

TOBY WALSH

IT'S ALIVE

Wie Künstliche Intelligenz
unser Leben verändern wird



Toby Walsh

IT'S ALIVE

Wie Künstliche Intelligenz
unser Leben verändern wird

Aus dem Englischen von
Naemi Schuhmacher und Sonja Schuhmacher

 Edition
Körber

Die englische Originalausgabe erschien 2017 unter dem Titel »It's Alive! Artificial Intelligence from the Logic Piano to Killer Robots« bei La Trobe University Press & Black Inc., Melbourne/Australia.

Copyright © Toby Walsh 2017

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://dnb.d-nb.de> abrufbar.

© Edition Körber, Hamburg 2018

Umschlag: Groothuis. www.groothuis.de

Coverfoto: iStock/mennovandijk

Herstellung: Das Herstellungsbüro, Hamburg |
www.buch-herstellungsbuero.de

ISBN 978-3-89684-545-0

Alle Rechte vorbehalten

www.edition-koerber.de

Inhalt

Prolog	7
Die Vergangenheit der Künstlichen Intelligenz	
1. Der Traum von der Künstlichen Intelligenz	23
2. Wie man Künstliche Intelligenz messen kann	67
Die Künstliche Intelligenz in der Gegenwart	
3. Künstliche Intelligenz heute	89
4. Die Grenzen der Künstlichen Intelligenz	129
5. Künstliche Intelligenz und die Folgen	192
Die Zukunft der Künstlichen Intelligenz	
6. Technologischer Wandel	263
7. Zehn Vorhersagen	281
Nachwort	301
Anmerkungen	305
Literatur	331
Stichwortverzeichnis	335
Über den Autor	346

Prolog

Ich möchte mit einem Zitat aus dem Jahr 1950 beginnen. Die Welt war damals in vieler Hinsicht noch bedeutend einfacher: Das Fernsehen sendete nur in Schwarzweiß, Verkehrsflugzeuge hatten noch keine Düsentriebwerke, der Siegeszug des Transistors stand noch bevor. Und weltweit gab es nicht einmal ein Dutzend Computer¹, jeder einzelne ein fantastisches, riesengroßes Gebilde aus Elektronenröhren, Relais, Schalttafeln und Kondensatoren.

So gehörte schon einige Kühnheit zu der Prophezeiung: »[Ich glaube,] dass am Ende des [20.] Jahrhunderts der Gebrauch von Wörtern und die allgemeinen Ansichten der Gebildeten sich so sehr geändert haben werden, dass man ohne Widerspruch von denkenden Maschinen reden können.«² Was für eine Idee! Denkende Maschinen! Wird es das in naher Zukunft tatsächlich geben? Und wenn, wie lange wird es dauern, bis sie uns überlegen sind?

Doch zunächst – von wem stammt diese kühne Prognose? Welche Relevanz hatte sie? Nun, die Zeitschrift *Time* zählte die Person, die sie traf, im Jahr 1999 immerhin zu den »100 wichtigsten Persönlichkeiten des 20. Jahrhunderts«.³ Ohne Zweifel gehört sie zu den rätselhaftesten Denkern des Jahrhunderts. Ein Mathema-

tiker. Ein Kriegsheld. Und vor allem ein Visionär, dessen Träume uns bis heute, lange nach seinem Tod, beeinflussen.

Ein oscarprämierter Film schildert, welchen entscheidenden Beitrag er im Zweiten Weltkrieg dazu geleistet hat, den Code der deutschen Verschlüsselungsmaschine »Enigma« zu knacken. Winston Churchill bezeichnete ihn und sein Team als »die Gänse, die das goldene Ei legten – ohne darüber ein großes Geschnatter zu veranstalten«. Historiker sind sich einig, dass die Entschlüsselung des Enigma-Codes den Zweiten Weltkrieg um mindestens zwei Jahre verkürzte. Und damit zweifellos Millionen Menschen das Leben rettete. Doch am Ende des 21. Jahrhunderts wird man sich an den Urheber dieses Zitats wohl nicht in erster Linie deshalb erinnern, weil er ein Meister-Dechiffrierer war.

Beim Knacken des Enigma-Codes legte er nicht nur das theoretische Fundament des Computers an sich, sondern trug entscheidend dazu bei, einen der ersten einsatzfähigen Computer zu bauen, *the bombe* genannt.⁴ Seine Ideen sind bis heute aus der Computerwissenschaft nicht wegzudenken. Noch bevor es überhaupt den ersten Computer gab, entwickelte er ein grundlegendes und vollkommen allgemeines Modell einer solchen Maschine.⁵ In Anerkennung seiner Leistungen trägt heute die höchste Auszeichnung der Informatik, der Turing-Award, seinen Namen. Doch trotz des enormen Einflusses, den Computer in allen Bereichen unseres Lebens haben, bezweifle ich, dass man sich dereinst hauptsächlich deshalb an ihn erinnern wird, weil er einen Großteil des Fundaments der Computerwissenschaft gelegt hat.

Mit einem einzigen Artikel, *The Chemical Basis of Morphogenesis*, der in einer altherwürdigen wissenschaftlichen Zeitschrift erschien, beeinflusste Turing zudem nachhaltig einen Zweig der Biologie, der Morphogenese genannt wird.⁶ In dieser Publikation hatten vor ihm schon Geistesgrößen wie Charles Darwin (des-

sen Evolutionstheorie das Selbstbild des Menschen entscheidend veränderte), Alexander Fleming (der durch seine Entdeckung des Penicillins vielen Millionen Menschen das Leben rettete) und James Watson, Francis Crick und Dorothy Hodgkin (die mit ihrer Beschreibung der Struktur der DNA die genetische Revolution auslösten) publiziert. Turing stellte eine Theorie über die Musterbildung in der Natur auf, die erklärte, wie sich Streifen, Flecken und Spiralen bei Pflanzen und Tieren bilden. Seine Arbeit wird bis zum heutigen Tag häufig zitiert. Doch auch seine Theorie über die Musterbildung wird wahrscheinlich nicht die Leistung sein, die man am Ende des 21. Jahrhunderts in erster Linie mit seinem Namen verknüpft.

Alan Turing ist auch eine wichtige Persönlichkeit der Schwulenbewegung. In den 1950er Jahren war Homosexualität in Großbritannien noch verboten. Viele geben bis heute der strafrechtlichen Verfolgung und der anschließenden chemischen Kastration, der sich Turing unterziehen musste, die Schuld an seinem frühen Tod und werfen der britischen Gesellschaft, für die er im Zweiten Weltkrieg so viel getan hatte, vor, ihn im Stich gelassen zu haben.⁷ Im Jahr 2009 entschuldigte sich Premierminister Gordon Brown nach einer öffentlichen Kampagne formell für den Umgang mit Turing. Vier Jahre später unterzeichnete die Queen eine königliche Begnadigung, ein äußerst seltener Vorgang. Für manche Schwule ist Turing begreiflicherweise ein Märtyrer. Und dennoch bezweifle ich, dass man sich am Ende des 21. Jahrhunderts vor allem wegen dieser Sache an ihn erinnern wird.

Aber weswegen wird man sich nun vor allem an ihn erinnern? Ich vermute, dass es ein Artikel sein wird, den er in der dem breiten Publikum eher unbekannteren philosophischen Zeitschrift *Mind* veröffentlichte, dem auch das Zitat zu Beginn dieses Kapitels entstammt. Diese Publikation war bis dahin vor allem

durch die Veröffentlichung von »Was die Schildkröte zu Achill sagte« hervorgetreten, einem Text von Lewis Carroll über die Logik eines Zenon-Paradoxons.⁸ Der Aufsatz, aus dem unser Zitat stammt, gilt heute als einer der wichtigsten in der Geschichte der Künstlichen Intelligenz.⁹ Er blickt in eine Zukunft, in der Maschinen denken. Als unser Autor ihn schrieb, gab es auf der ganzen Welt gerade mal ein Dutzend Computer, allesamt riesengroß und unvorstellbar teuer. Ihre Rechenleistung wird heute von jedem Smartphone übertroffen. Es war also keine Kleinigkeit, schon zum damaligen Zeitpunkt vorausszusehen, welchen Einfluss Computer dereinst auf unser Leben haben würden, und eine noch viel größere Leistung, sich vorzustellen, dass sie eines Tages selbst *denken* würden. Der Aufsatz nimmt bereits sämtliche wichtigen Argumente, die in der Folge gegen denkende Maschinen erhoben wurden, vorweg und widerlegt sie. Sein Autor wird deshalb von vielen als einer der Väter des Forschungsgebiets Künstliche Intelligenz angesehen.

Die Rede ist hier natürlich von Alan Mathison Turing, Mitglied der Royal Society, geboren 1912, 1954 von eigener Hand aus dem Leben geschieden.

Ich sage voraus, dass man am Ende des 21. Jahrhunderts Alan Turing vor allem deshalb gedenken wird, weil er das Fundament des Forschungsgebiets gelegt hat, das den Versuch unternimmt, denkende Maschinen zu bauen. Diese Maschinen werden unser Leben nicht weniger dramatisch umgestalten als die Dampfmaschine zu Beginn des Industriezeitalters. Sie werden verändern, wie wir arbeiten, was wir spielen, wie wir unsere Kinder erziehen, wie wir die Kranken und die Alten behandeln und schließlich auch wie man sich an die Menschheit erinnern wird. Mit an Sicherheit grenzender Wahrscheinlichkeit werden sie die Welt stärker umkrepeln als alle unsere bisherigen Schöpfungen. Die

denkenden Roboter, die bislang nur die Science-Fiction bevölkerten, stehen kurz davor, Einzug in die Realität zu halten.

Unser Leben bewegt sich in Riesenschritten auf die Zukunft zu, die sich Science-Fiction-Autoren erträumt haben. Die Taschencomputer, die wir nun ständig mit uns herumtragen, können die kompliziertesten Fragen beantworten, uns mit Spielen und Filmen unterhalten, uns den Weg zeigen, sie helfen uns, eine Arbeit oder einen Partner zu finden, sie spielen uns ein Liebeslied, wenn uns danach ist, und setzen uns jederzeit mit Freunden rund um die Welt in Verbindung. Dass man mit ihnen auch noch telefonieren kann, ist längst eine ihrer unspektakulärsten Eigenschaften geworden.

Natürlich wirft meine Prognose, dass Turings bleibender Beitrag seine Starthilfe für das Forschungsgebiet der Künstlichen Intelligenz sein wird, gleich mehrere Fragen auf. Werden sich am Ende des Jahrhunderts diese denkenden Maschinen an ihn erinnern? Steuern wir auf eine lebenswerte Zukunft zu? Werden Roboter alle mühseligen und gefährlichen Arbeiten übernehmen? Wird durch sie die Wirtschaft aufblühen? Werden wir weniger arbeiten und mehr Freizeit genießen? Oder wird Hollywood mit seinen düsteren Zukunftsvisionen recht behalten? Werden die Reichen noch reicher, alle anderen ärmer werden? Werden viele von uns ihre Arbeit verlieren? Oder kommt es noch schlimmer, werden die Roboter schließlich alles übernehmen? Sind wir vielleicht schon dabei, die Saat unseres eigenen Untergangs auszubringen?

Das sind die Fragen, denen in diesem Buch nachgegangen werden soll; zudem liefert es eine Prognose, wohin uns die Künstliche Intelligenz führen wird. In Teil I beschäftige ich mich damit, was wir von der Vergangenheit lernen können. Wie sich eine bestimmte Technik weiterentwickelt, versteht man wohl am bes-

ten, wenn man sich anschaut, wie sie entstanden ist. In Teil II stelle ich dar, wo die Künstliche Intelligenz (KI) heute steht, und wäge Risiken und Nutzen des Baus intelligenter Maschinen ab. Ich versuche, realistisch einzuschätzen, was wir damit erreichen können. Es ist zweifellos eine kühne und ambitionierte Idee, denkende Maschinen zu bauen, und wenn sie Erfolg hat, wird dies ungeheuren Einfluss auf die Gesellschaft haben. Schließlich werde ich in Teil III die Zukunft der KI mehr im Detail diskutieren. Welche der abenteuerlichen Zukunftsvorstellungen, die in Büchern und Filmen entwickelt wurden, werden vielleicht tatsächlich eintreten? Wird die Wirklichkeit die Fantasie womöglich noch übertreffen? Und um sozusagen auch ein wenig eigenes Kapital zu investieren, werde ich zehn Prognosen darüber wagen, was die KI bis zum Jahr 2050 leisten wird. Die eine oder andere wird Sie womöglich überraschen.

Auf dem Feld der KI tut sich derzeit viel. In den letzten fünf Jahren sind Unternehmen, die sich mit Künstlicher Intelligenz beschäftigen, Milliarden Dollar an Wagniskapital zugeflossen.¹⁰ Es wird hoch gepokert. Der Computergigant IBM hat eine Milliarde Dollar in seinen Superrechner Watson gesteckt, der Cognitive Computing, die Simulation menschlicher Denkprozesse, betreibt.¹¹ Toyota lässt für eine Milliarde Dollar autonome Fahrzeuge erforschen. OpenAI, eine Initiative, die sichere und alltags-taugliche KI entwickeln will, wird ebenfalls mit einer Milliarde Dollar unterstützt. Der SoftBank Vision Fund, gestartet im Oktober 2015, stellt mit wesentlicher Beteiligung Saudi-Arabiens insgesamt 100 Milliarden Dollar Kapital für Technologieunternehmen mit dem Schwerpunkt KI und »Internet der Dinge« bereit. Viele weitere Mitstreiter der Technologieszene, darunter Google, Facebook und Baidu, investieren in Künstliche Intelligenz. Es sind also spannende Zeiten für alle, die auf diesem Gebiet arbeiten.

Bei dem Einsatz solcher Summen ist es gut möglich, dass sich die Entwicklung hin zu denkenden Maschinen noch beschleunigt.

Wozu dieses Buch lesen?

Computer verändern unser Leben heute in Riesenschritten. Kein Wunder, dass weltweit das Interesse daran wächst, mehr über Künstliche Intelligenz zu erfahren. Kühne Prognosen sind im Umlauf. Im Mai 2016 verkündete Dave Coplin, der Chefvisionär von Microsoft UK, Künstliche Intelligenz sei »die wichtigste Technologie, an der heute auf diesem Planeten gearbeitet wird ... sie wird unser Verhältnis zur Technologie überhaupt verändern. Sie wird die Beziehungen der Menschen untereinander verändern. Ich denke, sie wird sogar verändern, was es für uns heißt, Mensch zu sein.«

Einen Monat zuvor hatte Google-CEO Sundar Pichai die zentrale Stellung der KI in Googles Strategie beschrieben. »Ein Hauptantrieb ... ist unsere Langzeit-Investition in Maschinenlernen und KI ... Mit Blick in die Zukunft ... werden wir uns von einer Welt, in der sich alles um das Mobiltelefon dreht, zu einer entwickeln, in der KI das wichtigste ist.«

Doch manche sehen mit der KI auch Gefahren, ja sogar das baldige Ende der Menschheit heraufziehen. Im Jahr 2014 warnte Elon Musk seine Zuhörer vom MIT: »Wir sollten sehr vorsichtig mit der Künstlichen Intelligenz umgehen. Wenn man mich fragt, was ich für die größte existenzielle Bedrohung halte, dann denke ich sofort an die KI.« Musk ist der Multiunternehmer, Erfinder und Investor, der mit der Gründung von PayPal, Tesla und

SpaceX bekannt wurde. Er hat den Bankensektor, die Auto- und die Raumfahrtindustrie aufgerüttelt, er weiß, wie Technik, insbesondere die Computertechnik, die Welt umkrepeln kann. Und Musk steht mit seinem Geld für die Annahme ein, dass die KI eine ernsthafte Bedrohung für die Menschheit sein könnte. Anfang 2015 spendete er dem Future of Life Institute 10 Millionen Dollar für Forschungsarbeiten an einer sicheren Künstlichen Intelligenz. Das ist vielleicht nicht so viel für jemanden wie Musk, der mit einem Vermögen von 10 Milliarden Dollar zu den 100 reichsten Personen der Welt gehört. Doch noch im gleichen Jahr erhöhte er sein Engagement in diesem Bereich beträchtlich, indem er ankündigte, sich als einer der Hauptsponsoren am 1 Milliarde Dollar schweren Projekt OpenAI zu beteiligen, dessen Ziel es ist, Künstliche Intelligenz ohne Gefahr für die Welt und auf Open-Source-Basis zu entwickeln.

Der jüngst verstorbene Physiker Stephen Hawking sah das Gefahrenpotenzial der Künstlichen Intelligenz ähnlich wie Musk. Es ist nicht ohne Ironie, dass Hawking ausgerechnet ein Update seines Sprachgenerators zum Anlass nahm, mit der elektronischen Stimme, die ihm dieses Gerät verlieh, zu warnen: »Die Entwicklung einer echten künstlichen Intelligenz könnte das Ende der Menschheit herbeiführen.«

Etliche Größen der Computerbranche, darunter Bill Gates von Microsoft und Steve Wozniak von Apple, prognostizieren, dass von der KI Gefahren ausgehen. Der Vater der Informationstheorie, Claude Shannon, schrieb 1987: »Ich sehe eine Zeit kommen, in der wir für die Roboter das sind, was heute die Hunde für uns Menschen sind ... ich bin ganz auf der Seite der Maschinen!«¹² Und auch Alan Turing persönlich warnte bereits 1951 in einem Radiobeitrag der BBC:

Wenn Maschinen denken können, dann tun sie das vielleicht auf intelligentere Weise als wir, und was wird dann aus uns? Selbst wenn wir die Oberhand gegenüber diesen Maschinen behalten, indem wir beispielsweise dafür sorgen, ihnen jederzeit den Strom abstellen zu können, sollte das die Menschheit doch sehr demütig stimmen ... Das ... kann einem wirklich Angst machen.

Doch nicht alle Technologen und Technokraten sehen wegen denkender Maschinen gleich das Ende der Menschheit heraufziehen. Facebook-Chef Mark Zuckerberg versuchte solche Bedenken im Januar 2016 zu zerstreuen: »Ich bin überzeugt, dass wir die KI so aufbauen, dass sie für uns arbeitet und uns hilft. Manche schüren Ängste angesichts großer Gefahren, die von der KI ausgehen, aber für mich klingen sie weit hergeholt und viel weniger wahrscheinlich als die Katastrophen, die uns von Seuchen, Gewalt usw. drohen.« Und Andrew Ng, ein führender KI-Forscher, tätig für Chinas Internet-Riesen Baidu, meint: »Angst vor der KI ist wie Angst vor der Überbevölkerung auf dem Mars.« (Wobei man allerdings erwähnen sollte, dass eines der spektakulären Projekte von Elon Musk die Besiedlung des roten Planeten ist ...)

Wem soll man nun glauben? Muss es uns nicht nachdenklich stimmen, dass sich Technologie-Giganten wie Musk und Zuckerberg in dieser Frage nicht einig sind? Die KI ruft schon lange Ängste hervor. Einer der größten Visionäre überhaupt, der Science-Fiction-Autor Arthur C. Clarke, warnte schon 1968 vor gefährlichen Folgen der Künstlichen Intelligenz. Clarke hat als Technologievisionär eine verblüffende Erfolgsbilanz. Er hat geostationäre Satelliten, eine globale digitale Bibliothek (die wir heute Internet nennen), Maschinenübersetzung und mehr vorausgesagt. Und der Computer HAL 9000 in dem Film 2001 –

Odysee im Weltraum, zu dem er das Drehbuch schrieb, ist eine eindringliche Veranschaulichung, was es heißen könnte, wenn die KI die Kontrolle übernimmt.

Visionäre wie Clarke haben in mir den Traum von der Künstlichen Intelligenz geweckt, als ich noch ein kleiner Junge war. Und ich habe mein ganzes Leben lang auf diesem Gebiet gearbeitet und versucht, diesen Traum wahr zu machen. Daher lässt es mich nicht kalt, wenn Leute von *außerhalb* dieses Forschungsgebiets, noch dazu, wenn es sich um äußerst intelligente Physiker und erfolgreiche Technologie-Unternehmer handelt, prophezeien, dass die KI das Ende der Menschheit herbeiführen wird. Vielleicht sollte man auch mal hören, was die Leute, die an vorderster Front stehen, zu sagen haben? Oder stecken wir einfach mit der Nase zu tief in dem, was wir tun, um die Risiken zu erkennen? Und warum sollten wir an etwas arbeiten, das unsere eigene Existenz zerstören könnte?

Womöglich löst die Idee der KI tiefsitzende Ängste bei uns aus, solche, wie sie beispielsweise im Mythos von Prometheus beschrieben werden, der Geschichte des griechischen Gottes, der den Menschen das Feuer brachte, mit dem sie so viel Gutes *und* so viel Böses anrichteten. Es ist die Angst, dass sich unsere eigenen Schöpfungen gegen uns wenden könnten, die Mary Shelley in ihrem Roman *Frankenstein* schilderte. Dass es sich dabei um uralte Befürchtungen handelt, heißt allerdings noch nicht, dass sie unbegründet sind. Die Menschheit hat vieles erfunden, über das sie gründlich hätte nachdenken sollen, und manchmal hat sie es sogar getan: die Atombombe, das Klonen, Laser, die zur Erblindung führen können, oder soziale Medien, um nur einiges zu nennen. Mit diesem Buch möchte ich Ihnen bei der Einschätzung helfen, wie viel Kopfzerbrechen uns die denkenden Maschinen der Zukunft tatsächlich bereiten sollten.

Ein Teil der Verantwortung liegt bei uns, den Wissenschaftlern, die an Künstlicher Intelligenz arbeiten. Wir haben zu wenig und oft auch zu unverständlich kommuniziert. Wir müssen besser darstellen, was wir tun und wohin es die Gesellschaft führen könnte. Die Wissenschaft muss hier mehr Verantwortung übernehmen. Und es ist umso wichtiger, wenn, wie ich in diesem Buch darstelle, diese Veränderung hauptsächlich eine gesellschaftliche ist, denn die Gesellschaft wandelt sich naturgemäß langsamer als die Technik. Die KI selbst ist moralisch neutral, wie Technik meistens. Aber sie kann zu Gutem wie zu Schlechtem führen.

Die vielen Missverständnisse, die über die KI im Umlauf sind, erschweren allerdings die Debatte. Einige hoffe ich zu klären. Zunächst einmal tendieren viele Menschen dazu, die derzeitigen und zukünftigen Möglichkeiten der Künstlichen Intelligenz zu überschätzen, und dies umso mehr, wenn sie nicht vom Fach sind. Sie sehen, dass ein Computer besser Go spielen kann als jeder Mensch, und da sie selbst nicht besonders gut Go spielen¹³, folgern sie, dass ein Computer auch alle möglichen anderen Aufgaben bewältigen kann, die menschliche Intelligenz erfordern. Oder zumindest dass es nicht allzu schwer sein kann, ihm auch andere anspruchsvolle Aufgaben zu übertragen. Doch das Go-Programm ist wie alle unsere derzeitigen Computerprogramme ein *Idiot savant*: Es verfügt nur über eine Inselbegabung, es ist bloß in einer einzigen Sache wirklich gut. Nicht einmal andere Spiele wie Schach oder Poker beherrscht es. Man müsste es komplett umschreiben, um ihm irgendein anderes Spiel beizubringen. Es wird nicht eines Tages aus dem Schlummer erwachen, feststellen, dass es langweilig ist, gegen Menschen Go zu spielen, und beschließen, sein Glück im Online-Poker zu versuchen. Und schon gar nicht wird es eines Tages damit beginnen, von

der Weltherrschaft zu träumen. Es hat keine Wünsche. Es ist ein Computerprogramm und kann ausschließlich das, wofür es programmiert wurde – nämlich besonders gut Go spielen.

Andererseits werde ich auch darlegen, dass wir dazu neigen, die Veränderungen, die neue Techniken auf Dauer mit sich bringen, zu unterschätzen. Smartphones gibt es erst seit einem Jahrzehnt, und sie haben unser Leben bereits gewaltig verändert. Und das Internet, dessen Siegeszug erst vor zwei Jahrzehnten begann, hat beinahe sämtliche Aspekte unseres Lebens verändert. Der Multiplikatoreffekt der Technik bringt es mit sich, dass für die nächsten zwanzig Jahre mit noch größeren Veränderungen zu rechnen ist. Exponentielles Wachstum ist für uns Menschen, die wir von der Evolution darauf getrimmt sind, uns mit unmittelbar drohenden Gefahren auseinanderzusetzen, nur schwer verständlich. Auf Langzeitrisiken sind wir so schlecht eingestellt wie auf die Begegnung mit schwarzen Schwänen.¹⁴ Würden wir wirklich begreifen, was langfristige Entwicklungen bedeuten, würde niemand mehr Lotto spielen und dafür mehr auf die hohe Kante legen. Die Verbesserungen, die akkumuliertes Wachstum mit sich bringt, sind für unser Gehirn, das auf das Streben nach Lust und die Vermeidung von Unlust programmiert ist, schwer zu erfassen. Wir leben für den Augenblick.

Doch bevor Sie tiefer in das Buch einsteigen, muss ich Sie warnen: Die Futurologie gehört zu den inexakten Wissenschaften. Der dänische Physiker und Nobelpreisträger Niels Bohr konstatierte dazu ironisch: »Voraussagen sind sehr schwierig, besonders, wenn sie die Zukunft betreffen.« Ich kann nur hoffen, dass das Bild, das ich mit breiten Strichen zeichne, wenigstens in groben Zügen korrekt ist, in manchen Details ist es mit Sicherheit falsch. Doch indem ich diese Ideen ausbreite, hoffe ich, Ihr Verständnis dafür zu fördern, warum ich und Tausende meiner

Kollegen unser Leben der Erforschung des spannenden Wegs gewidmet haben, der uns zu den denkenden Maschinen führt. Und ich hoffe, Sie werden verstehen, warum es ein Weg ist, den wir erforschen sollten – sogar müssen –, wenn wir weiterhin die Lebensqualität auf diesem Planeten verbessern wollen. Es gibt etliche Bereiche, die es uns geradezu gebieten, Künstliche Intelligenz zu entwickeln, denn sie wird viele Menschenleben retten.

Vor allem hoffe ich, dass Sie bedenken, wie sich die Gesellschaft selbst ändern muss. Die Botschaft dieses Buches ist letztlich, dass die Künstliche Intelligenz uns viele Pfade zeigen kann, gute ebenso wie schlechte, aber dass die Gesellschaft wählen muss, welchen sie einschlagen will, *und* dass sie dieser Wahl entsprechend handeln muss. Es gibt viele Entscheidungen, die wir Maschinen übertragen können. Doch ich möchte dafür plädieren, dass wir dies nur sehr sparsam tun – selbst dann, wenn Maschinen diese Entscheidungen besser treffen können als wir. Jedenfalls muss die Gesellschaft damit beginnen, sich darüber Gedanken zu machen, was wir Maschinen anvertrauen wollen.

Wer sollte dieses Buch lesen?

Dieses Buch wendet sich an interessierte Leser, aber nicht an Spezialisten des Fachs. An Leute, die verstehen möchten, wohin uns die Künstliche Intelligenz führen kann. Werden einige der fantastisch klingenden Vorhersagen über denkende Maschinen wirklich eintreten? Wird es zu einer technologischen Singularität kommen? Sollte man sich Sorgen darum machen, wohin uns die KI führt? Welchen Einfluss wird sie auf uns und unsere Kinder haben? Und wie lang wird es dauern, bevor einige dieser

Voraussagen eintreten? Um den Lesefluss nicht zu unterbrechen, habe ich Quellen und zusätzliche technische Anmerkungen ans Ende des Buches gestellt. Sie können sie beim Lesen des Buchs getrost ignorieren.¹⁵ Wenn Sie jedoch ein Punkt stärker interessiert, dann werden Ihnen diese Anmerkungen zusätzliche Informationen liefern und weiterführende Literatur empfehlen.

Fangen wir an.

Teil I

Die Vergangenheit der Künstlichen Intelligenz

1. Der Traum von der Künstlichen Intelligenz

Wenn wir wissen wollen, wohin uns die Künstliche Intelligenz führen wird, ist es hilfreich zu verstehen, woher sie kommt und wo sie heute steht.

Der offizielle Startschuss der KI fiel im Jahr 1956, als einer ihrer Väter, John McCarthy¹, den Begriff auf einer Konferenz in Dartmouth prägte.² Es ist kein ganz unproblematischer Begriff, wie ja auch Intelligenz an sich schon ungenügend definiert ist. Und es klingt nie gut, wenn man das Adjektiv *künstlich* vor etwas setzt. Witze über die natürliche Intelligenz und künstliche Dummheit blieben natürlich nicht aus. Doch nun müssen wir mit dem Namen *Künstliche Intelligenz* leben. Die Geschichte der KI reicht allerdings viel weiter zurück – sogar bis in die Zeit vor der Erfindung des Computers. Die Idee denkender Maschinen oder die Frage, wie wir Denken überhaupt nachbilden können, beschäftigt die Menschheit schon seit Jahrhunderten.

Wie so viele Geschichten hat auch diese keinen klaren Anfang. Allerdings ist sie eng mit der Geschichte der Logik verknüpft. Ein möglicher Ausgangspunkt ist das dritte Jahrhundert vor Christus, als Aristoteles das Fundament der formalen Logik legte. Ohne Logik gäbe es auch nicht die modernen Digitalrechner. Und Logik wurde und wird oft als Modell des Denkens gesehen,

ein Mittel, unseren Gedanken und Argumenten Präzision zu verleihen.

Abgesehen von einigen mechanischen Geräten, die für astronomische Bestimmungen und einfache Berechnungen eingesetzt wurden, hat die Menschheit nach Aristoteles 2000 Jahre lang kaum Fortschritte in Richtung denkender Maschinen gemacht. Doch wir wollen nicht zu streng sein, selbst in den am weitesten entwickelten Ländern hatten die Menschen ein paar andere Probleme: Kriege, Seuchen, Hunger und das finstere Mittelalter.

Eine Ausnahme ist der katalanische Autor, Dichter, Theologe, Mystiker, Mathematiker, Logiker und Märtyrer Ramon Llull, der im 13. Jahrhundert lebte.³ Manche sehen in Llull einen der Väter des Computers. Er konstruierte eine einfache, rechen-scheiben-artige »logische Maschine«, mit der er rein *mechanisch* sämtliche wahren Aussagen über einen Gegenstand ermitteln wollte. Dies kann als einer der ersten logisch-mechanischen Versuche zur Produktion von Wissen betrachtet werden. Llull fand damit zwar in seiner Zeit keinen großen Widerhall, soll aber die nächste Person in unserer Geschichte maßgeblich beeinflusst haben.

Lasst uns rechnen

Als sich der geistige Nebel des Mittelalters verzog, nahm unsere Geschichte Fahrt auf. Mit Gottfried Wilhelm Leibniz betritt eine ihrer herausragenden Figuren die Bühne.⁴ Eine seiner weitsichtigsten Beiträge zur Ideengeschichte war, dass er es für möglich hielt, menschliches Denken irgendwie auf Berechnungen zu reduzieren, und dass sich durch solche Berechnungen Fehler in unseren Überlegungen entdecken und Meinungsverschiedenheiten

lösen ließen. Er schrieb: »Das einzige Mittel, unsere Schlussfolgerungen zu verbessern, ist, sie ebenso anschaulich zu machen, wie es die der Mathematiker sind, derart, dass man seinen Irrtum mit den Augen findet und, wenn es Streitigkeiten unter Leuten gibt, man nur zu sagen braucht: ›Rechnen wir!‹ ohne eine weitere Förmlichkeit, um zu sehen, wer recht hat.«⁵

Leibniz schlug ein einfaches logisches Verfahren zur Durchführung solcher Berechnungen vor. Ihm schwebte dabei ein »Alphabet des menschlichen Denkens« vor, das jeden Grundbegriff durch ein spezielles Symbol darstellen sollte. Computer sind letztlich nichts anderes als Maschinen zur Manipulation solcher Symbole.⁶ Damit führte Leibniz eine wesentliche Abstraktion auf dem Weg zu denkenden Computern ein. Das ist so zu verstehen: Auch wenn Computer *nur* Symbole manipulieren, dann können sie doch, wenn diese Symbole für Grundbegriffe stehen, wie Leibniz es vorschlug, neue Begriffe ableiten und auf diese Weise menschenähnlich denken.

Ungefähr zur gleichen Zeit steuerte ein anderer Philosoph, Thomas Hobbes, seinen Beitrag zum philosophischen Fundament der denkenden Maschinen bei.⁷ Ganz wie Leibniz setzte auch Hobbes Denken mit Rechnen gleich: »Vernunft ... ist nichts anderes als Rechnen, das heißt Addieren und Subtrahieren.«⁸

Denken in Analogie zum Rechnen zu sehen, wie das Leibniz und Hobbes gleichermaßen taten, ist ein erster Schritt auf dem Weg zum Bau einer denkenden Maschine. Mechanische Rechengeräte gab es zwar schon vor Hobbes und Leibniz, aber es sollten noch fast zwei Jahrhunderte vergehen, bevor jemand versuchte, Denken durch Rechenoperationen in die Praxis umzusetzen.⁹

Ein weiterer Universalgelehrter der frühen Neuzeit ist René Descartes.¹⁰ Eine seiner großen philosophischen Gedanken, »cogito ergo sum«, »Ich denke, also bin ich«, ist heute noch Ge-

genstand von Diskussionen innerhalb der KI. Mit diesen drei lateinischen Wörtern verknüpft er elegant das Denken mit der (menschlichen) Existenz. Der Umkehrschluss lautet, dass man nicht denken kann, wenn man nicht existiert.¹¹ Damit stellt Descartes nicht weniger als die bloße *Möglichkeit* denkender Maschinen in Frage. Maschinen existieren nicht auf dieselbe Weise wie wir. Ihnen fehlen etliche spezifische Eigenschaften, die wir mit der menschlichen Existenz verknüpfen: Gefühle, Moral, Bewusstsein und Kreativität, um nur einige zu nennen. Diese Eigenschaften wurden als Argumente gegen die Möglichkeit denkender Maschinen ins Feld geführt, wie wir noch sehen werden. Wie zum Beispiel: Maschinen hätten kein Bewusstsein und könnten nicht denken. Oder ihre fehlende Kreativität spreche dafür, dass sie nicht denken könnten. Wir werden uns diese Argumente in Kürze genauer ansehen.

Boole und Babbage

Die nächste wichtige Figur in unserer Geschichte taucht erst 200 Jahre später auf. George Boole war als Mathematiker Autodidakt.¹² Obwohl er nie eine Universität besucht hatte, wurde er 1849 in Anerkennung einiger mathematischer Artikel, die er neben seiner Arbeit an einer von ihm selbst gegründeten Schule verfasst hatte, zum ersten Mathematikprofessor ans Queen's College im irischen Cork berufen. Dieser Lehrstuhl, der sich am Rand der damaligen akademischen Welt befand, gab ihm die Freiheit, Gedanken zu verfolgen, die sich als zentral für die Entwicklung der Computertechnik erweisen sollten, und vom Bau denkender Maschinen zu träumen. Boole hatte die Idee, Logik durch alge-

braische Operationen zu formalisieren, die auf zwei Werten beruhen: wahr oder falsch, an oder aus, 0 oder 1. Diese sogenannte Boole'sche Logik beschreibt die Operationen sämtlicher heutiger Computer; sie sind nichts anderes als hochentwickelte Maschinen, die einen unablässigen Strom Boole'scher Nullen und Einsen verarbeiten. Seinen Zeitgenossen war die Bedeutung von Booles Ideen nicht bewusst, doch es ist keineswegs übertrieben, ihn als den Vater unserer Informationsgesellschaft zu bezeichnen.

Boole hatte jedoch weit ambitioniertere Pläne, mit denen er seiner Zeit weit voraus war. Der Titel seines umfassendsten Werks über seine Ideen zur Logik gibt uns eine Idee davon: *An Investigation of the Laws of Thought*, zu Deutsch »Eine Untersuchung der Gesetze des Denkens«. Boole wollte nicht nur einfach eine mathematische Begründung der Logik liefern, er wollte das menschliche Denken selbst erklären. In der Einleitung zu seinem Werk schreibt er:

Der Zweck der folgenden Abhandlung ist es, die fundamentalen Gesetze jener Operationen des Geistes zu untersuchen, die das Denken ausmachen; ihnen in der symbolischen Sprache eines Kalküls Ausdruck zu geben und auf dieser Grundlage die Wissenschaft der Logik zu errichten und ihre Methode auszuarbeiten ... und schließlich aus den verschiedenen Elementen der Wahrheit, die im Laufe dieser Untersuchungen ans Licht gebracht wurden, einige plausible Hinweise auf die Natur und Zusammensetzung des menschlichen Geistes zu gewinnen.

Boole war es nicht vergönnt, diese ehrgeizigen Pläne zu verwirklichen. Schon zehn Jahre nach Veröffentlichung des Buchs starb er viel zu jung, ohne dass die Bedeutung seines Werks erkannt worden wäre.¹³ Doch selbst wenn Booles Wirkungsort nicht die

akademische Provinz von Cork gewesen wäre, eine Maschine zur Automatisierung dieser Träume hätte er wohl auch nicht entwickeln können.

Zwei Jahre vor Booles Tod kam es faszinierenderweise zu einer Begegnung zwischen ihm und der nächsten Figur in unserer Geschichte, Charles Babbage. Sie fand auf der Londoner Weltausstellung 1862 statt, wo die beiden Wegbereiter der KI sicher auch über Babbages Pläne einer denkenden Maschine sprachen. Es ist eine reizvolle Frage, was sie wohl gemeinsam ausgeheckt hätten, wenn Boole nicht kurz darauf gestorben wäre. Charles Babbage war ein Universalgenie: Mathematiker, Philosoph, Erfinder und Ingenieur.¹⁴ Er träumte davon, mechanische Rechenmaschinen zu bauen. Auch wenn er dies nicht verwirklichen konnte, wird er heute von vielen als Vater des *programmierbaren* Computers betrachtet. Die von ihm entworfene Analytical Engine sollte mithilfe von Lochkarten programmiert werden.

Die Verwendung austauschbarer Programme ist eine wesentliche Eigenschaft von Computern. Jedes Smartphone kann mit neuen Apps bestückt werden, darunter sicherlich Programme, die weder Steve Jobs noch sonst einem der Schöpfer des Smartphones in den Sinn gekommen waren. So kann es viele Dinge zugleich sein: Taschenrechner, Notizblock, Gesundheitsmonitor, Navigationsgerät, Kamera, Filmabspielgerät und sogar, man vergisst es manchmal schon, Telefon. In diese Richtung gingen schon Turings Vorstellungen, als er sein allgemeines Modell für Rechenmaschinen entwickelte. Er beschrieb den Computer als eine *universelle* Maschine, die für verschiedene Aufgaben programmiert werden kann. Und Computerprogramme können sich sogar selbst verändern, eine wichtige Voraussetzung, um den Traum von der Künstlichen Intelligenz Wirklichkeit werden zu lassen. Lernen ist eine wesentliche Komponente von Intelligenz. Wenn

ein Computer das Lernen simulieren soll, dann muss er irgendwie sein eigenes Programm modifizieren können. Zum Glück ist es relativ einfach, ein Computerprogramm zu schreiben, das sich selbst verändern kann. Ein Programm besteht schließlich nur aus Daten, und die können bearbeitet und verändert werden, so wie die Zahlen einer Tabelle, die Buchstaben in einem elektronischen Text oder die Farben der Pixel eines Digitalfotos. So können Computer lernen, neue Aufgaben zu erledigen – das heißt ihr Programm ändern, um Dinge zu tun, die ihnen ursprünglich nicht einprogrammiert waren.

Der erste Programmierer

Babbages wichtigste Mitstreiterin war Augusta Ada King, Countess of Lovelace, allgemein bekannt als Ada Lovelace.¹⁵ Sie verfasste eine Reihe von Erläuterungen, in denen sie Babbages Analytical Engine einem breiteren Publikum beschrieb und erläuterte. Dabei entwickelte sie auch das, was allgemein als das erste Computerprogramm gilt. Babbage sah die Aufgabe seiner Maschine vor allem darin, numerische Berechnungen durchzuführen, um astronomische und andere Tabellen zu erstellen. Doch Lovelace war in der Lage, sich Rechenmaschinen vorzustellen, die viel mehr konnten als bloß Zahlen verarbeiten. Babbages Erfindung, schrieb sie, »könnte auch auf andere Dinge als Zahlen anwendbar sein ... die Maschine könnte nach wissenschaftlichen Prinzipien komplizierte Stücke von Musik beliebiger Komplexität und Länge komponieren.«

Mit dieser Idee war sie ihrer Zeit um ein Jahrhundert voraus. Der kühne Gedanke von Lovelace hat sich mit unseren Smartphones

verwirklicht, die außer Zahlen auch Töne, Bilder und viele andere Dinge bearbeiten können. Und zugleich war Ada Lovelace auch eine der ersten Kritikerinnen der Künstlichen Intelligenz, sprach sie doch denkenden Maschinen die Kreativität ab. »Die Analytical Engine besitzt nicht den Ehrgeiz, irgendetwas Neues erschaffen zu wollen«, schrieb sie. »Sie kann alles, was wir ihr auszuführen befehlen. So kann sie analytischen Überlegungen folgen; aber sie besitzt keine Fähigkeit, analytische Beziehungen oder Wahrheiten zu entdecken.«

Ob sich dies tatsächlich so verhält – dass Computer nicht intelligent sein können, wenn sie nicht kreativ sind –, ist intensiv diskutiert worden. Turing ging darauf in seinem bahnbrechenden *Mind*-Artikel ein. Ich werde darauf in Kürze zurückkommen, doch zuvor will ich noch einige Bemerkungen zu dem Einwand von Lovelace machen. Der erste Mensch, der sich Gedanken über die Programmierung von Computern machte und, bereits über hundert Jahre bevor es solche Rechner gab, von Computern träumte, die mehr können als bloß Zahlen zu verarbeiten, war zugleich äußerst skeptisch, ob das Endziel, denkende Maschinen zu bauen, erreicht werden könne. Es ist eben kein simpler Traum, sondern einer, der unsere Position in der Welt tief berührt. Sind wir in irgendeiner Weise etwas Besonderes? Oder sind auch wir nur Maschinen, so wie unsere Computer? Die Antwort auf diese Fragen wird letztendlich Einfluss auf unser Selbstbild haben. Sie gefährdet unsere Stellung im Zentrum der Dinge, so wie die Entdeckung von Kopernikus, dass sich die Erde um die Sonne dreht, oder Darwins Erkenntnis, dass wir von den Affen abstammen. Eine der weniger bekannten Figuren, die im 19. Jahrhundert in unserer Geschichte mitspielt, ist William Stanley Jevons.¹⁶ Er war hauptsächlich Mathematiker und Ökonom. Aber für uns ist hier vor allem sein »logisches Piano« von Interesse, das er 1870 vor-