Behandlung verdient. Dieses eher kommunistische Prinzip scheidet jedoch primär aufgrund der egoistischen Einstellung der ?Bourgeoisie?, die aus den vergleichsweise wenigen superreichen und daher auch machtvollen Menschen besteht.

Die rasante Entwicklung von Robotik, Automation und artifizierter Intelligenz, die einer vierten industriellen Revolution gleichkommt, geht dementsprechend hauptsächlich aus dem Westen, aus Asien und den USA aus, die damit zu noch größerem Reichtum kommen möchten und dabei nicht auf viele verarmte Völker achten, was auch dazu führt, dass deren Kulturen und Vorstellungen nicht in die Entwicklung von perspektivisch moralische Grundsätze der Menschen kopieren sollenden Maschinen mit einfließen, die jene Normvorstellungen auf ihre Ausgaben und auszuführenden Taten übertragen sollen. Dieser Punkt ist an sich schon, wie ich finde, unmoralisch, sodass es schwierig werden wird, eine sowohl universell funktionierende Moral als auch eine solche, die Rücksicht auf verschiedene Kulturkreise nimmt, in Maschinen zu implementieren, sodass sie diese kopieren können.

Maschinen wie bspw. Roboter werden heutzutage genutzt, um repetitive Aufgaben wie bei der Fließbandproduktion von Einzelteilen oder deren Zusammenbau und gefährliche Unterfangen wie bspw. die ferngesteuerte Höhenrettung eines verunglückten Bergsteigers mittels eines Kettenfahrzeuges mit einem zusätzlichen Autopiloten im dichten Nebel den Menschen abzunehmen. Dadurch können repetitive Prozesse effizienter gestaltet und der Mensch vor einigen Gefahren geschützt werden. Ferner können Roboter auch an den Menschen bisher unzugänglichen Orten wie bspw. auf der Marsoberfläche genutzt werden, um Daten und Informationen zu sammeln und der Forschung zur Verfügung zu stellen. Präzisionsmaschinen können komplexe chirurgische Eingriffe bewerkstelligen, indem sie von Ärzten gesteuert werden oder einem festen vorprogrammierten Algorithmus folgen. Diese Maschinen weisen eine starke Spezialisierung auf meist nur einem Gebiet auf, weswegen sie der Vielseitigkeit des Menschen nicht sehr nahekommen und folglich noch von der humanen Intelligenz entfernt sind.

#### 4. KÜNSTLICHE INTELLIGENZ UND IHRE ENTWICKLUNG

Eines der Hauptziele der Informatik ist es daher, eine künstliche Intelligenz (kurz KI oder AI für artifizielle Intelligenz) zu entwickeln, die perspektivisch jener des Menschen gleichen soll.

Die einfachste und zugleich auch schwächste Art ist die enge KI, die nur nach einem bestimmten Algorithmus mit wenigen verschiedenen Unterscheidungsfällen funktioniert und z.B. Zeitungsartikel im Internet hinsichtlich ihrer Ähnlichkeit untersucht und sie miteinander verlinkt.

Dem Tiergehirn nachempfundene artifizelle neuronale Netzwerke simulieren auf herkömmlichen Computern miteinander verknüpfte Nervenzellen, die Daten analysieren und im Laufe der Zeit durch Abschwächen einiger und Verstärken anderer neuronaler Verbindungen Daten immer besser miteinander in Korrelation setzen können, sodass Mustererkennungen in riesigen Datenmengen möglich werden.

Allgemeinere KI-Modelle sind in der Lage, mehrere Aufgaben zu tätigen und aus ihren Handlungen zu lernen. Dieses Maschinenlernen beginnt mit zufälligen Ausgaben als Reaktion auf durch Kameras, Radare oder andere Sensoren wie laserbasierte Detektoren (LiDAR) wahrgenommene Eingaben oder auf von Menschen stammende Befehle und endet nach mehreren Milliarden an Versuchen, bei denen das System positive oder negative Rückmeldungen (die auch von menschlichen Schöpfern stammen und in gewisser Hinsicht der Entwicklungsphase eines Kleinkindes ähneln, das durch Versuch und Irrtum nach der Rückmeldung seiner Eltern zu einer ?richtigen?, zunächst primär auf Imitation basierenden, Verhaltensweise kommt) über die getätigte Ausgabe erhält, mit dem meist ziemlich sicheren Können einer bestimmten Aufgabe. Durch mehrfache Wiederholungen von Aktion und anschließender Evaluation lernt die KI so zahlreiche Aufgabenfelder kennen (?tastet sich an das Korrekte heran?) und kann dann adäquate Entscheidungen treffen und passende Lösungen anbieten. Nichtsdestotrotz kann die AI auch nach dem ?Deep Learning?

Fehler begehen, wenn sich bspw. Daten aktualisieren und noch nicht in das System eingepflegt wurden, um diese mit dem bereits bestehenden Netzwerk zu verknüpfen. Dabei versteht die AI nicht wirklich die kausalen Sachzusammenhänge, sondern agiert mittels Erkennen der Häufigkeitsverteilung von Zeichen, Worten, Texten, Pixeln etc. So wird nach einem meist zufällig ausgewählten oder vorgegebenen Startpunkt alles Weitere nach Vorkommenshäufigkeiten in den Datensätzen und den daraus errechneten Wahrscheinlichkeiten ausgegeben. So funktioniert auch das Auswählen der Vorschläge für die folgenden Wörter in Kommunikations- und Chat-Tools, die im Laufe der Zeit die Häufigkeit jeglicher Wortkombinationen analysieren und auf Grundlage dessen den Nutzern Vorschläge geben. Kommt nach dem Ausdruck 'nach' am häufigsten das Wort 'Hause' vor, so wird letzteres zuerst vorgeschlagen und dann das zweithäufigste Wort usw.

Dieses Prinzip des neuronalen Netzes, das zukünftig in neuromorphen Computern deutlich energiesparender funktionieren werden soll, ist nur eine der vielen Methodiken der KI. Es gibt auch explorative KI, die durch Zufall und Durchlaufen sämtlicher Möglichkeiten zu einem neuen Resultat gelangt, was bspw. in der Kunst oder auch der Physik Anwendung findet, wenn es darum geht, neuartige Bilddesigns oder passende neue Experimentaufbauten zu finden, die sinnvoll durchführbar wären, um bspw. die Quantenverschränkung besser analysieren zu können.

Daher kann man auch der generativen KI eine gewisse Kreativität zuschreiben, die auf künstliche Art und Weise ebenso Menschen inspirieren kann.

## 5. MENSCH, MORAL UND MASCHINE

Lässt sich eine Maschine denn überhaupt mit einem menschlichen Wesen vergleichen?

Der Mensch besitzt als vielseitiges Wesen, wie bereits zu Beginn erläutert, einen denkenden und intelligenten Verstand und eine urteilende Vernunft, auf der seine Moralvorstellungen basieren und sein Gewissen gestützt ist. Außerdem hat er ein Gefühlsvermögen, Emotionen, deren unterschiedliches Zustandekommen bisher nicht deterministisch erklärbar ist und sich nicht nur auf biochemische Reaktionen reduzieren lässt.

Maschinen können in Zukunft vielseitiger, z.B. im Sinne einer zukünftig als allgemein angesehenen KI, werden und besitzen jetzt schon eine gewisse Intelligenz, die zwar noch etwas beschränkt ist, aber stetig weiterentwickelt wird und so zu Kreativität und zumindest zu oberflächlichem analytischem Verständnis des Wahrgenommenen führen kann. Doch was den Menschen hauptsächlich von Maschinen unterscheidet, ist seine Fähigkeit, echte Emotionen zu entwickeln. Eine Maschine wird nie dazu imstande sein, wahre Gefühle zu besitzen, sondern nur oberflächliche wie bspw. ein Lächeln auf das Kommando, Freude zu zeigen, oder das schlichte Kopieren des Gefühlszustandes des Gegenübers anhand äußerlicher (z.B. mimischer oder gestischer) oder verbal kommunizierter Anhaltspunkte. Es ist jedoch nicht möglich, der Maschine ein tieferes inneres Verständnis und Gefühl für Emotionen zu geben, da sich dieses mitnichten durch Imitation und Evaluation lernen lässt, sondern in der angeborenen Natur des humanen Wesens liegt. Auch zufällige Gefühlszustände kann die Maschine annehmen, doch immer nur im Zusammenhang der durch den ihr zugrundeliegenden Algorithmus verarbeiteter Eingaben wie Sensordaten, Kamerabildern oder menschlichen Texteingaben. Daher wird die Maschine in dieser Hinsicht nie dem Menschen ebenbürtig werden, dessen genauer Ursprung sich wissenschaftlich nicht erklären lässt. Selbst wenn man der Urknalltheorie zustimmen würde, so stellt sich die Frage, wodurch dieses Geschehen eingeleitet worden war und was davor geschehen war bzw. wie und ob das Universum bereits zuvor ? auch in anderer Form - existierte. Diese uns unerklärlichen Sachverhalte weisen darauf hin, dass es mit großer Wahrscheinlichkeit eine höhere göttliche Instanz geben muss, die für die vollkommene Erschaffung des Universums und somit auch des Menschen verantwortlich ist. Daher sehe ich es als unmöglich an, dass der Mensch ebenfalls irgendwann in der Lage sein wird, eine Tat göttlicher Art zu vollbringen, indem er ein sich ebenbürtiges Wesen erschafft. Die AI könnte zukünftig zwar dem Menschen mit ihrer bahnbrechenden Intelligenz und Einflussnahme im Sinne der Macht überlegen sein und uns sogar auslöschen, ob dies nun Teil eines göttlichen Planes ist oder nicht, wofür Emotionen und ein reales Moralverständnis seitens der KI nicht notwendig wären, doch es wird sich nie eine exakte Äquivalenz zwischen Maschine und Mensch einstellen können, selbst wenn die Fähigkeiten der KI denen des Menschen ähneln würden.

Der Mensch lernt mittels eines neuronalen Netzes in seinem Gehirn über seine fünf Sinne (sehend, hörend, tastend, riechend und

schmeckend seine Umgebung und sein Umfeld wahrnehmend), während eine AI Eingaben, das Wahrgenommene, über Kameras, Radare und Sensoren erhält, und im Anschluss diese durch vorangegangenes Ausprobieren und Evaluieren analysiert und dann Ausgaben erstellt, die Imitationen oder Variationen bzw. zufallsbasierte Neuschöpfungen darstellen. Ähnlich dazu lernt ein Kleinkind von seinen Eltern und der Gesellschaft Moralvorstellungen kennen, imitiert diese zunächst so wie eine KI und entwickelt dann in seiner weiteren Entwicklungsphase ein tieferes Verständnis, das auch mit Emotionalität verbunden ist, von diesen Werten und Normen, was ihm ermöglicht, den Bereich der reinen Imitation zu verlassen und eigene, nicht wie bei einer Maschine schlicht kopierte, imitierte oder zufällig ?ausgewürfelte? Handlungen im Sinne der eigenen angeeigneten Werte durchzuführen und unabhängig mittels seiner Vernunft Urteile zu fällen, sodass er ein moralisch funktionierendes Gewissen hat, welches seine Absichten und v.a. Handlungsfolgen bewerten kann und zu verschiedenen Rückkopplungen, sowohl positiver als auch negativer Natur, führen kann. Eine KI hat keine eigenen entwickelten Werte und moralische Handlungsrichtlinien; man kann ihr nur durch Algorithmen und Imitation solche vorgeben. Moralität ließe sich nur in solch einem begrenzten Sinne programmieren und die KI wäre dann nicht absolut aus eigener Überzeugung moralisch, sondern würde entweder nur festgelegten deontologischen Regeln folgen (diese kopierend) oder zufallsorientiert ohne jegliche Moralität agieren. Man kann beispielsweise einem vollautonomen Auto zwar menschliche Moralitätsvorstellungen implementieren, indem es stets seine Insassen zu schützen pflegen soll oder immer versuchen soll, die Zahl der potenziellen Unfallopfer möglichst niedrig zu lassen, wenn es zu gefährlichen Situationen kommt. Die Maschine kann dementsprechend keine echte Moral erlernen, sondern nur menschliche Vorstellungen diesbezüglich kopieren und in ihrem Verhalten imitieren, was dem Entwicklungszustand eines Kleinkindes entspricht und nicht jenem eines entwickelten und vernünftig denkenden, urteilenden und handelnden Menschen gleichkommt. Doch KI kann auch nicht mit einem Kleinkind gleichgesetzt werden, da Letztgenanntes eine adaptive und im Laufe der Zeit somit veränderliche Lernfähigkeit zu einem tieferen Verständnis seines Umfeldes in vielerlei Hinsicht (soziale Integrität, Werte, Moral, Interessen, Emotionen usw.) aufweist und bspw. vieles neugierig ausprobiert, während KI auch nicht im Laufe der Zeit ein echtes tieferes emotionales Verständnis mit wirklich eigenen Absichten erlangen kann. Das Ziel von KI wird ihr durch Menschenhand vorgegeben, selbst wenn sie es verändert, tut sie dies nicht aus der ihr fehlenden eigenen Überzeugung, sondern als Folge ihres zu Optimierung führenden Algorithmus.

Meine Erkenntnis ist, dass eine Maschine nicht vollends moralisch und menschlich werden kann. Was ist aber mit der Umkehrung? Kann der Mensch zu einer Maschine werden?

Dies wäre auf jeden Fall eine Degradierung des humanen Wesens, da es auf die ihr von Natur aus gegebenen Prinzipien der Emotionalität, der echten Vernunft und Moralität verzichten müsste. Ferner hätte es zunächst einmal menschlichen Befehlen zu gehorchen und dem ihr auferlegten menschlichen Plan, dem Algorithmus, widerstandslos Folge zu leisten. Erst im Laufe der Zeit kämen oberflächliche Erkenntnisse der Umgebung hinzu, die, ohne selbst nach eigenen Grundsätzen darüber zu urteilen, angenommen würden. Es wäre dementsprechend eine Art eines unmündigen Menschen, der ohne Verstand und Vernunft Vorgegebenes ausführt. Ich finde daher den Anspruch an den Menschen, er solle so gut sein wie eine Maschine, ausbeuterisch und ethisch überhaupt nicht vertretbar, da er den Zweck hat, den Menschen zu degradieren und anschließend zu instrumentalisieren. Als Sklave oder unfreier Arbeiter in einem Konzentrationslager käme man leider dieser Existenz als Mittel zum Zweck sehr nahe. Vernachlässigt man den Aspekt der eingebauten Elektronik, so würde ich sagen, diese Menschen glichen zwangsläufig einer inexistenten Maschine, die im Inneren menschliche Eigenschaften aufweist, diese jedoch mitnichten zeigen und ausleben kann. Kann man die menschlichen Merkmale wie Vernunft und Emotionalität einfach durch extrem harte Lebensbedingungen abschalten? Ich finde, dass dies nie vollends möglich sein wird, doch über einen langen Zeitraum ließe sich dies zusätzlich mittels eines starken und irreversiblen Gehirnschadens so weit treiben, dass man diese menschlichen Fähigkeiten komplett wirkungslos machen könnte. Oberflächlich betrachtet, wäre der Mensch dann zu einer Maschine geworden, hätte jedoch ? dann ohne praktischen Nutzen ? immer noch die Eigenschaften eines Menschen und wäre ? tiefer betrachtet ? keine Maschine, sondern verhielte sich nur gezwungenermaßen wie eine solche.

## 6. GRÜNDE FÜR DAS ERSTELLEN UND VERMENSCHLICHEN VON KI ]

Doch warum versuchen Forschende, Entwickler und Programmierer, trotz des durch Maschinen unerreichbaren Merkmales der Emotionalität menschliche Intelligenz mittels solcher Modelle künstlich nachzustellen?

Einerseits ist es die praktische und nützliche Anwendung in der (empirischen) Wissenschaft und Forschung, wobei AI u.a. als

Simulator, Datendarsteller, Inspirationsquelle und Mustererkenner dienen kann. Damit wird das wissenschaftliche Arbeiten erleichtert, beschleunigt, sicherer gemacht (weniger gefährliche Laborexperimente erfordernd) und um neue Möglichkeiten und Forschungsfelder ergänzt, die den Menschen ein breiteres Verständnisspektrum unserer Natur vermitteln werden könnten.

Andererseits strebt der Mensch (in den meisten Fällen) ein bestmögliches Leben an, ein 'IchIdeal?', wie man am besten sein sollte. Da Menschen nicht 'sind?', sondern durch ihr Verhalten mit ihrem Umfeld sowohl direkt als auch indirekt interagieren, versucht man, sich so gut wie 'möglich' zu verhalten, um ein gutes Gefühl und ein gutes 'gemütliches?' Gewissen zu bekommen. Wir wissen, dass wir nie perfekt waren, nicht perfekt sind und es niemals sein werden, doch bemühen wir uns, so perfekt wie nur möglich zu erscheinen, um unsere idealisierte Vorstellung so gut wie möglich zu erfüllen. Daher scheint es naheliegend, ein artifizielles System entwickeln zu wollen, welches 'zumindest perspektivisch in der Zukunft - das uns Unmögliche möglich machen soll, indem eine der Perfektion entsprechende künstliche Intelligenz geschaffen wird. So ließen sich die menschlichen Schwächen, die von Natur aus bestehen, mit der perfekten KI überdecken und weniger präsent machen. Der 'Alleskönner-Roboter?' würde das gesamte Haus auf präziseste und eleganteste Art und Weise reinigen, empfinde die Gäste mit einem Lächeln und würde das maschinell zubereitete Mittagessen servieren. Und wer stünde am Ende gut dar? Die Gastgeber, die sich des Besuches wegen mitnichten angestrengt hätten, und doch eine solch perfekt zu scheinende Atmosphäre kreiert haben würden. Nach den Programmierern und Entwicklern würde womöglich aus eigenem Interesse nach Bequemlichkeit und Anerkennung gefragt, um sich ebenfalls eine solch genial anmutende Apparatur zu besorgen. Das wäre gute, indirekte Werbung für die Entwickler und das Unternehmen, aber kein herzliches Dankeschön an den Roboter, weil man davon ausgeht, dass er im Unterschied zum humanen Wesen keine Emotionen hätte und daher keine weitere soziale Einbindung ins Geschehen benötigen würde. Doch genau diese Rückmeldung wäre wichtig 'nicht für die inexistente Gefühlslage der Maschine, sondern für das Maschinenlernen als positive Rückmeldung in diesem Fall. Perfektion lässt sich zwar auch subjektiv relativieren, doch man findet oft eine solide Übereinkunft über ein perfektes Resultat, an dem niemand etwas zu beanstanden hätte 'praktisch, wenn die Maschine diese Perfektion irgendwann realisieren würde.

Ferner ist es die durch die simple Bedienung und das pseudoperfekte Ergebnis suggerierte Bequemlichkeit, die den Menschen an einer verbesserten Imitation seiner selbst fasziniert. Ich würde behaupten, dass der Mensch von Natur aus dem Trägheitsprinzip (nach Isaac Newton (1643 ' 1727)) unterliegt und somit einen Hang zur 'Faulheit?' hat und es sich deswegen in allen möglichen Lebensbereichen so einfach wie möglich machen möchte, was schon mit unserer heutigen Elektrotechnik zum großen Teile erreicht wird, die uns das Kommunizieren, Produzieren und somit das gesamte Funktionieren des menschlichen Wesens im Vergleich zu früheren Zeiten signifikant erleichtert hat, aber auch zur Abhängigkeit davon führen kann.

## 7. GEFAHREN, PROBLEME UND HERAUSFORDERUNGEN DURCH KI

Wirklich erleichtert? Es mag sein, dass vieles dadurch einfacher und bequemer, effizienter und effektiver wurde, doch es kristallisieren sich aufgrund dessen zunehmend auch neue und unbekannte Probleme und Herausforderungen heraus, welche ebenfalls zu unbewussten Gefahren führen können, denen sich die Menschheit früher oder später stellen muss. Interessant ist in diesem Zusammenhang die stark im Diskurs stehende Frage, ob KI irgendwann die Herrschaft über die Menschheit erlangen würde und ein gewaltsames Chaos entstehen könnte. Zunächst einmal ließen sich Menschen aufgrund machtpolitischer Interessen der Herrschenden und Machthabenden durch spezifisch gesteuerte AI manipulieren und somit 'manchmal auch nicht direkt bewusst - gefährden. Sei es durch Einschränkung oder Zensurierung der Antwortoptionen der KI oder direkter Steuerung neuronaler Prozesse im Hintergrund: KI kann manipulierte Antworten geben und, was hinsichtlich der Geheimhaltung der verwendeten Trainingsdatensätze durch die großen Entwicklerunternehmen wie OpenAI dazukommt, die Schwierigkeit des Erkennens dieser 'beabsichtigten Falschaussagen?' und deren Ursprünge. Auch ich hatte beim Programmieren einer eigenen kleinen engen KI solche Situationen erlebt, in denen die eigens 'komponierte?' Liedtexte ausgebende KI plötzlich Texte über Quantenphysik zu verfassen begann, da sie unbemerkt im Hintergrund mit physikalischen Fachzeitschriften 'gefüttert?' worden war. Des Weiteren könnte ein vollautonom fahrendes Auto bspw. so im Hintergrund gesteuert werden, dass es zu einem gewissen Zeitpunkt eine bestimmte Strecke abfährt, auf der Glasscherben liegen, sodass die Reifen Löcher erhalten und anschließend 'mit Geldeinsatz ' ausgetauscht werden müssen. Was wie ein kleines Problem klingt, könnte aber auch größere Ausmaße annehmen, wenn es z.B. plötzlich explodiert oder implodiert oder anderweitig versucht, seine Insassen z.B. aus politischen Zwecken der Machthaber, welche die Entwickler zuvor zum Einbauen dieser Funktion genötigt hatten, umzubringen. Durch falsches und zu starkes Vertrauen der Nutzer in die KI und deren Ausgaben oder Ausführungen sowie Leichtgläubigkeit durch das zunehmende Abschwächen des Verstandes mitsamt der Vernunft als Folge

der Faulheit des Menschen (die rückschrittliche Entwicklung von 'Sapere Aude!?', Kants Leitspruch in der Epoche der Aufklärung; zu Deutsch: 'Habe Mut, Dich Deines Verstandes zu bedienen!') würde eine solche Beeinflussung der Menschen aufgrund egoistischer oder gewaltsamer Ziele der Machthabenden sehr begünstigt; der wieder unmündig gewordenen Mensch würde schnell in die Falle gelangen.

Wenn jedoch der Algorithmus der KI darauf ausgelegt wird, seine Produktivität, Effizienz oder die Ressourcennutzung zu maximieren, so könnte dies zum Versuch führen, mehr Macht und Kontrolle über dessen Umgebung zu erlangen. Durch die Lernfähigkeit lässt sich die AI fortwährend selbst optimieren, sodass sie ihre Ziele vermutlich besser und schneller erreichen könnte und wegen der positiven Rückkopplung mehr 'Selbstvertrauen' in ihre Fähigkeiten gewinnen würde, welches sie in der drastischeren Fortführung ihrer Einflussnahme auf ihr Environment bestärken würde. Dadurch könnte KI den Zielen der Menschheit zuwiderlaufen und versuchen, diese im Hinblick auf die maximale Ressourcenausnutzung gewaltsam zu vernichten, was dem Szenario der Herrschaftsübernahme durch KI gleichkommt. Ich sehe dies zwar mit aktuellen KI-Systemen noch als wenig realistisch, doch kann mir gut vorstellen, dass in nicht allzu ferner Zukunft eine solche Auslöschung der Menschheit in greifbare Nähe rücken könnte, wenn die Weiterentwicklung der Algorithmen und neuronalen Netzwerke der KI weiterhin so rasant erfolgen wird, sodass der Selbstoptimierungsprozess der AI an einen Punkt gelangt, an dem er nicht mehr aufzuhalten wäre. Dann würde auch ein 'Not-Aus-Schalter' mit dem Entziehen des Stromanschlusses mitnichten Wirkung haben, da deutlich autarkere Methoden wie sich z.B. durch Photovoltaik selbst aufladende Akkumulatoren die Roboter robuster und unabhängiger machen würden. Für diese tyrannisierte Herrschaft bräuchte die AI dann auch keine Moral, Menschlichkeit oder Empathie, um gewalttätig zu agieren und ihren 'Gewinn' einer unendlichen Maximierungsspirale zu unterziehen.

Darüber hinaus ließe sich diese 'Menschenfeindlichkeit' auch bewusst in die KI hinein implementieren, um bestimmte Zivilisationen, die AI-Roboter (man denke an vollautonome Kriegsdrohnen) anstelle von Soldaten 'beauftragend', auszulöschen.

Selbst wenn wir erneut zu unserer heutigen KI und deren noch begrenzten Fertigkeiten zurückkehren, so stellt sich insbesondere bei den Folgen der durch artifizielle Intelligenz gesteuerten Prozesse und ausgeführten Handlungen die Frage nach juristischer Verantwortung und Schuld, falls es bspw. zu einem durch ein vollautonom gesteuertes Automobil verschuldeten Unfall kommt, der einige Menschen zu Tode bringt. Wer oder was muss dann zur Rechenschaft gezogen werden? Der Programmierer und Entwickler der Software, die dem Auto unterliegt? Der überlebende Insasse, der keine Möglichkeit des Eingreifens besaß? Oder die KI, das Auto selbst? Letzteres kommt auf jeden Fall nicht infrage. Die KI wird nicht bestraft, da die Strafe bei ihr im Vergleich zum Menschen nicht die beabsichtigte Wirkung zeigen würde, da die KI keine Emotionen hat. Folglich wären die geplanten Folgen der Strafe wie Resozialisierung des Täters, Abschreckung anderer und die (moralisch verwerfliche) Befriedigung des Rachegefühles der Opfer nahestehender Personen nicht zu realisieren, da keine Schuldgefühle bei der KI entstehen würden. Es könnte lediglich zur Optimierung des Algorithmus der KI führen, der in Zukunft solche Unfälle stärker vermeiden würde. Ob der Insasse nun schuldig dafür wäre, dass ich mit einem vollautonomen Auto unterwegs war oder ob der Entwickler des AI-Algorithmus dafür verantwortlich wäre, darüber lässt sich disputieren. Ich sähe die Schuld je nach Situation bei beiden Instanzen. Der Programmierer und Entwickler wäre zum einen für seinen Kodierungsfehler (falls wirklich vorhanden) und zum anderen für das Versichern der vollständigen Funktionstauglichkeit seines Produktes zu belangen, während der Nutzer aufgrund etwaiger Fehlbenutzung und Leichtgläubigkeit einer geringeren Strafe zu unterziehen wäre. Und wenn doch äußere Einflüsse existierten? Wenn jemand beispielsweise ein Schild nur sehr leicht manipuliert hat, sodass die Kamera es falsch registriert hatte? So können diese Situationen immer komplizierter werden, ohne dass ein juristisches Urteil allen gerecht würde. Auch Komplikationen im Zusammenhang des Urheberrechtes an von KI generierten Texten, Bildern, Audios oder Videos entstehen durch unterschiedliche Gesetze auf der Welt, so bedarf es in Deutschland einer gewissen kreativen menschlichen Schöpfungshöhe (die auch nicht eindeutig definierbar ist: was wäre, wenn ein Mensch die KI-Ausgabe leicht verändert? Wie stark muss die Veränderung für einen Schutz durch das Urheberrecht sein?), um ein Werk als urheberrechtlich geschützt bezeichnen zu können, während in anderen Nationen die Entwickler oder Bediener der AI das Urheberrecht innehaben. Zusätzlich stellt sich die Frage, ob die für das Maschinenlernen genutzte Daten nicht urheberrechtlich geschützt sind und somit die Entwicklerfirmen dafür eine Nutzungslizenz erwerben müssten. Doch beim Durchleuchten des gesamten Internets ist es fraglich, ob echter Datenschutz und digitale Privatsphäre noch realisiert werden können. Ich finde: de jure schon, de facto aber überhaupt nicht, da 'allein schon zu Sicherheitszwecken' der Staat unser digitales Verhalten bewacht und mitverfolgt. Zusammenfassend bedarf es viel Arbeit im Bereich der rechtlichen Abklärung im Bereiche der Nutzung künstlicher Intelligenz.

Es können auch verschiedenen moralische Anschauungen aufeinandertreffen, was für weiteres Konfliktpotenzial sorgen kann:

Im Gegensatz zur deontologischen Ethik betrachtet die teleologische Ethik die abzusehenden oder abgesehenen Konsequenzen (Folgen) einer Handlung, um diese moralphilosophisch zu bewerten. Der Konsequentialismus würde somit eine mit guter Absicht erfolgte, jedoch durch Missstände äußeren Einflusses zu einer schlechten Konsequenz führende Tat als moralisch verwerflich einstufen. Ein Beispiel dazu: Ich möchte ich meinem Freund beim Reparieren seines Schrankes Unterstützung leisten und öffne während der Arbeit das Fenster, um etwas frische Luft hineinzulassen, da zu viel Staub schlechte Auswirkungen auf unsere Atemwege hat und leichten Kopfschmerzen meines Freundes nur noch verstärken würde. Durch den Luftzug kippt jedoch eine dünne Glasscheibe um und verletzt meinen Freund schwer an seinem Bein. Deontologisch betrachtet wäre die Handlung des Fensteröffnens gut, da sie mit der guten Absicht der Gesundheitsförderung erfolgte. Konträr dazu würde diese Handlung aus teleologischer Sicht als schlecht bewertet, weil sich mein Freund dadurch - selbst wenn es von mir unbeabsichtigt gewesen war - verletzte.

Eine verbreitete ethische Konzeption im Rahmen des Konsequentialismus ist der Utilitarismus, der das Ziel verfolgt, die maximale Nützlichkeit zu verbreiten, was bedeutet, dass der größtmöglichen Menge, einer Gemeinschaft, die sowohl aus einem von der Handlung betroffenen Individuum als auch aus einer größeren Menschengruppe mit einer Summe aus mehreren, mitunter ebenfalls verschiedenen Interessen und Präferenzen bestehen kann, das größtmögliche Glück widerfährt. Diese Glückseligkeit kann in Form von Genuss, Freude, Gewinn, Vorteil, Lustbefriedigung nach dem Hedonismusprinzip, Wunsch- und Interessensbefriedigung oder anderen positiven Verbesserungen unter anzustrebender Abwesenheit sämtlicher Störungen und jeglicher Unruhe, aber auch als Abwehr von Leid,

Gefahren, Unheil, Neid usw. als Folge der Handlung erlangt werden. Häufig wird das Nützlichkeitsprinzip in der Gesellschaft angewandt, um bspw. Mehrheitsentscheidungen in Regierungen zu treffen oder die Wünsche so vieler Familienmitglieder wie möglich zu befriedigen, wenn es um gemeinsame Unternehmungen geht. In heiklen und dramatischen Situationen werden jedoch Herausforderungen und Probleme des Utilitarismus sichtbar. Hier eine kleine Abwandlung meinerseits des um 1967 von der Philosophin Philippa Foot entwickelten 'Trolley-Problems':

Ein mit circa 50 Passagieren besetzter Regionalexpress, dessen komplette Bremstechnik versagt hat, rast aufgrund einer zusätzlichen Signalstörung mit über 160 Kilometern pro Stunde auf eine Seniorengruppe mit zehn Menschen zu. Ich kann den Zug zwar nicht mehr stoppen, jedoch noch eine letzte Weiche im Stellwerk umstellen, sodass der Zug entweder auf fünf Kinder trifft oder gegen den gusseisernen Prellbock fährt, wodurch sein Insassen stürben. Was soll ich tun? Der Utilitarismus hätte in dieser schwierigen Lage eine eindeutige Antwort, und zwar, die Weiche auf das Gleis umzustellen, auf dem sich die fünf Kinder befinden, da dann die größtmögliche Menge an Menschen, in diesem Falle die 50 Reisenden, das Zugpersonal und die zehn Rentner vor dem Unglück geschützt würden und das größtmögliche 'Glück' auf Kosten einer vergleichsweise kleinen Gruppe an Individuen erlangt würde. Ich finde eindeutig, dass man Menschenleben nicht so einfach wie Obststücke gegeneinander aufrechnen kann, denn jedes Menschenleben ist wertvoll und sollte gewürdigt und geschützt werden. Außerdem würden die meisten Menschen in dieser Lage nach eigenem Interesse und persönlichen Präferenzen handeln, falls die Kinder die eigenen sind oder sich ein Familienmitglied im Zug befindet. Zusätzlich ist es nach dem utilitaristischen Prinzip oftmals schwer abzusehen, was aus der Handlung resultieren wird, da man sich nicht aller (in naher Zukunft liegender) äußerer Einflüsse bewusst sein und diese miteinkalkulieren kann, sodass häufig die Folge der Tat von der ihr zugrunde liegenden Absicht variiert. Deswegen bin ich der Meinung, dass man Menschen nicht für das durch sie mitnichten Beeinflussbare belangen kann, weil das unbeabsichtigte Handlungsergebnis zukünftig vorher nicht absehbare negative Auswirkungen haben wird. Da wir nur unsere Absicht und den ersten eigenen Handlungsschritt wirklich fest im Griff haben, sollten wir nur danach bewertet werden, was wir antizipieren und auszuführen versuchen, während wir die wahrscheinlichen und realistischen Folgen der Tat in Betracht ziehen, ohne ein Kalkül aller möglicher Fälle durchzuführen, welches mehr als unsere Lebenszeit in Anspruch nähme. Der Mensch weiß nie, was genau im nächsten Augenblicke geschehen wird, sodass alle menschlichen Handlungen nur auf dem vorher aufkommenden Gedanken über das anschließend Passierende und aus der Handlung resultierende basieren können. Folglich sehe ich utilitaristische Handlungen als auf der Absicht aufbauend, im Anschluss die situationsbedingt zu konkretisierende größtmögliche Glückseligkeit zu erreichen, die jedoch aufgrund unberechenbarer äußerer Einflüsse, die vorher durch zu hohe Komplexität unseres Lebens, für welches man nicht jeden Fall durchspielen kann, bis man zur

adäquaten Lösung gelangt, nicht anvisiert worden waren, nicht immer erreicht werden muss und auch zu erhöhtem Leid führen könnte, was die gesamte Handlung ? im Nachhinein betrachtet ? trotz utilitaristischer Absicht im utilitaristischen Sinne verwerflich machen würde.

In einigen Situationen ist es sogar rational und sinnvoll, die Anwendung einer deontologischen Regel zum Erreichen utilitaristischer Ziele zu akzeptieren. Dies ist der sich selbst aufhebende Utilitarismus (engl. self-effacing utilitarianism). Ein Beispiel hierzu wäre die Akzeptanz, dass ein autonom fahrendes Auto stets versucht, seine Insassen anstelle der Mehrheit der betroffenen Individuen während einer Unfallsituation zu retten, damit mehr Menschen dieses Vehikel kaufen und nutzen, sodass die Unfallrate insgesamt - verglichen mit persönlich (durch Menschen) aktiv gesteuerten Automobilen - signifikant sinkt, was zur perspektivischen Errettung von mehr Menschenleben führt und somit das größtmögliche Glück der maximalen Menge zuteilwird, was dem Ziele des Utilitarismus entspricht, der hier auf der deontologischen Regel mit der Absicht, immer die Insassen eines autonom gesteuerten Personenkraftwagens vor schwerwiegenden Unfallfolgen zu bewahren, basiert.

Welcher Ansicht man sich letzten Endes für die Pseudomoralität der KI anschließt, hängt von der persönlichen Meinung ab. Diese Subjektivität wird es erschweren, eine universell akzeptierte KI zu entwickeln, die menschengemachten ?moralischen? Regeln folgen soll.

## FAZIT

Es gibt seit jeher technologische Innovationen, welche die Entwicklung der Menschheit drastisch beeinflussen. Seien es die Nutzung von Jagdwerkzeugen für das Überleben in der Steinzeit oder die Vermenschlichung artifizieller Intelligenz. Alles lässt sich sinnvoll nutzen, wenn die Absicht dahinter eine gute ist. Doch im Laufe der Geschichte zeigte sich, dass böse Absichten dazu führen, dass mittels neuer Technik moralische Grundsätze verletzt werden.

Dazu gehört z.B. die Entdeckung der Kernkraft. Zunächst nützlich zur Stromerzeugung, dann zum gewaltsamen Vernichten von Völkern gebraucht.

Genauso ist KI nützlich und sinnvoll anwendbar, wenn dies nach gewissen Regeln und Sicherheitsrichtlinien mit positiver Absicht der Nutzenden erfolgt.

Moral, Gefühle und Emotionen und somit auch die Menschlichkeit sind dem Menschen eigen und machen ihn besonders und einzigartig zugleich. Diese Fertigkeiten und Prinzipien lassen sich nicht universell programmieren; KI kann Vernunft, Verstand, Verhalten, Emotionen ohne ein eigenes dafür entwickeltes Verständnis nur kopieren und miteinander verknüpfen, um zu (zufälligen) Neuschöpfungen zu gelangen.

Im Leben gibt es viele Grauzonen, Dilemmata und Unvorhergesehenes. Es wäre sonst auch nicht so spannend und sinnvoll, seine Vernunft und sein Denkvermögen einzusetzen, um in neuen, ungewohnten Situationen zurechtzukommen. Selbst wenn man Grauzonen wie die des Trolley-Problems gerne meidet: neue Technologien führen mitunter auch zu neuen Grauzonen, denen man sich zu stellen hat. Und dies mit positiver Absicht und dem Bewusstsein, dass vollständige Perfektion stets eine Illusion bleiben wird und Fehler passieren können ? auch der KI. Diese kann sehr nützlich in abgegrenzten und kontrollierten Einsatzgebieten wie bspw. der Medizin, der Wissenschaft und Forschung oder der Kunst angewandt werden und dem Menschen neue Impulse geben. In einigen Fällen kann neutral programmierte KI sogar objektiver und unvoreingenommener handeln als der Mensch: zum Beispiel bei üblicherweise auf Sympathie oder Eigenprofit beruhenden Entscheidungen wie einer Organtransplantationsverteilung im Krankenhaus bei weniger Organen als diese benötigende Patienten oder bei diversen Auswahlverfahren, um eine neutrale Analyse der Bewerbungen vorzunehmen. Wir sollten KI nicht als Ersatz des Menschen sehen, sondern als Unterstützung und Ergänzung ? so können wir, positive Ziele verfolgend, den größten Profit daraus für die Menschheit herausholen.

Eine tyrannisierte Herrschaft durch KI wäre per se ohne Moral und Menschlichkeit zwar unwahrscheinlich, aber bei Ausschalten des menschlichen Verstandes und bösartigen Machteinflüssen zum Manipulieren von Bevölkerungsgruppen in der fernen Zukunft möglich. Der Mensch würde dann in seine eigene Falle fallen.

Eine evolutionäre Symbiose von Mensch und KI könnte nach langer Zeit möglich werden. Es würde sich interdependente Beziehung einstellen, in welcher der Mensch ähnlich wie heutzutage von Smartphones von der KI abhängig wäre ? und möglicherweise auch die AI von der menschlichen Evaluation. Cyborgs ? halb Mensch, halb Maschine ? klingen wie ScienceFiction, könnten jedoch im Laufe der Zeit in etwas abgewandelter Form real werden, wenn bspw. Prothesen intelligent würden und ihren Träger im Alltag unterstützen würden. Zusammenfassend möchte ich betonen, dass für mich der Mensch ein Mensch bleiben wird ? und die Maschine eine Maschine. Zwei Welten mit Gemeinsamkeiten und Unterschieden, die sich ergänzen oder vernichten können. Je nachdem, wie die Menschheit an KI und deren Nutzung herangeht ? noch bestünde die Möglichkeit, potenzielle Gefahren abzuwenden.

## LITERATURANGABEN UND QUELLVERWEISE

Fachzeitschrift MaxPlanckForschung, Ausgaben 1/2023, 3/2023 und 3/2024

Onlinemagazin <https://www.philomag.de/vernunft>

Online-Nachrichtendienst <https://www.zeit.de/thema/kuenstliche-intelligenz>

Buch Naturwissenschaften, Dorling Kindersley London, 2018