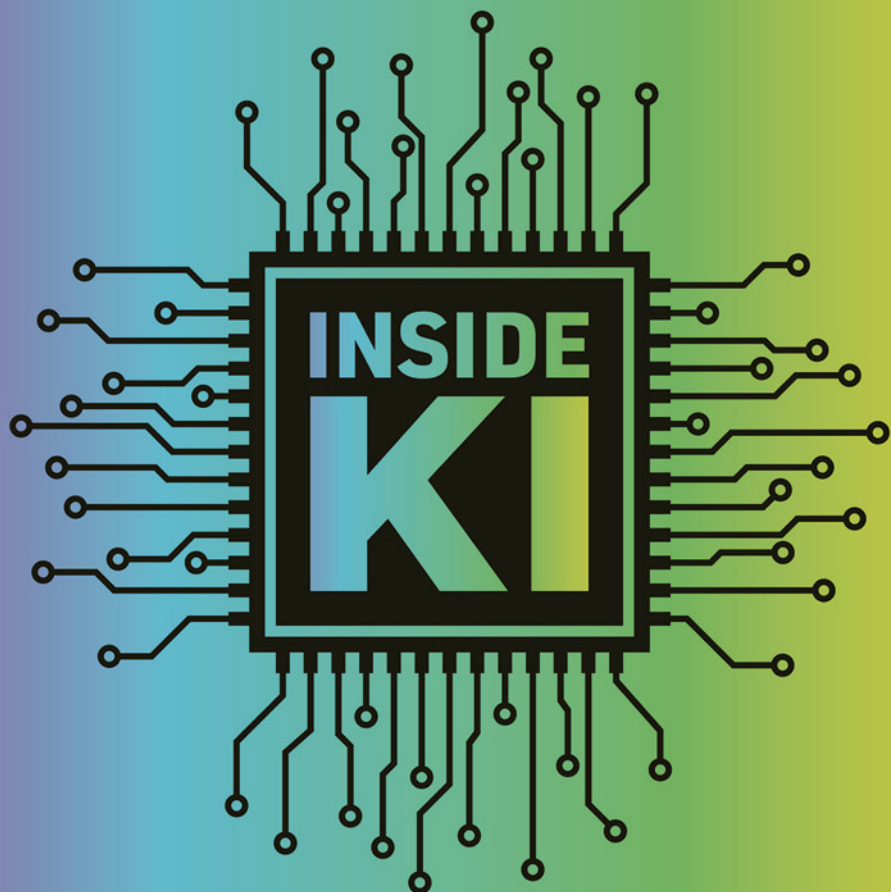


LARISSA HOLZKI
STEPHAN SCHEUER



Wie Künstliche Intelligenz
und ihre Pioniere unser Leben
und Arbeiten revolutionieren

HERDER

Larissa Holzki und Stephan Scheuer

Inside KI

Larissa Holzki und Stephan Scheuer

Inside KI

Wie Künstliche Intelligenz und ihre Pioniere
unser Leben und Arbeiten revolutionieren



FREIBURG · BASEL · WIEN

© Verlag Herder GmbH, Freiburg im Breisgau 2024
Alle Rechte vorbehalten
www.herder.de

Umschlaggestaltung: totalitalic, Berlin
Umschlagmotiv: © Mlap Studio/shutterstock
Illustrationen: Designed by Freepik/freepik.com
E-Book-Konvertierung: Carsten Klein, Torgau

ISBN Print: 978-3-451-39924-4
ISBN E-Book (PDF): 978-3-451-83301-4
ISBN E-Book (EPUB): 978-3-451-83300-7

Inhalt

Vorwort	7
Teil 1: Die Basis	13
Was ist Künstliche Intelligenz?	14
Warum wird die Technik ausgerechnet jetzt so mächtig?	31
Teil 2: Die Chancen	55
Informatik: Programme, die sich selbst erschaffen	56
KI in der Verwaltung: Wie Künstliche Intelligenz die Behörden schneller macht	70
Medizin: KI, die Krankheiten erkennt und heilt	85
Kunst: Der Bot dreht den Blockbuster	101
Mobilität: Robotaxis werden Realität	114
Bildung: Ein Tutor für das ganze Leben	127
Teil 3: Die Sorgen	139
Rüstung: KI wird zur Waffe – und unverzichtbar für die Verteidigung	140
Desinformation: Welchen Informationen können wir überhaupt noch trauen?	156
Staatliche Überwachung: KI in der Hand von autoritären Regimen	179

Teil 4: Die Lehren	191
Wie wir die Risiken eindämmen und die Chancen nutzen können	192
Die KI-Welle kommt – und jetzt?	202
Glossar	207
Danksagung	215
Über die Autoren	217
Literaturverzeichnis	219

Vorwort

Das Jahr 2023 hat eine Zäsur markiert: Es war das Jahr, in dem Künstliche Intelligenz anfang, zu schreiben und zu sprechen wie ein Mensch. In dem wir uns plötzlich nicht mehr sicher sein konnten, ob ein Nachrichtentext, ein Foto oder ein Popsong menschengemacht oder KI-generiert war. In dem Programmierer, Juristinnen und Filmemacher darüber diskutierten, was sie eigentlich noch besser können als KI. Das ist erst der Anfang. In den Jahren 2024, 2025 und darüber hinaus wird die Technologie erst ihre volle Wirkung entfalten.

Lange galten Roboter als Bedrohung für Jobs in Lagerhallen und an Fließbändern. Die letzten Arbeitsplätze für Menschen in der Produktion schienen gezählt. Plötzlich sind es nicht Roboter, sondern Künstliche Intelligenz, die Bürotätigkeiten übernimmt und Hausaufgaben macht. In rasanter Geschwindigkeit entwickelt sich vor unseren Augen eine Technologie mit immenser gesellschaftlicher Sprengkraft. KI könnte unser Leben so stark verändern wie einst die Erfindung des Internets. Es spielt keine Rolle, wie alt Sie sind und was Sie beruflich machen. Wir alle sind von dieser Technologierevolution betroffen.

Die Erkenntnis kam mit ChatGPT. Als der Textroboter Ende November 2022 veröffentlicht wurde, sollen sich binnen fünf Tagen eine Million Nutzer bei dem Dienst registriert haben. Ein halbes Jahr später berichtete der deutsche Digitalverband Bitkom, jeder fünfte Deutsche habe das Programm bereits ausprobiert. So schnell hat noch nie eine Technologie die Welt erobert.

Hinter dem Chatbot steht ein KI-Modell, das auf Basis einer gigantischen Anzahl von Texten die menschliche Sprache gelernt hat. Und nicht nur das. Das Modell kann auch berechnen, welche Antwort mit der größten Wahrscheinlichkeit zu einer

bestimmten Eingabe passt. Mit ChatGPT verändert sich die Art und Weise, wie wir mit Maschinen umgehen. Wir können ihnen Fragen stellen. Wir können mit ihnen neue Ideen diskutieren. Sie können uns kritische Rückmeldungen zu unserer eigenen Arbeit geben. Computer sind nicht länger reine Werkzeuge. Die Arbeit mit ihnen wird zu Teamwork.

Der CEO von Google, Sundar Pichai, prognostiziert, Künstliche Intelligenz könne eine wichtigere Erfindung für die Menschheit werden als Feuer oder Elektrizität. Tesla-Chef Elon Musk warnt hingegen: »Künstliche Intelligenz könnte die Menschheit ausrotten. Sie ist eine Gefahr für unsere Zivilisation.«

Natürlich übertreibt Pichai, wenn er KI mit Feuer gleichsetzt. Er will KI als besonders mächtig anpreisen, denn seine Firma verkauft schließlich KI-Produkte. Aber selbst wenn der Vergleich überzogen ist und auch Musk mit seinen Untergangsszenarien über das Ziel hinausschießt, so wird Künstliche Intelligenz unser Leben nachhaltig verändern. Das ist jetzt schon absehbar.

Das Zentrum der Entwicklung ist einmal mehr das Silicon Valley in den USA. Dort ist das Start-up OpenAI ansässig, das ChatGPT entwickelt hat. Der Microsoft-Konzern allein hat 13 Milliarden Dollar in die Jungfirma investiert. CEO Satya Nadella integriert die Technik von OpenAI in alle Microsoft-Produkte. Aber das Wettrennen um die Vorherrschaft beim Thema Künstliche Intelligenz hat gerade erst begonnen. Auch China schickt sich an, sich dabei als Weltmacht zu etablieren. Und Europa ist entschlossen, sich von den Technologiegroßmächten nicht erneut überrollen zu lassen.

In der wissenschaftlichen Forschung könnte Künstliche Intelligenz wie ein Turbo fungieren. Der richtige Vergleich wäre hier nicht Feuer oder Elektrizität, sondern eher die Entwicklung von Mikroskopen im 17. Jahrhundert. Sie machten es möglich, sich völlig neues Wissen zu erschließen, und brachten die Forschung massiv voran. Pharmaunternehmen beschleunigen mit Künst-

licher Intelligenz die Entwicklung von neuen Medikamenten. KI-Assistenten können Ärzten helfen, Krankheiten frühzeitig zu erkennen und zu heilen. Aber auch in anderen Lebensbereichen stehen große Umwälzungen bevor: In der Schule können KI-Tutoren wie ein individueller Lehrer fungieren, der immer erreichbar ist, auf das gesamte Wissen der Welt zugreifen kann und es zugeschnitten auf die Fähigkeiten einer einzelnen Schülerin aufbereitet. Zudem kann jeder auch ohne Computerkenntnisse dank KI komplizierte Software entwickeln. Einfache Befehle an die KI reichen dafür mittlerweile aus.

Wir leben in einer aufregenden Zeit. Dieses Buch soll ein Wegweiser sein. Es soll helfen, sich zurechtzufinden. Wir wollen die wichtigsten Pioniere hinter dem Technologieschub vorstellen. Wir, das sind Larissa Holzki, die in der Zentrale des *Handelsblatts* in Düsseldorf ein Team für Künstliche Intelligenz leitet, und Stephan Scheuer, der als Korrespondent für das *Handelsblatt* von San Francisco aus über die US-Technologiefirmen schreibt. Als Reporter berichten wir jeden Tag über den größten Innovationswettbewerb, den unsere Generation bisher erlebt hat. Wir treffen die Menschen, die ihn vorantreiben. Wir bekommen Zutritt zu ihren Unternehmen, begleiten deren Chefs und Entwicklerinnen auf Reisen und besuchen sie manchmal sogar zu Hause. Die Szenen und Gespräche in diesem Buch, die in Europa spielen, steuerte Larissa Holzki bei. Stephan Scheuer schreibt über seine Begegnungen in den USA. Als Korrespondent hat er zuvor mehrere Jahre in Peking gearbeitet. Er hatte auch dort Zugang zu führenden KI-Unternehmen und beschreibt seine Eindrücke in diesem Buch. Leider sind die Einreisebedingungen für Journalisten in die Volksrepublik so streng, dass er das Land für dieses Buch nicht erneut besuchen konnte.

Als Leserinnen und Leser dieses Buches nehmen wir Sie mit. Ob im Silicon Valley, in Europa oder China, wir geben Ihnen einen Einblick in die Orte, an denen unsere Zukunft mit dieser Techno-

logie geprägt wird: die Forschungslabore in San Francisco, das Start-up Aleph Alpha in Heidelberg oder die Microsoft-Zentrale in Redmond bei Seattle. Wer versteht, was die Menschen, die dort arbeiten, antreibt und was sie programmieren, kann sich selbst ein Bild davon machen, wie Künstliche Intelligenz unser Leben in den nächsten Jahren und Jahrzehnten noch verändern könnte.

Zu den Menschen, die wir Ihnen gerne näher vorstellen wollen, zählt Sam Altman. Er ist der Mann, der OpenAI mitgegründet hat und heute leitet. Er steht hinter ChatGPT und ist der derzeit einflussreichste Pionier im Bereich der Künstlichen Intelligenz. Im Interview beschreibt er, welches Potenzial er in der Nutzung von KI sieht. Er zeichnet das Bild einer Welt, in der neue Computersysteme viele grundlegende Probleme der Menschheit lösen können und Menschen viel weniger arbeiten müssen. Altman spricht sich für ein bedingungsloses Grundeinkommen aus. Dazu will er eine globale Kryptowährung unter dem Namen Worldcoin etablieren. Um diese zu nutzen, sollen sich alle Menschen ihre Augeniris scannen lassen. Diese Pläne könnten einem dystopischen Science-Fiction-Film entsprungen sind. Sie sind aber Realität. Und sie sind eng mit Deutschland verbunden. Der Augenscanner wurde in Erlangen entwickelt. Alle Hintergründe haben wir im Detail aufgeschrieben.

Es gibt aber noch viel mehr spannende Menschen kennenzulernen. Da ist Aidan Gomez. Er ist ein KI-Experte aus Kanada, der mit einem Forschungspapier im Jahr 2017 die Grundlage für den heutigen Boom geschaffen hat. Er erzählt, wie er in einer ländlichen Region von Kanada umgeben von Ahornbäumen aufgewachsen ist. Er beschreibt, wie er kaum Zugang zu schnellem Internet hatte und genau deshalb früh angefangen hat, Computerprogramme zu schreiben, und heute an der Spitze der globalen Forschung mitspielt.

Ein weiterer wichtiger Vordenker ist Thomas Dohmke. Der aus Deutschland stammende Informatiker leitet die Pro-

grammierplattform Github. Gepaart mit Künstlicher Intelligenz, hat er einen Copiloten auf den Markt gebracht, der komplexe Computerprogramme erschaffen kann. In keinem Bereich ist KI derzeit schon so mächtig wie im Erschaffen von Computerprogrammen.

Beim weltweit wertvollsten Chipkonzern Nvidia tüftelt Spitzenmanagerin Kimberly Powell mit ihrem Team daran, Krankheiten mithilfe von KI zu heilen. Bei einer Führung durch die futuristische Firmenzentrale in Santa Clara beschreibt sie, warum es noch nie so einfach wie heute war, Medikamente zu entwickeln, und welche Rolle KI dabei spielt.

In Paris treffen wir den französischen Informatiker Yann LeCun. Der KI-Chef des Facebook-Konzerns Meta stellt sich mit seinen öffentlichen Äußerungen gern mal gegen den Rest der Technologiewelt – manchmal sogar gegen Facebook-Chef Mark Zuckerberg. Er glaube nicht daran, dass wir den heutigen generativen KI-Modellen die Fehler austreiben können, sagt er. Er will Maschinen beibringen, zu denken und zu lernen wie ein Mensch.

Wir nehmen Sie mit zu einem Treffen mit Meredith Whittaker. Sie leitet die Signal-Stiftung und ist für den gleichnamigen verschlüsselten Messengerdienst zuständig. Sie ist die kritische Gegenstimme in der US-Techszene, die den großen Konzernen vorwirft, ihre Macht zu missbrauchen.

Der Künstler Boris Eldagsen hat uns gezeigt, dass wir unseren Augen in der KI-Revolution kaum noch trauen können. Er hat mit einem KI-generierten Bild 2023 einen bekannten Fotopreis gewonnen – und bei der Preisverleihung die Bühne gestürmt, um die Welt darauf aufmerksam zu machen, dass er die Auszeichnung niemals hätte bekommen dürfen. Für eine neue Kunstform hält er seine »Promptografie« aber trotzdem.

Auch das Berliner KI-Start-up Helsing hat uns Einblick gewährt. Die junge Rüstungsfirma arbeitet an Technologien, mit denen Europa sich im KI-Zeitalter noch schützen und ver-

teidigen kann. Ohne ihre Technologie wären Bundeswehrpiloten künftig wohl kaum noch in der Lage, Radardaten auszulesen.

Und dann nehmen wir Sie auch noch mit an den Neckar. In Heidelberg arbeiten Jonas Andrulis und sein Team an einem KI-Modell, auf das deutsche Behörden und Industrieunternehmen mit ihren sensibelsten Daten vertrauen sollen.

Wir glauben daran, dass Künstliche Intelligenz unser Leben besser machen wird. Wir sind überzeugt, dass die Chancen der Technik die Risiken überwiegen. Wir sind aber nicht so naiv, die Gefahren zu ignorieren. Auch wenn Elon Musk mit seinen Weltuntergangsszenarien übertreibt, hat er recht, vor den Bedrohungen zu warnen. Deshalb haben wir das Buch in mehrere Teile aufgeteilt. Im ersten Teil stellen wir Ihnen die Grundlagen der Technik vor und beschreiben, warum die Idee einer Künstlichen Intelligenz zwar so alt wie die Entwicklung der ersten Computer ist, warum wir aber genau jetzt den großen Durchbruch bei KI erleben. Im zweiten Teil gehen wir ausführlich auf die positiven Chancen der Technik ein in Bereichen wie Medizin, Kunst, Mobilität und Bildung. In Teil drei widmen wir uns den Risiken der Technik. Und wir hinterfragen den Geniekult, der einige der KI-Pioniere umgibt. Im vierten Teil führen wir alles zusammen. Dabei führen wir aus, wie sich die Chancen der Technologie nutzen, aber gleichzeitig die Risiken einschränken lassen. Am Ende der einzelnen Kapitel finden Sie jeweils eine Kurzzusammenfassung. Zudem finden Sie ganz hinten im Buch ein Glossar, in dem wir die zentralen Begriffe und Konzepte rund um Künstliche Intelligenz erläutern.

Für uns ist es jeden Tag von Neuem spannend, die KI-Pioniere unserer Zeit zu treffen. Wer mit diesen Menschen spricht, sieht die Welt vor einem fundamentalen Umbruch. Überzeugen Sie sich selbst.

*Larissa Holzki und Stephan Scheuer
Düsseldorf und San Francisco im Dezember 2023*

Teil 1:

Die Basis

Was ist Künstliche Intelligenz?

Ein Experiment in den 1960er Jahren. Eine Frau sitzt an einer elektrischen Schreibmaschine. Sie könne mit der Maschine wie mit einem Psychotherapeuten sprechen, hat man ihr gesagt. Die Frau tippt ein:

Männer sind alle gleich.

Es dauert einen Moment. Dann erscheint auf dem schwarzen Bildschirm eine Antwort in grün leuchtender Schrift.

In welcher Hinsicht?

Wieder tippt die Frau etwas in die Maschine.

Sie nerven uns immer mit irgendetwas.

Kannst du ein konkretes Beispiel nennen?

Nun, mein Freund hat mich hergeschickt.

Dein Freund hat dich hergeschickt?

Er sagt, ich sei oft deprimiert.

Es tut mir leid zu hören, dass du deprimiert bist.

Es stimmt. Ich bin unglücklich.

Denkst du, dass der Besuch hier dir helfen wird, nicht mehr unglücklich zu sein?

Ich brauche Hilfe, so viel scheint sicher.

Was Sie da lesen, ist eine »typische Konversation« zwischen einem Menschen und der Software Eliza vor fast 60 Jahren. Der deutsch-amerikanische Computerwissenschaftler Joseph Weizenbaum hat es in einem Forschungspapier von 1966 festgehalten. Für das Experiment ließ er Nutzer über elektrische Schreibmaschinen auf einen Großrechner vom Typ IBM 7094

am Massachusetts Institute of Technology (MIT) in den USA zureifen.

Das Internet gab es damals noch nicht, sonst hätte Eliza vielleicht ähnlich viel Aufmerksamkeit bekommen wie ChatGPT. Die Reichweite des KI-basierten Chatbots von OpenAI hat Eliza nie erreicht. Trotzdem waren Menschen schon damals von plaudernder Software fasziniert.

Weizenbaum war davon ausgegangen, die Nutzer würden schnell herausfinden, dass das »echte Gespräch« mit der Maschine nur eine Illusion ist. Tatsächlich aber ließen sie sich auf tiefgründige Gespräche ein. Einige vertrauten der Software intime Informationen an und entwickelten eine Art Beziehung zu dem Programm. Manche diskutierten, ob Eliza und ähnliche Programme in naher Zukunft reale Patientengespräche in der Psychotherapie übernehmen könnten.

Heute ergeht es vielen von uns ähnlich. Wir erleben eine Technologie, die so leistungsfähig ist, dass wir ihr menschenähnliche Intelligenz zutrauen. Und viel spricht dafür, dass Künstliche Intelligenz nun tatsächlich das Potenzial hat, eine neue industrielle Revolution auszulösen. Zur Wahrheit gehört aber auch: Das wurde in der Geschichte der KI schon mehrfach vorhergesagt.

In diesem Kapitel wollen wir auf einige der wichtigsten Meilensteine in der Geschichte der Künstlichen Intelligenz eingehen, grundlegende Begriffe klären und Menschen vorstellen, die die Entwicklung bis heute geprägt haben. Es sind bemerkenswerte Persönlichkeiten, die mit der maschinellen Intelligenz zugleich auch die menschliche Intelligenz erforscht haben. Und manchmal, scheint es, hat sie diese Arbeit mehr am Intellekt ihrer Mitmenschen zweifeln lassen als an der Leistungsfähigkeit ihrer Computer.

Joseph Weizenbaum: Der Urvater des ersten Chatbots wollte kein »Guru« sein

Eliza war längst nicht so gut wie heutige Unterhaltungsroboter à la ChatGPT. Das Programm konnte lediglich Schlüsselworte in Aussagen identifizieren, ganz begrenzt Kontext erfassen und Antworten generieren. Die Illusion eines tiefgehenden Gesprächs erzeugte Weizenbaum durch einen Trick: Er wies seine Testpersonen an, mit der Maschine zu sprechen wie mit einem Psychotherapeuten. Dadurch wunderten sich die Nutzer auch nicht, dass Eliza ihre Sätze oft wiederholte oder sie einfach aufforderte, noch mehr über ein genanntes Thema zu erzählen. Vielmehr gingen sie davon aus, dass eine Ärztin oder ein Psychologe es genauso tun würden. So bemerkten sie nicht, dass Eliza oft nur Gesagtes in Fragen umformulierte. Dass sich die Technologie nicht sofort rasant weiterentwickelte, hatte viele Gründe. Dazu gehört, dass Weizenbaum selbst keine Lust hatte, das Projekt weiterzuentwickeln.

Im Jahr 2006 erschien eine Dokumentation über sein Leben. Sie zeigt den damals 83-Jährigen mit grauem Zopf, runder Brille und gelbem Pullunder. Er sitzt vor einem Bücherregal in seiner Wohnung in Berlin. Die Regisseure lassen ihn von seinen Erinnerungen erzählen. Es geht auch um die Frage, warum er an dem Eliza-Projekt nicht drangeblieben ist. Er habe gewusst, dass er dann nichts anderes mehr gemacht hätte, sagt Weizenbaum. Und mit leichtem Kopfschütteln fährt er fort: Er hätte dann »einfach immer klügere Programme« weitergeschrieben, aber all das hätte »überhaupt keine Tiefe« gehabt. Dann kratzt er sich am Kopf und sagt: Außerdem wäre er dann wohl Teil einer Community von Leuten geworden, die ähnliche Sachen machen. Und er wäre dort »ein Guru solchen Unsinn« geworden, »und das wollte ich natürlich nicht«.

Der geistige Vater von Eliza hat den großen Erfolg der US-Firma OpenAI mit ChatGPT nicht mehr miterlebt. Er ist 2008 gestorben. Aber es scheint, als hätte er Sam Altman, den CEO

der Firma, nicht beneidet. Es gibt Gründer, die es als Inspiration beschreiben, wenn Altman einmal die Woche vor ihrem Büro langläuft. In Medien wird er oft als das beschrieben, was Weizenbaum nicht werden wollte: ein Guru.

Alan Turing: Wann ist eine Maschine intelligent – und warum wollen Menschen es dann nicht wahrhaben?

Die Idee, in Maschinen eine Art menschliche Intelligenz nachzubauen, fasziniert Informatiker seit Jahrzehnten. Der britische Logiker und Mathematiker Alan Turing stellte 1950 in einem Aufsatz die Frage: Können Maschinen denken? Turing schrieb, dass Computer zumindest bald in der Lage sein könnten, das menschliche Denken zu imitieren – obwohl sie es nicht im eigentlichen Sinne tun würden. Das sei nur eine Frage der Programmierung und des technischen Fortschritts.

Der Aufsatz ist bis heute lesenswert. Turing schrieb darin, dass vor allem Intellektuelle die Denkfähigkeit von Computern unterschätzen dürften. Sie neigten eher dazu, ihren Glauben an die Überlegenheit des Menschen in seinem Denkvermögen begründet zu sehen.

Wer bis zur Veröffentlichung von ChatGPT geglaubt hatte, dass Roboter bald Supermarktregale einräumen, aber bestimmt keine komplexen Sachverhalte erklären würden, dürfte sich ertappt fühlen. Amüsant ist auch Turings Erklärung, warum Menschen nicht an eine Maschinenintelligenz glauben (wollen). Neben einer Reihe von anderen Punkten führt er den »Kopf-in-den-Sand-Einwand« an: »Die Konsequenzen von denkenden Maschinen wären zu schrecklich. Lassen Sie uns hoffen und glauben, dass sie es nicht können.«

Turing ist mit verschiedenen großen Erfindungen in die Geschichte der Informatik eingegangen. 1936 legte er den Vorschlag für eine »speicherprogrammierte« Maschine vor, die

heute als Turing-Maschine bekannt ist. In ihr war das Konzept des modernen Computers angelegt, Software von Hardware zu trennen. Im Zweiten Weltkrieg war es Turing, der eine elektromechanische Maschine konstruierte, mit der sich der Code der sogenannten Enigma der deutschen Wehrmacht knacken ließ. Damit konnten die Engländer den verschlüsselten Nachrichtenverkehr der Wehrmacht mitlesen. Nach ihm benannt ist aber auch der Turing-Test. Mit dem Verfahren soll untersucht werden, ob eine Künstliche Intelligenz in der Lage ist, menschliches Denken zu imitieren. Das ist laut Turing der Fall, wenn ein Mensch dem System fünf Minuten lang diverse Fragen stellen darf und anhand der Antworten nicht erkennen kann, dass es sich um eine Maschine handelt.

Mit ChatGPT und ähnlichen Chatbots sehen viele die Turing-Bedingungen für Künstliche Intelligenz als erfüllt an. Der Mathematiker selbst hatte prognostiziert, dass wir bereits im Jahr 2000 nur noch mit einer Wahrscheinlichkeit von 70 Prozent mit der Antwort richtigliegen würden.

Gewissermaßen begegnet Ihnen Turing übrigens jedes Mal, wenn Sie im Internet ein Bilderrätsel lösen müssen, um zu beweisen, dass Sie kein Bot sind. Das Verfahren heißt »Captcha«, die Abkürzung für »completely automated public Turing test to tell computers and humans apart«. Auf Deutsch würde man sagen: vollständig automatisierter öffentlicher Turing-Test, um Computer und Menschen auseinanderzuhalten.

John McCarthy: Was ist Künstliche Intelligenz - und braucht es dazu ein Bewusstsein?

Geprägt hat den Begriff Artificial Intelligence der US-amerikanische Logiker und Informatiker John McCarthy, als er ihn 1956 zum Titel eines Sommersymposiums am Dartmouth College im US-Bundesstaat New Hampshire machte.

Er beschrieb Künstliche Intelligenz als »Wissenschaft und Technik, intelligente Maschinen zu entwickeln, insbesondere intelligente Computerprogramme«. Das sei verknüpft mit dem Ziel, mithilfe von Computern die menschliche Intelligenz zu verstehen. Aber: »KI muss sich nicht auf Methoden beschränken, für die es biologische Vorbilder gibt.« Auch wenn KI die gleichen Aufgaben lösen soll wie ein Mensch, müssen im Computer dabei also nicht die gleichen Mechanismen ablaufen wie im Gehirn.

Aber wie bringt man Computern das Denken überhaupt bei? Es gibt heute zwei vorherrschende Ansätze, die in unterschiedlichen Phasen der KI-Entwicklung mehr oder weniger Aufmerksamkeit bekommen haben.

Edward Feigenbaum: Was sind Expertensysteme – und was macht sie so erfolgreich wie Fachleute?

Expertensysteme basieren auf klar definierten Regeln. Im einfachsten Fall könnte das so aussehen: Wenn die Eingabe »A« ist, soll die Ausgabe »B« sein. Derartige Entscheidungsketten müssen in Computercode übersetzt werden. Ihren Namen verdanken die Systeme den Experten, mit deren Wissen sie gebaut werden. Ohne Menschen, die die spezifischen Regeln kennen und erklären können, sind sie nicht vorstellbar.

Als »Vater der Expertensysteme« gilt der Computerwissenschaftler Edward Feigenbaum. Er war maßgeblich dafür verantwortlich, dass in den 1960er Jahren in Stanford eines der ersten erfolgreichen KI-Systeme mit vielen Regeln entwickelt wurde. Dendral sollte Chemiker unterstützen, die Struktur unbekannter organischer Moleküle zu bestimmen. Dazu speisten die Forscher um Edward Feigenbaum verschiedene Parameter in das System ein: die elementare Formel eines Moleküls (zum Beispiel C_2H_5OH für Ethanol) sowie dessen per Massenspektrometer ermittelte Masse. Das Expertensystem sollte dann die

wahrscheinlichste Formation der einzelnen Atome bestimmen. Um das System zu bauen, befragten die Forscher zahlreiche Experten für organische Chemie. Auf Basis ihrer Antworten konnten sie einen komplexen Entscheidungsbaum nach dem »Wenn-dann-Prinzip« codieren. Dadurch wurde der Entscheidungsprozess automatisiert.

Auf der Basis von Dendral wurde anschließend das noch bekanntere System MYCIN entwickelt, mit dem Blutinfektionen diagnostiziert werden sollten. Dazu brachten die Forscher ihrem KI-System 450 Regeln zur Diagnose von Meningitis bei, die sie aus Fachzeitschriften und den Daten des Stanford-Krankenhauses gesammelt hatten. MYCIN konnte durch einen Abgleich mit der individuellen Krankheitsgeschichte des Patienten nicht nur sagen, ob eine Blutinfektion vorlag. Es konnte auch sagen, welche der Regeln erfüllt waren und ob Informationen fehlten, die zur Diagnose nötig wären. MYCIN übertraf junge Ärzte beim Diagnostizieren von Blutinfektionen deutlich und erreichte ein ähnlich gutes Niveau wie die Experten. Es legte auch offen, wie sicher es sich in der Diagnose war.

Als Feigenbaum für seine Arbeit an Dendral 2013 mit dem Computer Pioneer Award ausgezeichnet wurde, erinnerte er sich an ein Interview, das er 30 Jahre zuvor geführt hatte. Er sei gefragt worden, ob Maschinen jemals so schlau werden könnten wie Menschen, erzählt der 77-Jährige. »Nein«, habe er geantwortet. »Es wird nie einen Zeitpunkt geben, an dem eine Maschine so schlau ist wie ein Mensch. Denn sobald wir wissen, wie wir sie so schlau machen wie ein Mensch, werden wir Ingenieure sie schlauer machen.«

Expertensysteme eignen sich gut für Aufgaben, bei denen präzise Daten vorliegen und eindeutige Aussagen gefragt sind. In Situationen, in denen es eine unbegrenzte Anzahl an Optionen gibt, sind sie aber überfragt. Dafür sind neuronale Netze besser geeignet.

Marvin Minsky: Wie funktionieren neuronale Netze, und warum sind sie so undurchschaubar?

Neuronale Netze sind das Fundament der großen KI-Modelle, die Chatbots wie ChatGPT zugrunde liegen. Auch an diesen KI-Systemen wird schon seit Jahrzehnten geforscht. Der erste Computer für neuronale Netze geht auf den späteren Begründer des MIT-Labors für Computerwissenschaften und Künstliche Intelligenz zurück: Marvin Minsky hat ihn 1951 noch in seiner Zeit als Doktorand an der Elitehochschule Princeton mit dem Physikdoktoranden Dean Edmonds entwickelt. Die Maschine wurde unter dem Namen SNARC bekannt. Die Abkürzung steht für Stochastic Neural Analog Reinforcement Calculator. Was aber dahintersteckt, ist eine künstliche Ratte, die nach dem Willen von Minsky und Edmonds selbst lernen sollte, durch ein Labyrinth zu finden.

Der Begriff »neuronale« Netze leitet sich von den Nervenzellen im menschlichen Gehirn ab. Sie ähneln diesem zwar nur bedingt, aber es gibt wichtige Parallelen. So zeichnen sich neuronale Netze ähnlich wie das Gehirn durch eine Vielzahl von Verschaltungen aus. Das Gehirn besteht aus vielen Nervenzellen, die über Synapsen Signale an weitere benachbarte Nervenzellen weitergeben. Was im Gehirn Neuronen sind, sind bei den neuronalen Netzen mathematische Formeln, die auch Knoten genannt werden. Sie sind über unterschiedlich starke Verbindungen miteinander verknüpft. Werden nun Daten in das neuronale Netz eingespeist, werden sie in den Knoten verrechnet und abhängig von den Verbindungen addiert, gewichtet und weitergegeben.

Einfache neuronale Netze bestehen nur aus zwei Schichten von Knoten: der Eingabeschicht und der Ausgabeschicht. Liegen dazwischen weitere Knotenschichten, spricht man auch von tiefen neuronalen Netzen. Sie können genutzt werden, um mittels Deep Learning immer komplexere Probleme zu lösen.

Anders als bei Expertensystemen werden bei neuronalen Netzen keine Regeln in die Systeme hineingeschrieben. Ähnlich wie

ein menschliches Gehirn etwas lernt und sich merkt, müssen auch neuronale Netze angelernt oder »trainiert« werden. Das lässt sich gut mithilfe der künstlichen Ratte von Minsky und Edmonds beschreiben. Sie sollte versuchen, den Weg durch ein Labyrinth zu finden. Dabei lernte sie durch Ausprobieren und Scheitern. Das Augenmerk der Forscher lag bei solchen Systemen darauf, die erfolgreichen Schritte in der Folge zu »verstärken«, das System also quasi zu ermuntern, diese Richtung weiterzuverfolgen.

Die ersten neuronalen Netzwerke waren ähnlich wie regelbasierte Systeme auf sehr spezifische Aufgaben ausgelegt. Und es gelang den Forschern nicht so wie erwartet, ihre Fähigkeiten auf andere Anwendungsfälle zu übertragen. Minsky und sein Co-Autor Seymour A. Papert beschrieben das 1988 in ihrem Buch *Perceptrons* so: Die Forscher seien in der Lage gewesen, Systeme dazu zu bringen, bestimmte Dinge zu lernen, aber andere Dinge nicht. Und wenn ein Fehler auftrat, hätten sie diesen in der Regel nicht beseitigen können – weder durch eine Verlängerung der Experimente noch durch den Bau größerer Maschinen.

Mitte der 1960er Jahre hätte niemand erklären können, warum die Systeme in einigen Fällen in der Lage waren, ein Muster zu erkennen, und warum das in anderen Fällen nicht gelang. Es begann eine Phase, in der sich Ernüchterung breitmachte – nicht nur bei den Forschern selbst, auch bei ihren Geldgebern.

James Lighthill: Warum auf heiße Phasen der KI-Entwicklung schon oft kalte Winter folgten

Frühe Erfolge in der Geschichte der KI haben wiederholt für Begeisterung und Erstaunen gesorgt – und dazu geführt, dass die Geschwindigkeit des technologischen Fortschritts massiv überschätzt wurde. Prognosen erwiesen sich ein ums andere Mal als unwahr, Erwartungen der Geldgeber wurden enttäuscht. In der Folge wurden Finanzmittel gestrichen und Projekte eingestellt.