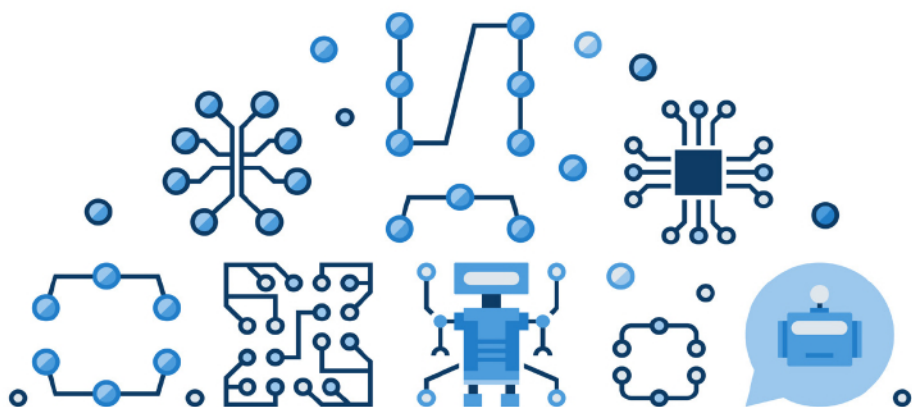
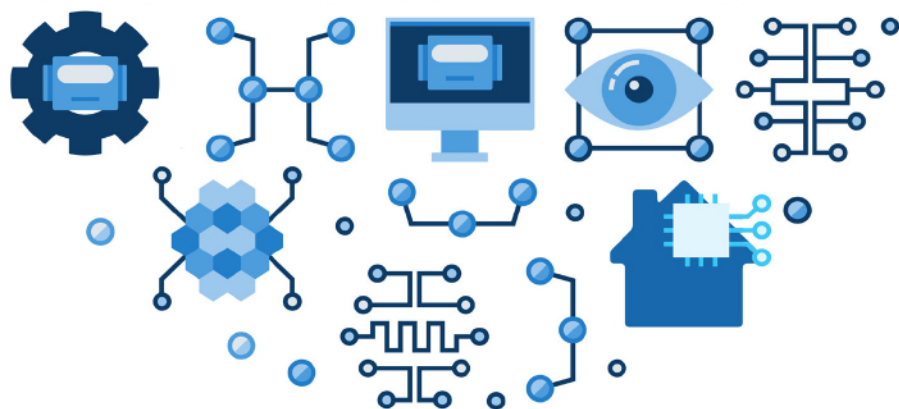


Ralf Otte



MASCHINEN bewusstsein



Die neue Stufe
der KI – wie weit
wollen wir gehen?

campus

MASCHINENBEWUSSTSEIN

Dr. Ralf Otte ist Professor für Industrieautomatisierung und Künstliche Intelligenz an der Technischen Hochschule Ulm. Seit über 25 Jahren arbeitet er an KI-Projekten in Industrie und Gesellschaft. Sein Arbeitsschwerpunkt liegt im Bereich des Maschinellen Sehens und der Erforschung von Bewusstsein auf Maschinen.

Ralf Otte

MASCHINEN BEWUSSTSEIN



Die neue Stufe der KI –
wie weit wollen wir gehen?

CAMPUS VERLAG
FRANKFURT/NEW YORK

ISBN 978-3-593-51470-3 Print
ISBN 978-3-593-44897-8 E-Book (PDF)
ISBN 978-3-593-44896-1 E-Book (EPUB)

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Trotz sorgfältiger inhaltlicher Kontrolle übernehmen wir keine Haftung für die Inhalte externer Links. Für den Inhalt der verlinkten Seiten sind ausschließlich deren Betreiber verantwortlich.

Copyright © 2021. Alle Rechte bei Campus Verlag GmbH, Frankfurt am Main.

Umschlaggestaltung: Guido Klütsch, Köln

Umschlagmotiv: © shutterstock/Venomous Vector

Redaktion: Karin Herber-Schlapp

Satz: Oliver Schmitt, Mainz

Gesetzt aus: Minion und DIN Next

Druck und Bindung: Beltz Grafische Betriebe GmbH, Bad Langensalza
Beltz Grafische Betriebe sind ein klimaneutrales Unternehmen.

Printed in Germany

www.campus.de

Inhalt

Einleitung	7
1. Von Intelligenz und Bewusstsein	15
Was ist Künstliche Intelligenz?	16
Bewusstsein als eine neue Stufe der Intelligenz	24
Was ist Maschinenbewusstsein und was ist es nicht?	30
2. Das Jahr 2021 – die KI der Datenauswertung erobert die Welt	37
Die zwei Säulen der KI – Denken und Lernen	39
KI-Anwendungen, so weit das Auge reicht	54
Die aktuelle KI birgt enorme Risiken	65
Digitalisierung und die unerträgliche Konzentration von Macht	75
Die technischen Grenzen jedweder Software-KI	88
3. Bewusstsein und seine Erzeugung auf Maschinen	93
Wahrnehmung – das Bewusstsein als Beobachter	96
Handlung – das Bewusstsein als Akteur	105
Wissenschaftliche Positionen zur Erzeugung von Bewusstsein	112
Neuromorphe Computer als Hardwarebasis für Maschinenbewusstsein	126

4. Das Jahr 2040 – Maschinenbewusstsein erobert die Welt	135
Maschinenbewusstsein überwindet bisherige Anwendungsgrenzen	136
Vom Nutzen bewusst wahrnehmender Computer	147
Die Gefahren des Maschinenbewusstseins	156
Die Manipulation der Gesellschaft und ihre sozialen Folgen	165
5. Vom Maschinenbewusstsein zur hybriden KI	171
Von der Außenwahrnehmung zur Innenwahrnehmung	173
Warum ein Stein nicht friert – vom Nutzen der Gefühle	175
Maschinenbewusstsein kann weder Gefühle noch Willen ausprägen	183
Die Erzeugung von KI-Chimären und ihre Folgen	191
6. Big Tech lässt den Schleier fallen	199
Der Transhumanismus – die wohl größte Gefahr seit Schaffung der KI	200
Technokratie und die neue dystopische KI-Religion	204
Was kann uns vor den transhumanistischen Gefahren retten?	211
7. Maschinenbewusstsein könnte unsere Sichtweise für immer verändern	221
Danksagung	229
Glossar	232

Einleitung

Der Mensch ist das klügste bekannte Lebewesen im Universum. Menschen haben eine so hohe Intelligenz, dass es noch vor einigen Jahren unmöglich erschien, ihre Intelligenz mit einer sogenannten Künstlichen Intelligenz (KI) technisch nachzubilden. Doch nach vielen großartigen Erfolgen der KI – zum Beispiel dem Sieg von Deep Blue im Schach gegen den damaligen Weltmeister im Jahre 1997, dem Sieg von AlphaGo im Go gegen den damaligen Weltranglistenersten 2016 und dem Sieg von IBMs Watson bei Jeopardy 2011 – kamen bei vielen erstmals Sorgen auf, dass Fachleute intelligente Maschinen bauen könnten, die sogar klüger als die Menschen werden. Man nennt jenen Zeitpunkt »Singularität«, um aufzuzeigen, dass dieser eine Zeitenwende in der Evolution einläuten könnte, die Erzeugung von intelligenten Maschinen, die für ihre eigene Weiterentwicklung keine Menschen mehr bräuchten.

Obwohl sich die Fachleute bis heute nicht auf eine präzise Definition von Intelligenz geeinigt haben, hat man verschiedene Intelligenztests entwickelt, um zumindest die rationale Intelligenz von Menschen miteinander vergleichen zu können. Der sogenannte Intelligenzquotient (IQ) wird dabei auf einen Mittelwert von 100 normiert. Natürlich entscheidet der IQ einer Person nicht über Erfolg oder Misserfolg im Leben, denn mit den IQ-Tests kann man nur rationale Intelligenz messen, Menschen besitzen aber noch weitere Intelligenzformen, wie zum Beispiel soziale oder empathische Intelligenz. Auch der Künstlichen Intelligenz wurden mittlerweile IQ-Werte zugesprochen, die im

Maximum zwischen 50 und 60 liegen. Es scheint also noch ein langer Weg der KI, die Intelligenzleistungen eines Menschen tatsächlich zu erreichen, unabhängig von Siegen in Einzeldisziplinen.

In diesem Buch wird nun die zentrale These vertreten, dass die heutige KI, also die als Software auf einem Computer implementierte Künstliche Intelligenz, niemals an zentrale Intelligenzleistungen von Menschen heranreichen wird. Das schränkt ihre Anwendungen stark ein: Die heutige Software-KI wird niemals vollautonom Auto fahren, niemals allein operieren, niemals die Bedeutung ihrer Datenauswertung erkennen, niemals mit einem Menschen über einen gemeinsamen Kinobesuch sprechen können. Die Sorgen über eine kommende Singularität sind also völlig unberechtigt. Und nicht nur das: KI-Systeme heutiger Bauart werden niemals Gefühle haben und niemals einen Willen ausprägen, von emotionaler Intelligenz fehlt bei ihnen jede Spur.

Weiterentwicklungen der heutigen, softwarebasierten KI zu einem Terminator oder HAL 9000 sind daher völlig unmöglich, sie sind schön oder (wenn Sie mögen) schlechte Science-Fiction. Der Grund für das Unvermögen ist vielgestaltig, letztlich gibt es jedoch einen evidenten Mangel heutiger Künstlicher Intelligenz: Die KI besitzt kein Bewusstsein, doch Bewusstsein ist die Voraussetzung für Wahrnehmung, Gefühle und Willen. Man kann der aktuellen KI zugestehen, denken und lernen zu können – immerhin zwei bedeutende Grundpfeiler von Intelligenz –, aber selbst auch das nur, wenn man Denken und Lernen als mathematische Verfahren versteht, weil auf einem Computer immer nur Algorithmen (letztlich Mathematik und Logik) ablaufen.

Die Künstliche Intelligenz von heute ist beim näheren Hinsehen ein Sammelsurium von mathematischen Verfahren der gehobenen Datenauswertung und der Anwendung von Logikbausteinen. Es geht bei KI, insbesondere dem maschinellen Lernen, um hochkomplexe Auswertungen wie beispielsweise Bildverarbeitungen im Straßenverkehr, die vor Jahren noch nicht vorstellbar waren. Dieser Art von (softwarebasierter) Datenauswertung sind nahezu keine Grenzen mehr gesetzt. Seit 1989 gibt es sogar den ultimativen Beweis, dass man mit bestimmten Verfahren – den neuronalen Netzen – »alles lernen kann, was überhaupt lernbar ist«. Seit dieser Zeit sind die Anwendungen von neuro-

nen Netzen in Industrie und Gesellschaft geradezu explodiert. Aber dennoch handelt es sich immer »nur« um Datenauswertungen. Deswegen kritisieren viele Menschen mittlerweile den Hype um die KI. Die KI sieht für Kritiker sogar wie ein Kartenhaus aus. Tritt man zu dicht heran, könnte man es aus Versehen zum Einsturz bringen.

Doch das würde der KI nicht gerecht werden, denn sie besteht mittlerweile aus so cleveren mathematischen Verfahren mit so überzeugenden Ergebnissen, dass sowohl Laien als auch Fachleute begeistert sind. Man nennt diese KI-Anwendungen oftmals Data Mining, Data Science oder Big Data, und diese Anwendungen werden eine goldene Zukunft haben.

Der Vorbehalt vieler Kritiker gegenüber der KI entsteht daher auch nicht durch die Technologie selbst, sondern durch die falsche Interpretation ihrer Leistungen. Wenn alle Menschen wüssten, dass die KI nur Daten auswertet, dies zwar sehr clever macht, aber dennoch nichts von den Daten und den Ergebnissen *verst*ht, wären sowohl der Hype als auch die Vielzahl von fehlgeleiteten Anwendungen vorbei. Niemals würde ein Manager zum Beispiel auf die Versprechungen hereinfallen, dass bald vollautonome Autos durch Deutschland fahren, und das sogar durch unbekanntes Terrain. Auch würde die Europäische Kommission nicht über Persönlichkeitsrechte von Robotern diskutieren¹, denn alle würden wissen, dass die heutige KI lediglich smarte mathematische Verfahren realisiert, die auf physikalischen Geräten als Software ablaufen. Die Annahme, dass aus mathematischen Verfahren eine Persönlichkeit entstehen könnte, wäre eher lächerlich als diskussionswürdig. Nehmen wir nur das Beispiel Sophia. Sophia ist eine weltberühmte »Roboterfrau«, die bereits Staatsbürgerrechte in Saudi-Arabien genießt. Wer darin aber mehr als einen lustigen Marketing-Gag sieht, wird bewusst oder unbewusst irreführt. Sophia ist eine Maschine, so wie ein lernender Kühlschrank oder ein adaptiver Controller in einem Fahrzeug, nur eben mit einem ansprechenden menschlichen Silikongesicht ausgestattet.

Darum haben viele Menschen auch kein klares Bild von KI: Die heutige KI kann bestimmte Intelligenzeigenschaften so gut simulieren, dass man verkürzt sagt, sie simuliert nicht Intelligenz, sie *ist* intelligent. In

den meisten Anwendungsfällen ist es auch egal, ob die Ergebnisse durch Simulation von Intelligenz entstanden sind oder durch echte Intelligenz. Die Diskussion darüber roch bisher nach akademischer Haarspalterei. Aber nun müssen Fachleute feststellen, dass die mediale Gleichsetzung von Simulation und Original dazu geführt hat, dass viele Menschen – selbst hochintelligente Politiker, Firmenchefs oder Autoren – glauben, KI könne wirklich denken und lernen.

Die großen Technologiekonzerne (ich nenne sie im Buch Big Tech) blasen mit all ihrer Kraft in das gleiche Horn, doch nicht ganz uneigennützig. Für Big Tech wird es nämlich von großem Vorteil sein, wenn Menschen denken, die KI sei intelligent, und die KI bekomme bald sogar einen Willen und vieles mehr. Der Grund dafür liegt auf der Hand. Falls die Menschen das irgendwann wirklich glauben, kann man bei der Einführung neuartiger KI-Technologien sehr unangenehmen Haftungsfragen aus dem Wege gehen. Während heute noch Tesla oder andere Hersteller belangt werden, wenn eines ihrer autonom fahrenden Fahrzeuge mal wieder einen Unfall mit Todesfolge verursacht hat, so wird das nicht mehr so einfach sein, wenn man die KI selbst verantwortlich machen und »bestrafen« kann.

Die Täuschung ist schon jetzt fast perfekt. Wer Siri benutzt, um sein Telefon zu steuern, könnte glauben, Siri *verstünde*, was man sagt. Das ist natürlich nicht so. Es ist alles nur clevere Software, bei der die menschlichen Laute als Muster einer zeitlichen Abfolge von Daten erkannt werden und bei denen (je nach erkanntem Muster) zu bestimmten Unterprogrammen der Software gesprungen wird. Siri und Alexa und all die anderen Chatbots verstehen nichts von dem, was sie sagen, und trotzdem machen sie das, was wir von ihnen wollen.

Uns Menschen ist die Gleichsetzung von Simulation und Wirklichkeit eigentlich nicht egal. Wer von seinem Partner »ich liebe dich« gesagt bekommt, will, dass das stimmt, in der ganzen Bedeutung des Satzes. Wem hier aber tatsächlich die Simulation reicht, der darf zukünftig natürlich auch einen Roboter heiraten.

Vielen Entwicklern und Nutzern der KI ist mittlerweile klar geworden, dass es so nicht weitergehen wird. Die KI-gestützte Datenauswertung kommt langsam an ihr natürliches Ende. Damit ist nicht gemeint,

dass es bald keine KI-Anwendungen mehr geben wird. Nein, bereits in naher Zukunft werden wir von digitalen KI-Anwendungen des Big Data durchdrungen sein, das ist keine Frage. Mit »natürlichem Ende« ist gemeint, dass die Erfolge mit immer größerem Aufwand erzielt werden, was jeder Laie beim autonomen Fahren deutlich erkennen kann. Selbst in 10 Jahren wird es noch keine vollautonomen Serienautos geben, die in unbekanntem Terrain navigieren und unterwegs sein können, selbst wenn man jetzt eine nationale Kraftanstrengung einleiten und Hunderte Milliarden Euro spendieren würde. Ich werde Ihnen zeigen, warum das so ist, und insbesondere, warum das mit der heutigen Software-KI-Technik immer so bleiben wird.

In früheren Büchern haben Kollegen und ich den Begriff Mind Data eingeführt, um zu verdeutlichen, dass die Datenauswertung der Zukunft anders sein wird, als alle heutigen Technologien zu Datenauswertungen vermuten lassen. Warum? Für viele Tätigkeiten wie Reden, Streiten, Witze erzählen, Bilder oder Straßenverkehr wahrnehmen, Dichten und Lieben benötigen wir Menschen nämlich Geist (Mind), und den haben heutige KI-Maschinen nicht. Es wird also immer unmissverständlicher deutlich, dass die zukünftigen Maschinen Geist benötigen, um gewisse Tätigkeiten des Menschen nachahmen zu können.

Im vorliegenden Buch wird es daher um den »Geist in der Maschine« gehen oder, wie wir in der Fachwelt sagen, um Künstliches Bewusstsein (Artificial Consciousness – AC). Der Begriff Geist ist zu abstrakt, um ihn in der Technik zu verwenden. Der Begriff Bewusstsein ist zwar auch komplex, aber jeder Leser weiß, dass er ein Bewusstsein besitzt, daher hat er auch einen intuitiven Zugang zu diesem Begriff. Das ist ein großer Verständnisvorteil.

Die heutige KI hat in keiner einzigen Anwendung weltweit jemals Bewusstsein bewiesen – das ist die These, die ich in diesem Buch erklären und begründen werde. Doch ich werde noch einen gewichtigen Schritt weitergehen. Ich möchte Sie, meine Leserinnen und Leser, in einen neuen Zweig der KI-Forschung einführen und aufzeigen, woran einige Forscher heutzutage arbeiten. Die Rede ist von neuromorphen Computern, also einer KI, die nicht mehr auf Software basiert, sondern vollständig in Hardware realisiert wird, und bei der neuartige Effekte

entstehen können. Der wichtigste Effekt, den ich vorstellen werde, ist das – fast – unvermeidliche Aufkommen von elementaren Bewusstseinsformen in diesen Maschinen.

Es geht also um *künstliches Bewusstsein* auf neuromorphen Computern und ähnlichen Systemen. Doch um welche Art von Bewusstsein wird es sich handeln? Menschliches Bewusstsein kann es nicht sein, denn menschliches Bewusstsein hat nur der Mensch. Doch auch tierisches Bewusstsein kann es nicht sein, denn eine Maschine – auch eine neuromorphe – ist eindeutig unbelebte Materie. Es kann sich bei dieser Art von künstlichem Bewusstsein also nur um die Erzeugung einer völlig neuen Form von Bewusstsein handeln, um Maschinenbewusstsein eben.

In diesem Buch werden wir die Eigenschaften dieses Maschinenbewusstseins kennenlernen und ihre Nützlichkeit für bestimmte Anwendungen. Doch neben den Chancen für viele neue Anwendungen müssen wir die Risiken dieser neuen KI ausführlich diskutieren. Ich werde zeigen, dass die Risiken dieser neuromorphen KI beherrschbar sein werden. Es könnte jedoch zu einer Katastrophe kommen, wenn unverantwortliche Forscher über eine Erzeugung von Maschinenbewusstsein hinausgehen, indem sie die KI mit biologischen Organismen verbinden. Das Schaffen von Transorganismen (im Buch steht der Begriff nicht für genmanipulierte Organismen, sondern für die Verschmelzung von Maschinen mit biologischen Organismen) und die Entwicklung zum Transhumanismus (die Verknüpfung von Maschinen mit menschlichen Zellen, vorzugsweise des Gehirns) könnte sich als das größtmögliche Unheil erweisen, das die KI-Forscher der Gesellschaft jemals antun werden. Und das alles lange vor der hypothetischen Singularität. Wir werden sehen, dass durch den Transhumanismus sehr viele Horrorszenerarien wahr werden könnten. Aber wir können diese Entwicklung stoppen. Niemand zwingt eine Gesellschaft, derartige Wege zu gehen. Wir werden im Buch daher der Frage nachgehen, ob wir wirklich solche Risiken eingehen wollen.

Das Buch ist für alle Menschen geschrieben, die verstehen möchten, was demnächst auf sie zukommen könnte. Und natürlich ist es ein Buch für Entscheider und Politiker. Das Buch zeigt sehr detailliert, was in

Zukunft auf uns alle zukommt, sollte die KI bereits morgen mit einem Bewusstsein erwachen. Natürlich hat niemand eine Glaskugel, ich als Autor ganz gewiss nicht, aber darum geht es auch nicht. Sie, liebe Leserinnen und Leser, sollen viele verschiedene Szenarien kennenlernen, von möglichst vielen Fachleuten hören, um sich daraus ein eigenes Bild der Zukunft zu bauen. Sie sollen nach dem Lesen dieses Buches ein Verständnis davon bekommen haben, was ich unter Maschinenbewusstsein verstehe und was es für Sie ganz konkret bedeuten könnte, wenn Sie demnächst einer KI mit Bewusstsein begegnen, egal aus welchen KI-Laboren diese dann kommt.

Wie ist das Buch aufgebaut? Kapitel 1 definiert gleich zu Beginn wichtige Fachbegriffe, Kapitel 2 berichtet einführend nochmals über Grundlagen und Funktionsweise der heutigen KI, über ihre Risiken und Chancen. Kapitel 3 stellt das wissenschaftliche Fundament zum Maschinenbewusstsein zur Verfügung. Dort werden konkrete Konzepte zur Erzeugung von Bewusstsein erläutert. Viele der vorgestellten Ergebnisse werden in den Fachdisziplinen noch sehr kontrovers diskutiert, denn Maschinenbewusstsein – sowohl der Begriff als auch sein Inhalt – sind Gegenstand der Grundlagenforschung. Welche Verfahren zur Erzeugung von Bewusstsein sich daher durchsetzen werden, ist noch völlig offen. Kapitel 4 wagt einen Blick in die Zukunft in 20 Jahren, es führt Sie in die Vielfalt der Chancen und Risiken von Maschinenbewusstsein ein. Sie sollen verstehen, was in den nächsten Jahren in den KI-Laboren dieser Welt entstehen könnte, und was das für jeden Einzelnen bedeutet. Kapitel 5 und 6 führen über das Maschinenbewusstsein hinaus und diskutieren die riskante Verschmelzung von KI-Maschinen mit biologischen Systemen, im Kapitel 6 werde ich die großen Gefahren des Transhumanismus beschreiben und die Frage erörtern, ob wir so weit gehen dürfen.

Doch das Buch soll mit der Beschreibung von Gefahren nicht beendet werden, mir geht es noch um ein weiteres, sehr wichtiges Thema. Ich werde im Kapitel 7 aufzeigen, dass ein tieferes Verständnis von Geist und Bewusstsein uns selbst betrifft. Denn was bedeutet es eigentlich, wenn wir eine (tote) mineralische Maschine mit Bewusstsein gebaut haben? Können wir einem Regenwurm – der immerhin *lebt* – noch

Bewusstsein absprechen? Oder einer anderen Kreatur? Müssen wir unser Verhalten gegenüber diesen Mitbewohnern unserer Welt dann aber nicht überdenken?

Doch, wir werden es überdenken. Wir werden die Natur mit anderen Augen sehen. Die Regierung von Großbritannien will die Gefühle von Wirbeltieren ab 2021 bereits gesetzlich anerkennen, ein Anfang ist also gemacht.

Künstliches Bewusstsein, Maschinenbewusstsein, ist schon lange kein rein akademisches Thema mehr, denn die Bedeutung der neuen Erkenntnisse über die (physikalische) Natur des Bewusstseins tritt langsam an die Oberfläche, und das wird einen erneuten Wettlauf der Systeme einläuten. Bei diesem Wettlauf könnten die asiatischen Länder im Vorteil sein, weil sie keine Berührungsangst vor den »spirituellen« Ergebnissen haben, die durch die neue Natur des Bewusstseins zwangsläufig entstehen. Es wäre trotzdem schön, wenn Europa wenigstens mitgestalten könnte; in einem Rahmen, den diesmal die Gesellschaft und nicht die Wissenschaft vorgibt. Deshalb sollten alle Menschen eine begründete Meinung dazu haben.

Dafür ist dieses Buch da!



Von Intelligenz und Bewusstsein

Viele KI-Forscher und Neurowissenschaftler verknüpfen die Begriffe Intelligenz und Bewusstsein miteinander. Ihr Argument: Der Mensch besitze eine sehr hohe Intelligenz, und er hat ein sehr komplexes Bewusstsein. Was liegt da näher, als hier einen grundlegenden Zusammenhang zu vermuten? Hohe Intelligenz verursacht hohes Bewusstsein, sogar Selbstbewusstsein, mittlere Intelligenz verursacht mittleres Bewusstsein, und niedere Intelligenz führt dazu, dass das System kein Bewusstsein ausprägen kann. Das klingt plausibel.

Aber muss das so sein?

Es könnte ja auch sein, dass Bewusstsein gar nichts mit Intelligenz zu tun hat und der obige Zusammenhang nur ein scheinbarer ist – etwas, was in der Wissenschaft viel häufiger auftritt, als man sich zugestehen möchte. Zusammenhänge zwischen gefundenen Erscheinungen (Fachleute sprechen oft auch von Korrelationen) sind sehr häufig gar keine kausalen, sondern sie verweisen oftmals auf eine dritte, unbekannte Ursache. Es könnte daher sein, dass der beobachtete Zusammenhang zwischen hoher Intelligenz und hohem Bewusstsein auf grundlegendere Prinzipien der Natur basiert.

Falls das so wäre, könnten vielleicht auch Systeme mit niedrigerer Intelligenz Bewusstsein ausbilden und Systeme mit relativ hoher Intelligenz, wie die KI-Systeme der neuen Generation, vielleicht niemals Bewusstsein erlangen.

Und wie Sie, liebe Leserinnen und Leser, der Einleitung bereits entnommen haben, vertrete ich genau diese These: Heutige KI-Systeme,

egal wie intelligent sie noch werden, werden niemals Bewusstsein erlangen. Warum ich das so sehe, werde ich im Buch aufzeigen, es soll schließlich eine Reise zu zwei der schillerndsten Themen der heutigen Natur- und Geisteswissenschaft werden: Intelligenz und Bewusstsein.

Vorher müssen wir jedoch noch wichtige Begriffe definieren. In diesem Kapitel werde ich den Begriff der KI erläutern, danach den des Bewusstseins, und später werde ich zum Begriff des Maschinenbewusstseins überleiten.

Was ist Künstliche Intelligenz?

Der Begriff Künstliche Intelligenz (KI) wurde im Jahr 1956 auf einer Konferenz in Dartmouth (New Hampshire, USA) von führenden Wissenschaftlern ihrer Zeit geprägt. Er war in einem Antrag an die Rockefeller-Stiftung enthalten, bei dem es darum ging, finanzielle Mittel zu bekommen, um neue Verfahren auf Maschinen zu entwickeln, die intelligent sein sollten. (Zu Ihrem besseren Verständnis: IT-Fachleute sprechen oftmals davon, Verfahren »auf« Maschinen oder »auf« Daten zu entwickeln, anstatt die Präpositionen »für« oder »aus« zu nutzen.)

Die Antragsteller waren Enthusiasten und der Meinung, innerhalb weniger Jahre Computersysteme realisiert zu haben, die es mit dem Menschen aufnehmen könnten. Das ist zwar nicht passiert, aber die KI hat dennoch einen unvorstellbaren Siegeszug angetreten. Dabei ist bis heute noch nicht einmal ganz klar, was der Begriff beinhaltet und was nicht.

Daher werde ich gleich zu Beginn einige Präzisierungen vornehmen. Mit dem Begriff »künstlich« ist in der Regel »technisch« gemeint. Es geht um technische Maschinen, also um Verfahren der Physik. Man könnte Intelligenz auch chemisch oder biologisch konstruieren, aber das ist mit dem heute üblichen Begriff KI nicht gemeint. In späteren Kapiteln werden wir sehen, dass von zahlreichen Forschern tatsächlich eine Integration von KI-Maschinen mit biologischen Zellen – zu einer hybriden KI – angedacht ist.

Wie steht es um den Begriff der Intelligenz? Die Definition dieses Begriffs ist schwieriger, weil Intelligenz keine objektiv messbare Eigenschaft wie das Gewicht eines Objektes ist. Über den Begriff »Gewicht« muss man sich mit seinem Gegenüber nicht handelseinig werden, der ist präzise durch die Physik definiert. Bei »Intelligenz« ist das nicht so. Wir legen selbst fest, was wir darunter verstehen wollen. Fragt man beispielsweise Psychologen, so sehen sie in der Intelligenz die Fähigkeiten einer Person zum Denken, Lernen, Wahrnehmen und weiterhin auch zu Konzentration, Abstraktion, Merkfähigkeit und Anpassungsfähigkeit. Aber letztlich hilft das für technische Entwicklungen nicht weiter, denn es gibt ganz offensichtlich sehr viele Merkmale von Intelligenz.

Für Techniker ist auch keine maximale, sondern eine minimale Festlegung von Intelligenzeigenschaften interessant, auf die sich alle einigen könnten. Eine allgemein anzuerkennende Definition wäre beispielsweise folgende: Intelligenz ist die Summe von Denk- und Wahrnehmungsprozessen eines Objektes oder Subjektes, um auf Umwelteinflüsse angemessen zu reagieren.

Wenn man so etwas technisch bauen will, hat man immer noch einige Hürden zu bewältigen. Man muss klären, was die Begriffe »Denken« und »Wahrnehmung« exakt bedeuten. Dass das System angemessen reagieren soll, liegt auf der Hand. Aber jeder Computeralgorithmus reagiert bereits heute schon angemessen auf die Eingaben seiner Umgebung, sonst würde man ihn deinstallieren und die Entwickler nach Hause schicken. Angemessen zu reagieren ist also eher als Minimalanforderung zu verstehen; sie ist notwendig, aber sicher keinesfalls hinreichend, um von intelligenten Systemen zu sprechen.

Kommen wir zum Denken. Menschen können denken, darin sind sich alle einig, aber wie wollte man Maschinen Denken beibringen? Auch hier müssen wir präzisieren. Wir können auf Maschinen heute nur denjenigen Teil des Denkens implementieren, den man algorithmisch erfassen kann, denn auf einem Computer laufen letzten Endes immer nur Algorithmen ab. Wir legen also fest, dass wir mit dem Begriff des Denkens ab jetzt das logische Denken meinen, denn logisches Denken ist mathematisierbar (was bedeutet, dass es durch mathematische Verfahren ausgedrückt werden kann) und damit auf einen Computer

portierbar. (An dieser Stelle einige Begriffshinweise: IT-Fachleute verwenden oft den Begriff »portieren«, und meinen damit »etwas von A nach B übertragen«. Wenn wir von implementieren sprechen, meinen wir einbauen; eine KI-Software in einen Computer implementieren bedeutet also, diese dort einzubauen. Ein weiterer Fachbegriff, der in diesem Buch häufig verwendet wird, ist codieren. Dies bedeutet, eine Information in einem vorgegebenen Zeichenvorrat zu hinterlegen. Das können die Buchstaben des Alphabets sein, aber auch die Bits und Bytes eines Computers beziehungsweise einer Programmiersprache oder der »Zeichenvorrat« an chemischen Molekülen in der menschlichen DNA.)

Dass der Mensch auch zu unlogischem Denken fähig ist, ist allgemein bekannt. Solche Denkprozesse können wir jedoch nicht in mathematischen Verfahren modellieren und auf einem Computer ablaufen lassen.

Erschwerend wird in der Umsetzung noch hinzukommen, dass es verschiedene Arten von mathematischen Logiken gibt, also auch hier werden wir weiter eingrenzen müssen. Eine Logik, die sehr gut auf einen Computer übertragen werden kann, ist die Aussagenlogik, schwieriger wird es bereits bei der sogenannten Prädikatenlogik, doch dazu später.

Und was ist Wahrnehmung? Wahrnehmung bedeutet vereinfacht gesagt, Informationen aufzunehmen und subjektiv bewerten zu können. Subjektives Bewerten können wir bei Computern jedoch nicht nachweisen, da wir nicht wissen, wie ein Computer ein Signal subjektiv aufnimmt. Man könnte die Veränderung der elektrischen Ströme und Spannungen in den Schaltkreisen eines Computers beim Eintreffen von Zahlenkolonnen aus einer Kamera natürlich als Wahrnehmung des Computers definieren, nimmt dann aber in Kauf, dass dies nicht mit dem allgemein üblichen Begriff der Wahrnehmung übereinstimmt. Es führt in der Wissenschaft immer zu Schwierigkeiten, wenn man Begriffe undefiniert, daher ist davon abzuraten.

Wahrnehmung wird im Allgemeinen als subjektiver Erlebnisinhalt, als Erkennung der Bedeutung der Signale durch kognitive Weiterverarbeitung definiert.² Ob Computer das bereits können, ist jedoch strittig und Inhalt des Buches. Obwohl Wahrnehmung einen Großteil der

Intelligenz eines Menschen ausmacht, werde ich diese Eigenschaft bei einer Maschine vorerst wieder ausklammern.

Wir reduzieren den Begriff der Künstlichen Intelligenz also auf das logische Denken und führen folgende Definition ein: Künstliche Intelligenz ist die Summe von logischen Denkprozessen eines Systems, um auf Umgebungseinflüsse angemessen zu reagieren. Konkret geht es hierbei um die Fähigkeit des logischen Schlussfolgerns und die Fähigkeit, selbstständig neue Regeln zu generieren. Diese Art der Künstlichen Intelligenz wurde seit dem Jahre 1956 bis spät in die 1980er-Jahre geschaffen. Es ist die KI der ersten Welle, die sogenannte deduktive Künstliche Intelligenz. Mithilfe dieser KI konnte man riesige Expertensysteme (Datenbanken mit hinterlegten mathematischen Expertenregeln) aufbauen, automatische Beweise in der Mathematik durchführen und sogar das Schachspielen automatisieren. Der Sieg des KI-Computers Deep Blue der Firma IBM gegen den damaligen Weltmeister Garri Kasparow im Schach durch deduktive Verfahren galt als Zäsur.

Selbstverständlich ist die Entwicklung der KI nicht bei einfachen logischen Schlussfolgerungen stehen geblieben. Unter Intelligenz verstehen wir noch mehr. Wir betrachten jemanden auch dann als intelligent, wenn er aus seinen Fehlern lernt. Wenn jemand immer das gleiche Verhalten wiederholt, obwohl das Ergebnis nicht erfolgreich war, dann ist das für uns unintelligent. Formal kann man diese Erkenntnis wie folgt ausdrücken:

1. Intelligenz niederer Stufe besitzt ein System, wenn es adäquat (angemessen) auf Reize seiner Umgebung reagieren kann, und zwar derart, dass das System einen Vorteil, aber zumindest keinen Nachteil aus der Interaktion mit der Umgebung zieht.
2. Intelligenz höherer Ordnung besitzt ein System, wenn es im Fall einer nicht-adäquaten Reizantwort seine Reaktion selbstständig so nachjustieren kann, dass es bei einer erneuten Interaktion mit der Umgebung adäquat reagiert.

Der zweite Punkt bedeutet, dass das System selbstständig lernen können muss. Lernen ist eine zentrale Form der Intelligenz. Technische Systeme sollten das damit auch können. Der Einschub »einen Vorteil ... zieht« ist notwendig, denn ein Lebewesen reagiert dann adäquat, wenn ihm die Interaktion mit der Umwelt nicht zum Schaden gereicht. Hintergrund dieser Präzisierung ist, dass man Intelligenz nicht nutzen kann, um Wahrheiten zu entdecken, wie wir später noch sehen werden, weil der Wahrheitsbegriff in der Wissenschaft nicht eindeutig definiert ist. Ein intelligentes System braucht aber auch keine Wahrheiten über die Welt zu finden. Für eine Fliege im Raum reicht es aus, bei der »Interaktion« mit einer Fensterscheibe zu erkennen, dass es nützlich ist, nicht permanent gegen diese anzufliegen. Falls sie das dennoch bis zur totalen Erschöpfung macht, reagiert sie unangemessen, da sie nicht in der Lage ist, aus ihren Fehlversuchen zu lernen. Das ist eindeutig nicht intelligent. KI-Systeme der zweiten Generation lernen aus ihren Fehlern. Man nennt diese lernenden Systeme auch induktive Systeme.

Neben dem Denken (Deduktion) ist also das Lernen (Induktion) ein wichtiger Bestandteil von Intelligenz. Die Frage ist nun, ob man das Lernen so weit algorithmieren kann, dass auch Lernprozesse auf einem Computer umgesetzt werden können. Die Antwort ist ein klares Ja. Die Mathematik hält eine Vielzahl von Verfahren des Lernens bereit, man nennt sie maschinelles Lernen und/oder Modellierung. Logisches Denken und Lernen können automatisiert werden und die lernende KI ist die KI, die wir seit 20 Jahren im praktischen Einsatz haben, es ist die KI der zweiten Welle. Diese KI-Generation nutzt selbstverständlich auch die KI-Verfahren der ersten Generation. Von dritter Welle oder dritter Generation spricht man, wenn denkende und lernende Systeme zu sogenannten kognitiven und erklärbaren Systemen verschmolzen werden.³ Ein Beispiel dafür ist die berühmte KI IBM Watson. Die Wellen korrespondieren in etwa (aber nicht exakt) mit den im Folgenden eingeführten Intelligenzstufen.

Die Einordnung der Künstlichen Intelligenz

Künstliche Intelligenz besitzen technische Systeme, die logisch korrekt schlussfolgern können und selbstständig in der Lage sind, die Umgebung zu modellieren, um auf Signale der Umgebung adäquat zu reagieren und bei beobachteten Fehlern das eigene Modell über die Umgebung selbstständig so anzupassen, dass bei der nächsten Interaktion eine adäquate Reaktion erfolgen kann.

Extrahiert man den Begriff der heutigen KI also auf seine Essenz, bleiben wie gerade hervorgehoben Denken und Lernen übrig. Um diese Form der Intelligenz besser zu verstehen, muss sie im Kontext anderer Intelligenzen eingeordnet werden.

Man kann im Allgemeinen fünf Intelligenzstufen unterscheiden:⁴

1. Die angemessene Intelligenz (I1) – Adäquate Reaktion auf Reize der Umgebung

Eine berühmte, jedoch nicht lernende KI war Deep Blue, die Schach-KI von IBM, die gegen Garri Kasparow gewonnen hat.

2. Die lernende Intelligenz (I2) – Selbstständiger Erwerb von neuem Wissen

Diese Intelligenzstufe erreichen heute die meisten KI-Systeme. Als Paradebeispiel kann man die Spiele-KI von DeepMind AlphaGo, ansehen; auch die meisten Chatbots fallen in den Bereich, Alexa und Siri, die gesamte KI des autonomen Fahrens, zahlreiche Übersetzungsprogramme wie DeepL und Google Translate und nahezu alle KI-Systeme der Industrie.

3. Die kreative Intelligenz (I3) – Wissensgeneration außerhalb formaler Induktion und Deduktion

Zwei Beispiele für gut simulierte Kreativität im Bereich der Musik sind Ampere Music und IBM Watson Beat, im Bereich der Kunst zum Beispiel das Bild »Edmond de Belamy« als das erste KI-generierte Ölgemälde, das in einem berühmten Aktionshaus versteigert wurde.

4. Die bewusste Intelligenz (I4) – Bewusste Wahrnehmung und bewusstes Verstehen

Hier gibt es heute noch keine Systeme, diese KI-Stufe ist das Thema des Buches.

5. Die selbstbewusste Intelligenz (I5) – Bewusstes Verstehen des ICH-Konzepts

Hier gibt es heute noch keine Systeme, es sind in diesem Jahrhundert auch keine zu erwarten.

Die heutige KI erreicht mindestens die Intelligenzstufen I1 und I2, wie wir bereits gesehen haben. Es gibt mittlerweile aber auch KI-Systeme, die man für kreative Zwecke einsetzt (I3). Dies ist aber eher eine Pseudokreativität. KI-Systeme, die sich vollständig neue Erkenntnisse durch Schlussfolgern und Lernen erschließen können, gibt es (nahezu) nicht. Medienmeldungen, dass ein KI-System kreative Bilder gemalt oder Musik komponieren konnte (siehe oben), stimmen zwar auf den ersten Blick, aber bei näherem Hinsehen erkennt man, dass der KI vorher eine Vielzahl von Kompositionen oder Bildern eintrainiert wurden und sie nun auf Basis der eintrainierten Informationen neue erzeugt. Beim Projekt »The Next Rembrandt« der ING-Bank, Microsoft und anderen bekam die KI typische Portraits Rembrandts zu sehen (die vorher eingescannt und digitalisiert wurden) und erstellte daraus ein neues Portrait im Rembrandtstil.⁵ Genauso war es beim schon erwähnten über 400 000 Dollar teuren KI-Bild »Edmond de Belamy« des Pariser Kollektivs Obvious.⁶ Solche Kunststücke sind technisch möglich, aber wirklich kreativ ist das nicht. Kreativität bedeutet, in völlig neue Gebiete vorzudringen und nicht, bereits Bekanntes zu etwas Neuem zu kombinieren (das nennt man interpolieren).

Dennoch sind die Pseudokreativitäts-Ergebnisse der KI oftmals nicht schlecht. Jeder kann sich selbst einen Eindruck verschaffen. Unter aiva.ai kann man sich zahlreiche KI-Sounds anhören, die AIVA Technologies erzeugt hat.

Vielleicht fragen Sie sich jetzt, ob man KI-generierte Musik von »echter« Musik in einem Blindversuch noch unterscheiden könnte? Solche

Versuche wurden natürlich gemacht. Berühmt geworden sind dadurch Kompositionen des KI-Computers Iamus der Universität Malaga, da Testhörer (des Guardians) nicht herausfinden konnten, was von Menschen komponiert war und was von Iamus. Dennoch handelt es sich hier ebenfalls um Pseudokreativität, da auch bei Iamus Musiksequenzen vorgegeben waren, die der Computer lediglich mittels bestimmter Verfahren miteinander kombinierte. Einige Hörer sagten auch, der KI-Musik fehle die »Seele«, aber das muss jeder selbst entscheiden.⁷ Einen wirklich guten »Verriss« der KI-Kreativität in Kunst und Musik gibt das Video von Christian Ebbertz auf YouTube, das ich nur jedem empfehlen kann und von dem auch die Musikbeispiele oben entstammen.

Als Ingenieur möchte ich noch hinzufügen, dass es bis heute nahezu keine kreative KI in der Industrie gibt, das heißt, die KI hat bisher (fast) keine korrekten Ergebnisse hervorgebracht, die sich außerhalb des Raums der Überlagerung von altbekanntem Wissen befinden.

Systeme der Intelligenzstufen I1 bis I4 gehören zur sogenannten »schwachen« KI. Mit ihr versucht man, menschliche Intelligenz zu simulieren, um sie für den Menschen gewinn- und nutzbringend einzusetzen. Es sind keine Grenzen für das Vordringen dieser KI in Technik und Gesellschaft abzusehen.

KI-Systeme mit Selbstbewusstsein, Gefühlen und eigenem Willen (I5) nennt man »starke« KI. Im Unterschied zu KI-Systemen der Stufen I1 bis I4 haben diese nicht mehr das primäre Ziel, den Menschen zu unterstützen, sondern treten als autonome, selbstbewusste Systeme in direkte Konkurrenz zum Menschen selbst. Aber hier handelt es sich nur um Konstrukte in den Köpfen von Grundlagenforschern, Autoren oder Filmproduzenten, sie liegen völlig außerhalb der Reichweite heutiger technischer Entwicklungen. Wahrscheinlich dauert es noch weit über hundert Jahre, bis die Menschheit solche Systeme bauen kann – wenn überhaupt. Für Systeme auf rein mineralischer (anorganischer) Basis kann man Selbstbewusstsein mit Wünschen und Wollen sogar komplett ausschließen, wie ich in diesem Buch noch zeigen werde.