KAI-FU LEE

KI

ZEHN ZUKUNFTSVISIONEN

2041

QIUFAN CHEN

campus

KI 2041

KAI-FU LEE ist CEO der Risikokapitalgesellschaft Sinovation Ventures, Co-Vorsitzender des Artificial Intelligence Councils beim Weltwirtschaftsforum und Autor des Bestsellers *AI-Superpowers* (2019, Campus Verlag). Er hat einen Doktortitel der Carnegie Mellon University und hatte Führungsrollen bei Microsoft, SGI und Apple inne. Lee ist ehemaliger Präsident von Google China und stand bereits auf der Time-100-Liste. Er lebt in Beijing.

QIUFAN CHEN ist preisgekrönter Autor, Übersetzer, kreativer Produzent, Kurator und Präsident der World Chinese Science Fiction Association. Sein Roman *Die Siliziuminsel* erschien 2019 in deutscher Übersetzung (Heyne). Chen ist Gründer und COO des Content-Produktionsstudios Thema Mundi und lebt in Beijing und Shanghai.

KAI-FU LEE
QIUFAN CHEN

KI 2041

ZEHN ZUKUNFTSVISIONEN

Aus dem Englischen
von Thorsten Schmidt

CAMPUS VERLAG
FRANKFURT/NEW YORK

Copyright © 2021 by Kai-Fu Lee und Qiufan Chen
All rights reserved.
This translation published by arrangement with Currency, an imprint
of Random House, a division of Penguin Random House LLC.

ISBN 978-3-593-51549-6 Print
ISBN 978-3-593-45026-1 E-Book (PDF)
ISBN 978-3-593-45027-8 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 © 2022. Alle deutschsprachigen Rechte bei Campus Verlag GmbH, Frankfurt am Main.
Umschlaggestaltung: total italic, Thierry Wijnberg, Amsterdam/Berlin
nach einem Design von Will Staehle
Redaktion: Anne Büntig
Satz: Oliver Schmitt, Mainz
Gesetzt aus: Minion, Barlow und Aglet Mono
Druck und Bindung: Beltz Grafische Betriebe GmbH, Bad Langensalza
Beltz Grafische Betriebe ist ein klimaneutrales Unternehmen
(ID 15985-2104-1001).
Printed in Germany

www.campus.de

INHALT

EINLEITUNG VON KAI-FU LEE
Die wahre Geschichte der KI 9

EINLEITUNG VON CHEN QIUFAN
Ein optimistischer Blick in die Zukunft 19

KAPITEL 1
Der goldene Elefant 25

Analyse: Deep Learning / Big Data / Internet- und Finanzanwendungen / KI-Externalitäten 47

KAPITEL 2
Götter hinter den Masken 61

Analyse: Computer Vision / Konvolutionsnetze / Deepfakes / Generative Adversarial Networks / Biometrik / KI-Sicherheit 85

KAPITEL 3
Ungleiche Zwillinge 99

Analyse: linguistische Datenverarbeitung / selbstüberwachtes Training / GPT-3 / Künstliche allgemeine Intelligenz und KI-Erziehung 143

KAPITEL 4
Berührungslose Liebe 159

Analyse: KI-basierte Gesundheitsversorgung / Alphafold / Roboteranwendungen / KI-basierte COVID-Bekämpfung 193

KAPITEL 5
Das gespenstische Idol 211

Analyse: virtuelle Realität (VR) / erweiterte Realität und vermischte Realität / Gehirn-Computer-Schnittstelle / ethische und gesellschaftliche Fragen 247

KAPITEL 6
Der heilige Fahrer 261

Analyse: selbstfahrende Fahrzeuge / vollständige Autonomie und intelligente Städte / ethische und soziale Fragen 301

KAPITEL 7
Quantengenozid 313

Analyse: Quantencomputer / Bitcoin-Sicherheit / autonome Waffen und existenzielle Bedrohungen 367

KAPITEL 8
Der Jobretter 381

Analyse: KI-bedingter Stellenabbau / bedingungsloses Grundeinkommen / was KI nicht kann / fit für die Arbeitswelt der Zukunft 417

KAPITEL 9
Die Insel der Glückseligkeit 433

Analyse: KI und Glück / Datenschutz-Grundverordnung / personenbezogene Daten / Privacy Computing mit föderalem Lernen und vertrauenswürdiger Laufzeitumgebung 469

KAPITEL 10
Der Traum vom Überfluss 485

Analyse: Überfluss / neue Wirtschaftsmodelle / die Zukunft des Geldes / Singularität 509

Danksagung 529

Register 531

Was wir wollen, ist eine Maschine, die aus Erfahrungen lernen kann.

– Alan Turing

Jede hinreichend fortgeschrittene Technologie ist nicht von Magie zu unterscheiden.

– Arthur C. Clarke

EINLEITUNG VON KAI-FU LEE
DIE WAHRE GESCHICHTE DER KI

```
Künstliche Intelligenz (KI) ist intelligente
Soft- und Hardware, die Aufgaben ausführen
kann, die normalerweise menschliche Intelligenz
erfordert. KI gibt uns Aufschluss darüber,
wie wir lernen, quantifiziert menschliche
Denkprozesse, erklärt menschliches Verhalten und
lässt uns verstehen, was Intelligenz ermöglicht.
Es ist der letzte Schritt der Menschheit auf
ihrem Weg, sich selbst zu erkennen. Ich hoffe,
einen Beitrag zu dieser neuen, verheißungsvollen
Wissenschaft leisten zu können.
```

Ich schrieb diese Worte vor fast 40 Jahren als naiver Student, um mich für das PhD-Programm der Carnegie Mellon University zu bewerben. Geprägt wurde der Begriff »Künstliche Intelligenz« noch weit früher, von dem Informatiker John McCarthy beim legendären Dartmouth Summer Research Project on Artificial Intelligence im Sommer 1956. Vielen Menschen erscheint KI als *die* Technologie des 21. Jahrhunderts schlechthin, aber einige von uns haben schon vor Jahrzehnten darüber nachgedacht. In den ersten 35 Jahren meiner KI-Karriere war Künstliche Intelligenz als Forschungsgebiet weitgehend auf Universitäten beschränkt, es gab nur wenige erfolgreiche kommerzielle Anwendungen.

Noch bis vor kurzem wurden KI-Forschungsergebnisse nur langsam in praktische Anwendungen umgesetzt. Doch in den letzten fünf Jahren ist KI zur gefragtesten Technologie der Welt geworden. Der bedeutende

Meilenstein war im Jahr 2016, als AlphaGo, eine von DeepMind-Ingenieuren entwickelte Maschine, bei einem Go-Turnier über fünf Spiele, dem sogenannten Google DeepMind Challenge Match, den Lee Sedol schlug. Go ist ein Brettspiel, das 1 Million Billion Billion Billion Billion Mal komplexer ist als Schach. Nach Ansicht von Millionen begeisterter Fans des Spiels erfordert Go zudem, anders als Schach, wahre Intelligenz, Weisheit und eine Zen-artige intellektuelle Verfeinerung. Es schockierte viele, dass das KI-Programm den menschlichen Meister besiegte.

Wie die meisten bahnbrechenden kommerziellen KI-Anwendungen basiert auch AlphaGo auf Deep Learning, einer Technologie, die sich mithilfe großer Datenmengen selbst Dinge beibringt. Deep Learning wurde vor vielen Jahren erfunden, aber erst seit kurzem verfügen wir über ausreichend Rechenkapazität, um seine Leistungsfähigkeit aufzuzeigen und ausreichend Trainingsdaten, um außergewöhnliche Ergebnisse zu erzielen. Verglichen mit meinem Kaltstart im Bereich KI vor 40 Jahren steht uns heute etwa 1 Billion Mal mehr Rechenleistung für KI-Experimente zur Verfügung, und die Speicherung der notwendigen Daten ist 15 Millionen Mal billiger. Die Anwendungen von Deep Learning – und der damit verbundenen KI-Technologien – werden beinahe jeden Aspekt unseres Lebens berühren.

KI steht jetzt an einem Wendepunkt. Sie hat den Elfenbeinturm der Wissenschaft verlassen, und die Tage langsamer Fortschritte sind vorüber.

Allein in den vergangenen fünf Jahren hat KI menschliche Champions in Go, Poker und dem Videospiel Dota 2 geschlagen, und sie ist so leistungsfähig geworden, dass sie in nur 4 Stunden Schach lernt und von keinem menschlichen Spieler mehr besiegt werden kann. Aber es geht nicht nur um Spiele. Im Jahr 2020 hat KI ein kniffliges Problem gelöst, an dem sich Biologen 50 Jahre lang die Zähne ausgebissen haben: die sogenannte Proteinfaltung. Die Technologie ist Menschen auch in der Sprach- und Objekterkennung überlegen, sie hat »digitale Menschen« mit einem geradezu unheimlich realistischen Aussehen und Sprachvermögen erschaffen, und sie hat nicht nur Hochschulzulassungs-, sondern auch ärztliche Approbationsprüfungen bestanden. Außerdem

übertrifft KI menschliche Richter bezüglich der Fairness und Konsistenz von Strafurteilen und Radiologen bei der Diagnose von Lungenkrebs, und sie steuert Drohnen, die die Zukunft der Warenzustellung, der Landwirtschaft und der Kriegsführung revolutionieren werden. Außerdem ermöglicht KI überhaupt erst autonome Kraftfahrzeuge, die auf Schnellstraßen sicherer fahren als Menschen.

Während KI weitere Fortschritte macht und ständig neue Anwendungen für sie erschlossen werden, stellt sich die Frage, wo dies alles hinführt.

In meinem Buch *AI Superpowers: China, Silicon Valley und die neue Weltordnung*, das 2019 in deutscher Sprache erschien, bin ich auf die starke Zunahme von Daten eingegangen, das »neue Öl«, das die KI antreibt. Die Vereinigten Staaten und China führen die KI-Revolution an, wobei die Vereinigten Staaten bei der Forschung führend sind, während China sich zügiger große Datenmengen zunutze macht, um Anwendungen für seine große Bevölkerung einzuführen. In *AI Superpowers* habe ich neue Fortschritte vorhergesagt, von Entscheidungsfindung auf der Grundlage großer Datenmengen über maschinelle Wahrnehmung bis hin zu autonomen Robotern und Fahrzeugen. Ich rechne damit, dass neue KI-Anwendungen in der Digitalbranche, im Finanzsektor, im Einzelhandel und im Verkehrswesen beispiellosen ökonomischen Wert schaffen, aber auch Probleme mit sich bringen werden, die mit dem Verlust von Arbeitsplätzen und anderen negativen Begleiterscheinungen zusammenhängen. KI ist eine vielseitig einsetzbare Technologie, die praktisch alle Branchen durchdringen wird. Ihre Auswirkungen haben sich in bislang vier Innovationswellen niedergeschlagen, beginnend mit Internet-Anwendungen, gefolgt von wirtschaftlichen Anwendungen (zum Beispiel Finanzdienstleistungen), Wahrnehmungsanwendungen (etwa in Smart Cities) und autonomen Anwendungen, etwa bei Kraftfahrzeugen.

Wenn Sie dieses Buch Ende 2021 oder später lesen, werden die Vorhersagen, die ich in *AI Superpowers* gemacht habe, größtenteils Wirklichkeit geworden sein. Wir müssen uns jetzt nach neuen Herausforderungen umsehen. Bei meinen weltweiten Vorträgen zum Thema KI werde ich immer wieder gefragt: »Was kommt als Nächstes? Was wird

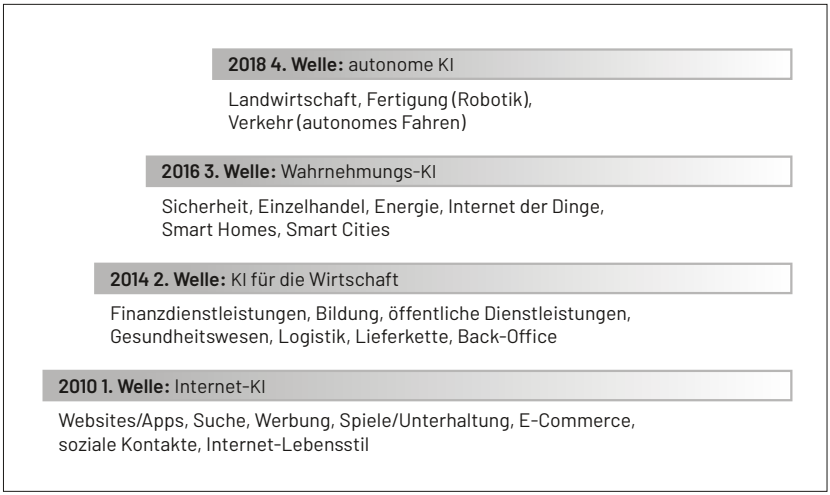

Abbildung 1: Vier Wellen von KI-Anwendungen als »Disruptoren« praktisch sämtlicher Wirtschaftszweige

in fünf, zehn oder zwanzig Jahren geschehen? Was hält die Zukunft für uns Menschen bereit?«

Dies sind wichtige Fragen für unseren Moment in der Geschichte, und jeder, der im Bereich Technologie arbeitet, hat eine Meinung dazu. Einige sind der Überzeugung, wir befänden uns mitten in einer »KI-Blase«, die über kurz oder lang platzen oder zumindest erschlaffen wird. Diejenigen, die extremeren und dystopischen Ansichten zuneigen, glauben alles Mögliche, angefangen von der Vorstellung, dass KI-Giganten »unser Gehirn kapern« und eine utopische neue Rasse »menschlicher Cyborgs« erschaffen werden, bis hin zu einer KI-getriebenen Apokalypse. Diese verschiedenen Vorhersagen mögen echter Neugier oder verständlicher Furcht entspringen, aber sie sind für gewöhnlich spekulativ oder übertrieben. Sie sehen nicht das ganze Bild.

Es gibt die unterschiedlichsten Spekulationen, weil KI komplex und undurchsichtig zu sein scheint. Mir ist aufgefallen, dass sich Menschen oft aus drei Quellen darüber informieren: aus Science-Fiction-Literatur, den Nachrichten und einschlägigen Äußerungen einflussreicher Personen. In Science-Fiction-Büchern und Fernsehsendungen werden

Roboter dargestellt, die Menschen kontrollieren oder überlisten wollen, und böse gewordene Superintelligenzen. In Medienberichten, die sich auf negative, extreme Beispiele und nicht auf die kleinen täglichen Fortschritte konzentrieren, lesen Menschen von autonomen Kraftfahrzeugen, die Fußgänger totfahren, von Technologieunternehmen, die mithilfe von KI Wahlen beeinflussen, und von Menschen, die KI dazu benutzen, Falschinformationen und Deepfakes zu verbreiten. Eigentlich sollten »Vordenker« die beste Informationsquelle sein, aber leider sind die meisten, die diesen Titel für sich beanspruchen, Experten für Wirtschaft, Physik oder Politik, nicht für KI-Technologien. Ihren Vorhersagen mangelt es oft an wissenschaftlicher Genauigkeit. Verschlimmert wird alles noch dadurch, dass Journalisten diese Meinungsmacher oftmals aus dem Zusammenhang gerissen zitieren, um auf möglichst große Resonanz zu stoßen. Daher ist es nicht weiter verwunderlich, dass die öffentliche Meinung über KI – die auf Halbwahrheiten beruht – mittlerweile von Bedenken oder sogar offener Ablehnung geprägt ist.

Selbstverständlich sollten gewisse Aspekte der KI-Entwicklung sorgfältig überprüft werden, aber es ist wichtig, bei all den Befürchtungen nicht das große Ganze und das Potenzial dieser unglaublich bedeutenden Technologie außer Acht zu lassen. Wie die meisten Technologien ist KI an sich weder gut noch schlecht. Und wie bei den meisten Technologien werden auch bei KI langfristig die positiven gesellschaftlichen Auswirkungen stärker zu Buche schlagen als die negativen. Denken Sie nur an den enormen Nutzen von Elektrizität, Mobiltelefonen und Internet. Im Lauf der menschlichen Geschichte haben wir uns oft vor neuen Technologien gefürchtet, die den Status quo bedroht haben. Im Lauf der Zeit verschwinden diese Befürchtungen in der Regel, und die entsprechenden Technologien werden Teil unseres Alltags und verbessern unseren Lebensstandard.

Ich glaube, dass es viele spannende Anwendungen und Szenarien gibt, wo KI unsere Gesellschaft grundlegend verbessern kann. Erstens wird KI, global gesehen, eine enorme Wertschöpfung generieren – PriceWaterhouseCoopers beziffert diese auf 15,7 Billionen Dollar bis 2030 –, was dazu beitragen wird, Hunger und Armut zu bekämpfen. KI wird überdies effiziente Dienstleistungen erbringen, die uns unsere

wertvollste Ressource zurückgeben werden: Zeit. Sie wird Routineaufgaben übernehmen und uns die Freiheit geben, uns anregenderen oder herausfordernderen Tätigkeiten zu widmen. Zu guter Letzt werden Menschen symbiotisch mit KI zusammenarbeiten; KI wird dabei quantitative Analysen, Optimierungen und Routinearbeiten erledigen, während wir Menschen Kreativität, kritisches Denken und Leidenschaft beisteuern. Die Produktivität jedes Menschen wird sich erhöhen, sodass wir unser ganzes Potenzial ausschöpfen können. Die bedeutenden Beiträge, die KI zum Wohl der Menschheit leisten dürfte, müssen genauso gründlich analysiert werden wie ihre Risiken.

Angesichts der zahllosen negativen Berichte über KI halte ich es für wichtig, auch diese anderen Geschichten zu erzählen und die Frage »Was geschieht als Nächstes?« zu beantworten. Deshalb beschloss ich, ein weiteres Buch über KI zu schreiben. Dieses Mal wollte ich den Horizont etwas erweitern und mir den Zustand der Welt und unserer Gesellschaft in 20 Jahren, also 2041, ausmalen. Mir geht es darum, die tatsächliche Geschichte von KI zu diesem zukünftigen Zeitpunkt auf eine unvoreingenommene und ausgewogene, aber auch konstruktive und hoffnungsvolle Weise zu erzählen. Dieses Buch basiert also auf *realistischer KI* beziehungsweise auf Technologien, die es entweder schon gibt oder von denen man vernünftigerweise annehmen kann, dass sie innerhalb der nächsten 20 Jahre anwendungsreif sein werden. Die hier versammelten Erzählungen zeichnen ein Porträt unserer Welt im Jahr 2041, basierend auf Technologien, die mit einer Wahrscheinlichkeit von über 80 Prozent in diesem Zeitraum am Markt eingeführt werden. Ich mag einige über- oder unterschätzen, aber ich bin fest davon überzeugt, dass die in diesem Buch geschilderten Szenarien plausibel sind.

Wieso bin ich mir so sicher? Im Verlauf der letzten 40 Jahre war ich an KI-Forschungsprojekten und -Produktentwicklungen bei Apple, Microsoft und Google beteiligt und managte Investments im Technologiesektor in Höhe von 3 Milliarden Dollar. Dementsprechend habe ich praktische Erfahrungen mit der Zeit und den Prozessen, die erforderlich sind, um eine Technologie von ihrer Beschreibung in einem Fachaufsatz zu einem Massenprodukt zu machen. Als Berater diverser Regierungen auf dem Feld der KI-Strategie kenne ich zudem den

politischen und regulatorischen Ordnungsrahmen und die dahinterstehenden Überlegungen. Ich mache keine spekulativen Vorhersagen über bahnbrechende Neuerungen und beschränke mich überwiegend darauf, wahrscheinliche oder potenzielle zukünftige Nutzanwendungen bestehender Technologien zu beschreiben. Da KI bislang in weniger als 10 Prozent der Wirtschaftszweige Fuß gefasst hat, gibt es viele Möglichkeiten, wie diese Sektoren mit KI-Anwendungen in Zukunft aussehen könnten. Kurzum, ich bin fest davon überzeugt, dass KI auch ohne beziehungsweise mit nur wenigen bahnbrechenden Neuerungen tiefgreifende Auswirkungen auf unsere Gesellschaft haben wird.

AI Superpowers wurde meines Wissens nach von seiner Leserschaft unter anderem deshalb so positiv aufgenommen, weil es auch für Menschen ohne KI-Vorkenntnisse verständlich war. Als ich dieses neue Buch in Angriff nahm, fragte ich mich daher: Was kann ich tun, um Geschichten über KI in einer Weise zu erzählen, die noch mehr Menschen anspricht? Die Antwort lautete: Ich müsste mit jemandem zusammenarbeiten, der gut erzählen kann! Ich beschloss daher, meinen früheren Google-Kollegen Chen Qiufan zu kontaktieren. Nach meinem Ausscheiden bei Google gründete ich eine Wagniskapitalfirma. Qiufan tat etwas Gewagteres – er wurde zu einem preisgekrönten Science-Fiction-Autor. Ich war hocherfreut, als sich Qiufan bereit erklärte, mit mir an dem Projekt zu arbeiten und seine Kreativität mit meiner Einschätzung dessen zu verbinden, was KI-Technologien in 20 Jahren zu leisten vermögen. Wir waren beide der Meinung, dass es faszinierend wäre, sich auszumalen, welche Technologien innerhalb der nächsten 20 Jahre wohl in unser Leben Einzug halten werden, und sie in Erzählungen einzubetten. Wir müssten dabei nicht auf Teleportation oder Aliens zurückgreifen, um unsere Leserinnen und Leser zu fesseln.

Qiufan und ich trafen eine ungewöhnliche Absprache. Ich würde zunächst eine »Technologiekarte« entwerfen, die vorhersagen sollte, wann gewisse Technologien die Marktreife erreichen würden, wie lange es dauern sollte, Daten zu sammeln und KI zu optimieren, und wie leicht es wäre, in verschiedenen Branchen ein KI-Produkt zu etablieren. Ich berücksichtigte auch mögliche externe Faktoren – Herausforderungen, Regulierungen und andere erschwerende Faktoren sowie erzählens-

werte Konflikte und Dilemmata, die womöglich in Verbindung mit diesen Technologien auftreten werden. Mit meinem Input hinsichtlich der technologischen Komponenten hat Qiufan dann seine Talente ausgespielt – sich die Figuren, Schauplätze und Handlungsstränge ausgedacht, die diese Themen zum Leben erwecken würden. Wir bemühten uns darum, jede Geschichte spannend, provokant und technologisch stimmig zu machen. Im Anschluss an jede Erzählung analysiere ich die darin behandelte KI-Technologie und beschreibe ihre Auswirkungen auf das menschliche Leben und die Gesellschaft. Wir haben die Geschichten so ausgewählt, dass sie sämtliche Schlüsselaspekte der KI abdecken, und wir haben sie grob so geordnet, dass wir von den einfachen zu den fortgeschrittenen Technologien übergehen. Wir hoffen, dass die Summe dieser Teile eine fesselnde und verständliche Einführung in KI ist.

Wir haben unser Buch *KI 2041* genannt, weil dann seit der (englischsprachigen) Erstveröffentlichung dieses Buches 20 Jahre vergangen sein werden. Aber es ist unserer Aufmerksamkeit nicht entgangen, dass die Ziffern »41« ein wenig wie »AI« aussehen.

Viele unserer Leserinnen und Leser mögen vielleicht packende Science-Fiction-Erzählungen, aber manche haben womöglich seit ihrem Studium nicht mehr zu einem Roman oder einer Sammlung von Kurzgeschichten gegriffen. Das macht nichts. Betrachten Sie *KI 2041* dann nicht als Science-Fiction-Literatur, sondern als »Scientific Fiction« (wissenschaftliche Fiktion). Die Geschichten spielen an weit verstreuten Schauplätzen rund um den Globus. In einigen erkennen Sie vielleicht eine Welt wieder, die sich nicht allzu sehr von Ihrer eigenen unterscheidet – mit altvertrauten Bräuchen und Gewohnheiten, wenn auch sozusagen mit einer KI-Einfärbung. In anderen hat KI das menschliche Leben bereits tiefgreifend verändert. Sowohl KI-Anhänger als auch KI-Skeptiker finden hier jede Menge Stoff zum Nachdenken. Ein Buch mit einer erheblichen fiktionalen Komponente ist per se ein größeres Wagnis als ein Sachbuch, das einfach die Gegenwart beschreibt und Fragen über die Zukunft stellt. Qiufan und ich wollten mit unseren Erzählungen wagemutig sein, und wir glauben, dass die nachfolgenden Geschichten bei jedem aufgeschlossenen Leser einen Nerv treffen wer-

den, dessen Vorstellungskraft groß genug ist, um darüber nachzudenken, was die Zukunft bereithält.

Die ersten sieben Erzählungen befassen sich mit KI-Anwendungen zunehmender technologischer Komplexität in verschiedenen Branchen sowie mit deren ethischen und gesellschaftlichen Folgen. Die letzten drei Erzählungen (sowie Kapitel 6, »Der heilige Fahrer«) konzentrieren sich mehr auf soziale und geopolitische Probleme, die KI aufwirft, wie etwa den Verlust traditioneller Arbeitsplätze, ein nie dagewesenes Angebot an Gütern, eine sich verschärfende Ungleichheit, ein Wettrüsten bei autonomen Waffensystemen, Konflikte zwischen Datenschutz und Lebenszufriedenheit und das Streben des Menschen nach Sinnerfüllung. Dies sind tiefgreifende Veränderungen, und die Menschen mögen ihnen mit Verständnis begegnen, sie mit Boshaftigkeit ausnutzen, sich resigniert in sie fügen oder sich durch sie dazu inspirieren lassen, sich selbst neu zu erfinden. In den letzten vier Geschichten wollten wir vier mögliche Variationen und abweichende Entwicklungspfade aufzeigen, um zu unterstreichen, dass die Zukunft noch offen ist.

Wir hoffen, dass diese Erzählungen Sie unterhalten und zugleich Ihr Verständnis der KI und der damit verbundenen Herausforderungen vertiefen werden. Wir hoffen auch, dass unsere Vorhersagen für die kommenden Jahrzehnte Ihnen dabei helfen werden, Chancen zu ergreifen und sich den Herausforderungen zu stellen, die die Zukunft mit sich bringen wird. Vor allem aber hoffen wir, dass die in *KI 2041* versammelten Erzählungen Sie in dem Glauben bestärken werden, dass es an uns selbst liegt – dass wir unser Schicksal selbst in der Hand haben. Keine technologische Revolution wird jemals etwas daran ändern.

Lassen Sie uns aufbrechen zu einer Reise ins Jahr 2041.

EINLEITUNG VON CHEN QIUFAN

EIN OPTIMISTISCHER BLICK IN DIE ZUKUNFT

Als ich im August 2019 das Barbican Centre in London besuchte, war dort gerade eine Ausstellung mit dem Titel »AI: More Than Human« (KI: mehr als menschlich). Wie ein erfrischender sommerlicher Regenguss schärfte sie meine Sinne – und räumte mit den meisten meiner Vorurteile und irrigen Vorstellungen über KI auf. Der trügerisch einfache Name der Ausstellung wurde ihrer Vielfalt und Komplexität nicht einmal annähernd gerecht. Jeder Ausstellungsraum enthüllte neue Wunder, die alle irgendwie mit der weiten KI-Definition der Kuratoren in Verbindung standen. Da war Golem, eine sagenhafte Figur des jüdischen Volksglaubens; Doraemon, der beliebte japanische Anime-Held; Charles Babbages frühe Experimente mit Rechenmaschinen; AlphaGo, das Programm, das entwickelt worden war, um zu zeigen, dass intellektuelle Fähigkeiten kein Monopol des Menschen sind; Joy Buolamwinis Analyse des Gender-Bias bei Gesichtserkennungssoftware; und die von Shinto-Philosophie und -Ästhetik inspirierte großflächige interaktive digitale Kunst von teamLab. Die Ausstellung veranschaulichte auf wunderbare Weise, dass interdisziplinäres Denken unseren geistigen Horizont enorm erweitern kann.

Erinnern wir uns an das sogenannte Gesetz von Amara: »Wir neigen dazu, die kurzfristigen Auswirkungen einer (neuen) Technologie zu überschätzen und die langfristigen Folgen zu unterschätzen.« Die meisten von uns denken in stark vereinfachter Weise über KI: der Mörderbot aus *Terminator*, inkompetente Algorithmen, die es niemals mit den intellektuellen Fähigkeiten des Menschen aufnehmen oder die Existenz des Menschen bedrohen könnten, seelenlose technologische Erfindungen, die in keiner Weise in der Lage sind, die Welt so wahrzunehmen,

Emotionen so mitzuteilen, Institutionen so zu führen und andere Möglichkeiten des Lebens so zu erkunden, wie dies Menschen können.

Tatsächlich hat der Mensch seit den Anfängen der Zivilisation nach Künstlicher Intelligenz gesucht, schon lange bevor es die Informatik als ein Fachgebiet gab oder der Begriff »KI« Eingang in den Wortschatz gefunden hat. Überliefert ist diese Suche in Geschichten, etwa in dem chinesischen Volksmärchen von Yan Shi, dem Handwerker, der einen humanoiden Roboter baut, bis hin zu Talos, dem bronzenen Automaten in der griechischen Mythologie. Von der fernen Vergangenheit bis zum heutigen Tag hat die unbezwingbare Kraft der KI jede Dimension der menschlichen Zivilisation revolutioniert, und sie wird dies weiterhin tun.

Science-Fiction, mein Metier, spielt eine recht zwiespältige Rolle bei der Erkundung des Mensch-Maschine-Paradigmas. Der 1818 erschienene Roman *Frankenstein*, der oft als der erste moderne Science-Fiction-Roman gerühmt wird, befasst sich mit Fragen, die auch heute noch relevant sind: Haben Menschen das Recht, mithilfe von Technologien intelligentes Leben zu erschaffen, das sich von allen gegenwärtig existierenden Lebensformen unterscheidet? In welcher Beziehung stünden Geschöpf und Schöpfer zueinander? Der Archetypus des verrückten Wissenschaftlers, der der Welt seine Geschöpfe aufzwingt, geht auf Mary Shelleys 200 Jahre altes Meisterwerk zurück.

Während manche Science-Fiction verteufeln und sie für die beschränkte und oftmals negative Wahrnehmung von KI verantwortlich machen, sehen sie nur eine Seite. Science-Fiction kann uns als Warnung dienen, aber spekulatives Geschichtenerzählen vermag auch auf einzigartige Weise, die Grenzen von Raum und Zeit zu überwinden, Ingenieur- und Geisteswissenschaften miteinander zu verbinden, die Trennlinie zwischen Fiktion und Wirklichkeit zu verwischen und Empathie und tiefschürfendes Denken hervorzubringen. Der Historiker und Bestseller-Autor Yuval Noah Harari hat Science-Fiction »das wichtigste künstlerische Genre« unserer Zeit genannt.

Das ist eine hohe Messlatte, der man nur schwer gerecht werden kann. Für Science-Fiction-Schriftsteller wie mich selbst besteht die Herausforderung darin, Geschichten zu erzählen, die nicht nur verbor-

gene Wahrheiten über unsere Gegenwart enthüllen, sondern auch Zukunftsszenarien ausmalen, die noch abenteuerlicher und fantastischer anmuten.

Als sich daher mein ehemaliger Google-Kollege Kai-Fu Lee bei mir meldete und mir die Zusammenarbeit bei *KI 2041* anbot – einem einzigartigen Buchprojekt, das Science-Fiction mit der Analyse bedeutender technologischer Ideen verknüpft –, war ich sofort begeistert. Der Kai-Fu, den ich kenne, ist ein globaler Visionär, ein smarter und richtungsweisender Investor und ein einfallsreicher, aufgeschlossener Tech-Prophet.

Kai-Fu, der über ein profundes Verständnis der Spitzenforschung und ihrer Anwendungen in der Wirtschaft verfügt, beschreibt, auf welche Weise KI innerhalb von 20 Jahren die Gesellschaft auf den verschiedensten Gebieten verändern könnte – von Medizin und Bildung bis hin zu Unterhaltung, Beschäftigung und Finanzen. Seine Idee zu diesem Projekt war ambitioniert, aber es war auch eine Art magische Koinzidenz. Jahre zuvor hatte ich in meiner eigenen schriftstellerischen Tätigkeit den Begiff des »Science-Fiction-Realismus« entwickelt. Für mich ist Science-Fiction nicht nur deshalb faszinierend, weil sie Eskapisten einen Raum der Fantasie eröffnet, in dem sie ihr prosaisches Alltagsleben hinter sich lassen, die Rolle von Superhelden spielen und weit entfernte Galaxien erkunden können, sondern auch deshalb, weil sie ihnen Gelegenheit gibt, sich zeitweise aus der Alltagsrealität zurückzuziehen und kritisch darüber nachzudenken. Wir können in die Zukunft, die wir uns mithilfe von Science-Fiction ausmalen, eingreifen, Dinge verändern und unsere Wirklichkeit aktiv mitgestalten.

Anders gesagt, bevor wir die Zukunft unseren Wünschen entsprechend gestalten können, müssen wir zunächst einmal lernen, sie uns vorzustellen.

Meine Fantasie begann sich zu entwickeln, als ich in jungen Jahren klassische Science-Fiction-Werke wie *Star Wars, Star Trek* und *2001: Odyssee im Weltraum* sah. Seitdem ich zehn Jahre alt bin, sind diese Werke meine Pforte zu den unermesslichen Tiefen des Alls und unbekannten Welten gewesen. Meiner Überzeugung nach sollte man, bevor man eine Erzählung niederschreibt, diese unbedingt in die Geschichte ihres Genres und den größeren gesellschaftlichen Zusammenhang

einordnen. Als jemand, der von den Fantasiewelten der Science-Fiction fasziniert, ja geradezu besessen ist, erstaunt es mich immer wieder, wie breit gefächert das Themenfeld der Science-Fiction ist. So gut wie jedes Sujet beziehungsweise jeder Stil kann in dem Genre seinen Platz finden.

Bevor ich die Schriftstellerei zu meinem Beruf machte, arbeitete ich im Technologiesektor. Viele Leute meinen, Ingenieure und Informatikgenies würden sich kaum für Belletristik interessieren – weil ihre Gehirne für Naturwissenschaften und nicht für Literatur geschaffen seien. Aber während meiner mehr als zehnjährigen Tätigkeit für Technologiefirmen habe ich viele Ingenieure und Techniker kennen gelernt, die aus ihrem Faible für fantastische Literatur keinen Hehl gemacht haben. Diese Begeisterung zeigte sich manchmal in den Namen von Konferenzräumen, wie etwa »Enterprise« oder »Neuromancer«, aber sie war auch bei den intellektuellen Überfliegern anzutreffen, die hinter Projekten wie Google X und Hyperloop standen. Viele Naturwissenschaftler werden bereitwillig zugeben, dass sie sich bei der Entwicklung von zahllosen technischen Neuerungen, angefangen von modernen U-Booten, Laserkanonen und Mobiltelefonen bis zu CRISPR, von literarischen Werken inspirieren ließen. Tatsächlich ist die Fantasie eine Kraft, die unsere Welt prägt.

Ich wollte von Anfang an, dass *KI 2041* das Stereotyp des dystopischen KI-Narrativs infrage stellt – die Art von Erzählung, in der die Zukunft unwiderruflich düster ist. Ohne die Mängel beziehungsweise die Nuancen der KI zu ignorieren, wollten Kai-Fu und ich eine Zukunft schildern, in der KI-Technologien den Einzelnen und Gesellschaften positiv beeinflussen könnten. Wir wollten eine Zukunft darstellen, in der wir gerne leben und die wir gerne mitgestalten würden. Wir malten uns eine Zukunft aus, in der die kommenden Generationen von den Früchten der technologischen Entwicklung profitieren, darauf hinarbeiten, die Welt zu einem angenehmeren Ort zu machen, und ein sinnerfülltes und glückliches Leben führen.

Es war nicht immer leicht, sich die Zukunft unserer Träume vorzustellen. Die Herausforderung für uns bestand darin, uns in die neuesten KI-Forschungsergebnisse zu vertiefen und auf deren Grundlage eine wissenschaftlich fundierte, stringente und vor allem *realistische*

Vision der KI-Szene in 20 Jahren zu entwerfen. Kai-Fu und unser Team verbrachten viele Stunden damit, aktuelle Forschungsarbeiten zu studieren, mit Experten, Fachkräften und Denkern, die in der KI-Branche tätig sind, zu sprechen, an den KI-Workshops des Weltwirtschaftsforums teilzunehmen und führende KI-Unternehmen zu besuchen, um sicherzustellen, dass wir die technologischen und philosophischen Grundlagen der KI-Entwicklung umfassend verstanden.

Die zweite Herausforderung bestand darin, sich die Zukunft der menschlichen Zivilisation auszumalen. Wir wollten darstellen, wie Menschen aus verschiedenen Kulturen und Wirtschaftszweigen und mit unterschiedlichen Identitäten wohl auf den durch KI hervorgerufenen Zukunftsschock reagieren werden. Subtile psychische Reaktionen lassen sich nur schwer logisch-rational erschließen. Um das Porträt der Gefühlswelt der Figuren in unseren Geschichten abzurunden, orientierten wir uns an historischen Vorbildern und ließen uns von ähnlichen weltverändernden Ereignissen der Vergangenheit inspirieren. Um die Fantasie der Lesenden und ihre Fähigkeit, sich andere Lebensbedingungen vorzustellen, anzuregen, mussten wir mit unseren Geschichten auch Empathie wecken, wenn wir unsere Vision und unsere Einstellung umfassend vermitteln wollten. Kai-Fus Analyse dient als die Schnur, die den sich in die Höhe schwingenden Drachen der Fantasie mit der greifbaren Spule der Wirklichkeit verbindet.

Nach monatelanger intensiver Arbeit und mehrmaligem Feinschliff finden Sie nachfolgend die zehn von uns zusammengestellten Zugänge zur Raumzeit des Jahres 2041. Wir hoffen, Sie treten diese Reise mit Neugier, offenem Geist und auch offenem Herzen an.

Eine letzte Sache noch: Für mich besteht der größte Nutzen von Science-Fiction nicht darin, Antworten zu liefern, sondern Fragen aufzuwerfen. Nachdem Sie das Buch beendet haben, werden Sie hoffentlich zahlreiche anregende neue Fragen umtreiben, wie etwa: Kann uns KI dabei helfen, die nächste globale Pandemie zu verhindern, indem sie diese gleich zu Anfang stoppt? Wie können wir mit dem drohenden Verlust von Arbeitsplätzen umgehen? Wie können wir in einer Welt, die von Maschinen beherrscht wird, kulturelle Diversität bewahren? Wie können wir unseren Kindern beibringen, in einer Gesellschaft zu

leben, in der Menschen und Maschinen koexistieren? Wir hoffen, dass die Fragen unserer Leserschaft uns auf dem Weg in eine glücklichere und strahlendere Zukunft ein Stück weiterbringen.

Willkommen im Jahr 2041!

KAPITEL 1
DER GOLDENE ELEFANT

»Es ist besser, sein eigenes Schicksal auf unvollkommene Weise zu leben, als das eines anderen auf eine vollkommene Weise nachzuahmen.«

– Bhagavad Gita (भगवद्गीता, Gesang des Erhabenen, Schrift des Hinduismus), Kapitel 3, Vers 35

ANMERKUNG VON KAI-FU

Den Auftakt bildet eine Erzählung, die uns nach Mumbai führt, zu einer Familie, die sich für ein Versicherungsprogamm entschieden hat, das auf Deep Learning basiert. Dieses dynamische Versicherungsprogramm interagiert mit den Versicherten über eine Reihe von Apps, die darauf abzielen, ihnen eine risikoärmere Lebensführung nahezubringen. Aber die halbwüchsige Tochter der Familie findet, dass die aufdringlichen Empfehlungen des Programms ihre Suche nach Liebe erschweren.

»Der goldene Elefant« führt in die Grundlagen der KI und des Deep Learning ein und vermittelt eine Vorstellung von deren größten Stärken und Schwächen. Insbesondere verdeutlicht die Geschichte, wie eine KI bei der konsequenten Optimierung gewisser Ziele mitunter schädliche Nebeneffekte erzeugen kann. Die Erzählung zeigt auch die Risiken, die auftreten, wenn ein Unternehmen sehr viele Daten über seine Nutzer besitzt. In meiner Analyse am Ende des Kapitels gehe ich auf diese Themen ein, ich werde auch einen kurzen Überblick über die Geschichte der KI geben und mich mit der Frage befassen, warum viele Menschen KI aufregend finden, während ihr andere mit Misstrauen begegnen.

Auf dem Bildschirm schwankte die 3 Meter hohe Ganesha-Figur in der Brandung des Chowpatty Beach, als würde sie sich im Rhythmus der Sitar-Klänge bewegen. Mit jeder Welle tauchte das Abbild der Gottheit tiefer ein, bis es vom Arabischen Meer verschlungen wurde. Im Salzwasser löste sich die Statue in gold- und burgunderfarbenen Schaum auf, der an den Strand gespült wurde, wo die Farben wie Segnungen an den unzähligen Gläubigen hafteten, die sich hier für das Visarjan-Eintauchritual versammelt hatten, mit dem das Ende des Festes Ganesh Chaturthi gefeiert wurde.

Nayana sah in der Wohnung ihrer Familie in Mumbai ihren Großeltern dabei zu, wie sie, vor dem Fernseher sitzend, in die Hände klatschten und mitsangen. Ihr jüngerer Bruder Rohan schob sich Cassava-Chips in den Mund und nahm einen großen Schluck von seiner Diät-Cola. Obwohl er erst acht war, hatte der Arzt empfohlen, seinen Fett- und Zuckerkonsum streng zu kontrollieren. Als er vor Begeisterung mit dem Kopf wackelte, spritzten Krümel aus seinem Mund und flogen über den Fußboden. In der Küche trommelten Papa Sanjay und Mama Riya auf Töpfen und summten vor sich hin, als würden sie in einem Bollywood-Film mitspielen.

Nayana versuchte, sie alle auszublenden. Stattdessen konzentrierte sich die Zehntklässlerin vollkommen auf ihren Smartstream, wo sie FateLeaf heruntergeladen hatte. Nayanas Klassenkameradinnen redeten in letzter Zeit über nichts anderes mehr als über diese neue App. Angeblich hatte sie dank der seherischen Gaben der bedeutendsten Wahrsager Indiens die Antwort auf so gut wie jede Frage.

Die App war – laut ihrer Marken- und Werbekampagne – von dem hinduistischen Weisen Agastya inspiriert, der vor Tausenden von Jahren das vergangene, gegenwärtige und zukünftige Leben aller Menschen in Sanskrit auf Palmblätter, sogenannte Nadi-Blätter, geschrieben haben soll. Der Legende zufolge konnte ein Nadi-Blatt-Wahrsager einer Person, die ihm lediglich ihre Daumenabdrücke und ihr Geburtsdatum zur Verfügung stellen musste, aus dem entsprechenden Blatt ihre Lebensgeschichte weissagen. Das Problem war, dass viele Blätter kolonialistischen Machenschaften, Kriegen und dem Zahn der Zeit zum Opfer gefallen waren. Im Jahr 2025 machte ein Hightech-Unternehmen sämtliche Nadi-Blätter ausfindig, die noch existierten. Das Unternehmen nutzte KI für Deep Learning, das Übersetzen und die Analyse der erhalten gebliebenen Blätter. Das Ergebnis waren virtuelle Nadi-Blätter, die in der Cloud gespeichert wurden – eins für jeden der 8,7 Milliarden Menschen auf der Erde.

Nayana hielt sich nicht mit der uralten Geschichte der Nadi-Blätter auf. Sie hatte ein dringlicheres Anliegen. Nutzer der FateLeaf-App konnten versuchen, das auf ihrem Nadi-Blatt gespeicherte Wissen in Erfahrung zu bringen, indem sie mehrere Fragen stellten. Während ihre Familie die Ganesh-Visarjan-Feierlichkeiten im Fernsehen verfolgte, tippte Nayana nervös eine Frage in der App ein: »Mag mich Sahej?« Bevor sie auf »Senden« klickte, poppte ein Fenster mit dem Hinweis auf, dass eine Antwort auf die Frage 200 Rupien kosten würde. Nayana klickte auf »Abschicken«.

Nayana hatte Sahej gleich von dem Moment an gemocht, als sein Stream sich erstmals mit ihrem virtuellen Klassenzimmer verband. Ihr neuer Klassenkamerad benutzte weder einen Filter noch einen Augmented-Reality-Hintergrund. Hinter Sahej, an der Wand, sah Nayana reihenweise bunte Masken hängen, die, wie sie erfuhr, Sahej selbst geschnitzt und bemalt hatte. Am ersten Tag des neuen Schulhalbjahrs hatte der Lehrer Sahej Fragen zu den Masken gestellt, und der neue Schüler hatte schüchtern ein paar gezeigt und erklärt, dass die Masken indische Gottheiten und Geister mit den Kräften von Superhelden verbanden.

In einem Raum ihres ShareChat, der geladenen Gästen vorbehalten war, lästerten einige von Nayanas Mitschülerinnen gerade über Sahej.

Aus der Möblierung seines Zimmers und der Tatsache, dass sein Familienname in den Schulakten nicht öffentlich einsehbar war, folgerten diese Mädchen, dass Sahej der »vulnerablen Gruppe« angehörte, für die aufgrund staatlicher Auflagen mindestens 15 Prozent der Plätze an ihrer Schule reserviert war. In Privatschulen überall in Indien bekamen diese Kinder praktisch immer einen Platz, und ihr Schulgeld, ihre Bücher und Uniformen wurden mit Stipendien finanziert. »15 Prozent« und »vulnerable Gruppe« waren Euphemismen für die Dalit.

Von Dokumentarfilmen, die sie sich online angeschaut hatte, wusste Nayana von dem uralten indischen Kastensystem, das tief in der hinduistischen Religion und Kultur verankert war. Die Kaste, der eine Person angehörte, hatte ehedem ihre Bildung, ihren Beruf und ihren Ehepartner – ihr ganzes Leben – bestimmt. Ganz unten in dieser Hierarchie standen die Dalit, die manchmal auch spöttisch »Unberührbare« genannt wurden. Über viele Generationen hinweg waren Mitglieder dieser Gemeinschaft gezwungen, die schmutzigsten Arbeiten zu erledigen: Sie mussten Abwasserkanäle säubern, Tierkadaver entsorgen und Tierhäute gerben.

Die indische Verfassung von 1950 verbot Diskriminierung auf Basis der Kastenzugehörigkeit. Aber noch Jahre nach der Unabhängigkeit hatten die Dalit von anderen, als höherrangig geltenden Gruppen getrennte Trink-, Ess-, Wohn- und sogar Bestattungsbereiche. Angehörige der höheren Kasten weigerten sich mitunter sogar, mit den Dalit im selben Raum zu sein, selbst wenn sie zusammen zur Schule gingen oder zusammen arbeiteten.

In den 2010er Jahren bemühte sich die indische Regierung darum, diese Ungerechtigkeiten zu beseitigen, indem sie Behörden und Schulen dazu verpflichtete, 15 Prozent ihrer Stellen beziehungsweise Plätze an Dalit zu vergeben. Die gut gemeinte Maßnahme löste jedoch Kontroversen und sogar gewalttätige Auseinandersetzungen aus. Eltern aus höheren Kasten beschwerten sich, dass solche Zulassungen nicht auf schulischen Leistungen beruhten. Sie behaupteten, ihre Kinder würden den Preis für die Sünden der vorherigen Generationen bezahlen, und Indien tausche lediglich eine Form der Ungleichheit gegen eine andere ein.