

EINLEITUNG

MARTIN FORD

AUTOR, ZUKUNFTSFORSCHER

Künstliche Intelligenz (KI) ist zunehmend keine Science Fiction mehr, sondern wird schnell zur alltäglichen Realität. Unsere Geräte verstehen, was wir sagen, antworten uns und übersetzen Fremdsprachen immer flüssiger. KI-gestützte Bilderkennungsalgorithmen übertreffen Menschen und finden überall Anwendung, wie in selbstfahrenden Autos oder Systemen, die anhand von medizinischen Aufnahmen Krebs diagnostizieren. Große Medienkonzerne nutzen zunehmend einen »automatisierten Journalismus«, der Rohdaten in zusammenhängende Meldungen umwandelt, die von denen eines menschlichen Journalisten praktisch nicht zu unterscheiden sind.

Die Liste der Fähigkeiten wird immer länger, und es wird offensichtlich, dass KI eine der bedeutendsten Kräfte sein wird, die unsere Welt prägen. Im Gegensatz zu spezialisierteren Innovationen wird KI zu einer echten Allzwecktechnologie. Mit anderen Worten: KI entwickelt sich zu einem Bedarfsgut, ähnlich der Elektrizität, das in allen Industriezweigen und in allen Wirtschaftssektoren Verwendung findet und praktisch alle Bereiche der Wissenschaft, der Gesellschaft und der Kultur durchdringt.

Die in den letzten Jahren unter Beweis gestellte Leistungsfähigkeit der KI hat zu umfangreicher Medienberichterstattung und einer Vielzahl von Kommentaren geführt. Unzählige Zeitungsberichte, Bücher, Dokumentarfilme und Fernsehsendungen zählen die Erfolge der KI auf und verkünden den Anbruch eines neuen Zeitalters. Das Ergebnis war manchmal eine unverständliche Mischung aus sorgfältiger Analyse, zusammen mit Hype, Spekulationen und dem, was man fast schon als Panikmache bezeichnen kann. Es heißt, dass schon in ein paar Jahren selbstfahrende Autos auf unseren Straßen unterwegs sein werden – und dass Millionen Jobs für Fernfahrer, Taxifahrer und Uber-Fahrer bald verschwinden werden. Bestimmte Machine-Learning-Algorithmen zeigen ein rassistisches oder frauenfeindliches Verhalten. Bedenken, welchen Einfluss KI-gestützte Technologien wie Gesichtserkennung auf die Privatsphäre haben, erscheinen also wohlbe-gründet. Die Medien berichten regelmäßig über Warnungen, dass Roboter bald bewaffnet sein werden oder dass wirklich intelligente (oder superintelligente) Maschinen eine existenzielle Bedrohung für die Menschheit darstellen. Eine Reihe sehr bekannter Personen des öffentlichen Lebens – von denen keine tatsächlich ein KI-Experte ist – haben sich zu Wort gemeldet. Elon Musk hat sich besonders extrem dazu geäußert und erklärt, dass die KI-Forschung »einen Dämon heraufbeschwört« und dass »KI gefährlicher als Nuklearwaffen« sei. Auch weniger launenhafte Persönlichkeiten, wie Henry Kissinger oder zuletzt Stephen Hawking, haben düstere Warnungen ausgesprochen.

Dieses Buch möchte das Forschungsgebiet Künstliche Intelligenz (KI) sowie die damit verbundenen Chancen und Risiken beleuchten. Dazu dient eine Reihe ausführlicher, weitreichender Gespräche mit einigen der bekanntesten KI-Forschern und in diesem Bereich tätigen Unternehmern. Viele dieser Menschen haben wegweisende Beiträge geleistet, auf denen die überall stattfindenden Umwälzungen unmittelbar beruhen. Andere haben Unternehmen gegründet, die die Forschung in den Bereichen KI, Robotik und Machine Learning vorantreiben.

Das Erstellen einer Liste der bekanntesten und einflussreichsten Persönlichkeiten eines Fachgebiets ist natürlich ein subjektives Unterfangen, und zweifelsohne gibt es viele andere Leute, die entscheidende Beiträge für die Weiterentwicklung der KI geleistet haben oder leisten. Dessen ungeachtet bin ich mir sicher, dass jeder mit

fundierten Kenntnissen des Fachgebiets, der eine Liste der führenden Köpfe erstellen sollte, die die aktuelle KI-Forschung prägen, zu einer Namensliste gelangen würde, auf der sich die meisten der in diesem Buch interviewten Persönlichkeiten befinden. Die hier vorgestellten Männer und Frauen sind tatsächlich die Architekten der Künstlichen Intelligenz – und darüber hinaus der Revolution, die sie bald auslösen wird.

Die hier wiedergegebenen Gespräche waren grundsätzlich zeitlich nicht begrenzt, sind aber so angelegt, dass die dringendsten Fragen zur Sprache kommen, mit denen wir uns bei der Fortentwicklung der KI konfrontiert sehen: Welche Ansätze und Technologien sind am vielversprechendsten, und welche Durchbrüche sind in den kommenden Jahren zu erwarten? Stellen »denkende Maschinen«, also KI auf menschlichem Niveau, eine echte Möglichkeit dar, und wann könnte ein solcher Durchbruch erfolgen? Welchen Risiken oder Bedrohungen, die mit KI verbunden sind, sollten wir ernsthafte Beachtung schenken? Und wie können wir diese Probleme in Angriff nehmen? Welche Rolle spielt eine staatliche Regulierung? Wird KI in der Wirtschaft und auf dem Arbeitsmarkt drastische Umwälzungen verursachen, oder sind solche Bedenken übertrieben? Könnten superintelligente Maschinen sich eines Tages unserer Kontrolle entziehen und zu einer ernsthaften Bedrohung werden? Sollten wir uns Sorgen machen, dass es zu einem »KI-Rüstungswettlauf« kommt, oder dass andere Länder mit autoritären politischen Systemen, insbesondere China, irgendwann die Führung übernehmen?

Es versteht sich von selbst, dass niemand die Antworten auf diese Fragen kennt. Niemand kann die Zukunft vorhersagen. Aber die KI-Experten, mit denen ich gesprochen habe, wissen mehr über den Status quo der Technologie und über die am Horizont erkennbaren Innovationen, als irgendjemand sonst. Sie verfügen über oft jahrzehntelange Erfahrung und haben zur Entstehung der Revolution beigetragen, die sich jetzt anbahnt. Deshalb sollte man ihren Überlegungen und Ansichten Beachtung schenken. Neben den Fragen zum Fachgebiet KI und dessen Zukunft habe ich mich auch mit den Hintergründen, den beruflichen Laufbahnen und den derzeitigen Forschungsinteressen dieser Persönlichkeiten befasst, und ich glaube, dass die diversen Werdegänge und die vielfältigen Lebenswege, die zu ihrer Bekanntheit geführt haben, für eine faszinierende und anregende Lektüre sorgen.

Künstliche Intelligenz ist ein umfassendes Forschungsgebiet mit einer ganzen Reihe von Teilgebieten, und viele der interviewten Forscher waren auf mehreren davon tätig. Einige von ihnen haben auch fundierte Erfahrung in anderen Fachgebieten, wie etwa der Untersuchung der menschlichen kognitiven Fähigkeiten. Dennoch versuche ich im Folgenden kurz zu beschreiben, in welcher Beziehung die hier interviewten Persönlichkeiten zu den wichtigsten Innovationen der KI-For-

schung und den zukünftigen Herausforderungen stehen. Weitere Informationen über die Personen können Sie den Kurzbiografien entnehmen, die Sie jeweils nach dem Interview finden.

Der größte Teil der bahnbrechenden Fortschritte, die im letzten Jahrzehnt erzielt wurden (Bild- und Gesichtserkennung, Übersetzung von Fremdsprachen, Alpha-Gos Erfolge beim Go-Spiel), beruht auf einer Technologie, die als *Deep Learning* bezeichnet wird. Der Begriff *tiefe neuronale Netze* ist ebenfalls gebräuchlich. Künstliche neuronale Netze, in denen, vereinfacht gesagt, Software die Strukturen und Interaktionen biologischer Neuronen im Gehirn emuliert, gibt es schon seit den 1950er-Jahren. Einfache Versionen dieser Netze können rudimentäre Aufgaben der Mustererkennung erledigen und sorgten damals bei den Forschern für ziemlichen Enthusiasmus. In den 1960er-Jahren verloren die Forscher jedoch das Interesse an neuronalen Netzen (was zumindest teilweise auf Kritik an der Technologie von Marvin Minsky, einem der ersten KI-Pioniere, zurückzuführen ist), und sie wurden kaum noch verwendet, während sich die Forscher mit anderen Ansätzen befassten.

Während eines etwa 20-jährigen Zeitraums, der in den 1980er-Jahren begann, befasste sich eine kleine Gruppe von Forschern weiterhin mit neuronalen Netzen und konnte Fortschritte erzielen. Zu dieser Gruppe gehörten insbesondere Geoffrey Hinton, Yoshua Bengio und Yann LeCun. Diese drei Männer leisteten nicht nur maßgebliche Beiträge zur mathematischen Theorie, auf der Deep Learning beruht, sie waren auch die bedeutendsten Fürsprecher der Technologie. Zusammen gelang es ihnen, sehr viel ausgeklügeltere – tiefe – neuronale Netze zu entwickeln, die aus vielen Schichten künstlicher Neuronen bestanden. Fast wie die mittelalterlichen Mönche, die klassische antike Texte verwahrten und kopierten, führten Hinton, Bengio und LeCun neuronale Netze durch ein dunkles Zeitalter – bis nach jahrzehntelangem exponentiellen Wachstum der Rechenleistung und einer nahezu unfassbaren Zunahme der Menge verfügbarer Daten eine Renaissance des Deep Learnings möglich wurde.¹ Aus dem Fortschritt wurde 2012 eine Revolution, als ein Team von Hintons Doktoranden an der University of Toronto an einem bedeutenden Bilderkennungswettbewerb teilnahm und die Konkurrenz dank des Einsatzes von Deep Learning weit hinter sich ließ.

In den darauffolgenden Jahren wurde Deep Learning allgegenwärtig. Alle bedeutenden Technologiekonzerne – Google, Facebook, Microsoft, Amazon, Apple sowie

1 Für ihre herausragenden konzeptionellen und technischen Arbeiten im Bereich Deep Learning erhielten Yoshua Bengio, Geoffrey Hinton und Yann LeCun Ende März 2019 den Turing Award für 2018.

führende chinesische Firmen wie Baidu und Tencent – investierten in die Technologie und machten sie sich zunutze. Die Unternehmen, die Mikroprozessoren und Grafikkchips (GPUs) entwickeln, wie NVIDIA und Intel, mussten ihr Geschäft anpassen und für neuronale Netze optimierte Hardware produzieren. Deep Learning ist, zumindest bis jetzt, die wichtigste Technologie, die für die KI-Revolution verantwortlich ist.

In dieses Buch sind Gespräche mit den drei Deep-Learning-Pionieren Hinton, LeCun und Bengio sowie mit einigen anderen sehr bekannten Spitzenforschern wiedergegeben. Andrew Ng, Fei-Fei Li, Jeff Dean und Demis Hassabis haben Fortschritte bei neuronalen Netzen ermöglicht, die in Bereichen wie Websuche, Computer Vision, selbstfahrenden Autos und eher allgemein intelligenten Systemen Verwendung finden. Sie sind auch als Lehrende, Leiter von Forschungsorganisationen und Unternehmer im Bereich Deep-Learning-Technologie tätig.

In den übrigen Interviews in diesem Buch kommen Personen zu Wort, die man nicht gerade als Deep-Learning-Experten bezeichnen würde, sondern vielleicht sogar als Kritiker. Sie erkennen zwar die im letzten Jahrzehnt erzielten Erfolge an, sind jedoch der Ansicht, dass es nur ein Verfahren unter vielen sei und dass weitere Fortschritte die Integration von Ideen aus anderen Bereichen der KI erfordern. Einige von ihnen, wie Barbara Grosz und David Ferrucci, haben sich stark auf das Problem des Verstehens natürlicher Sprache konzentriert. Gary Marcus und Josh Tenenbaum haben der Untersuchung der kognitiven Fähigkeit des menschlichen Gehirns viel Zeit gewidmet. Andere, wie Oren Etzioni, Stuart Russell und Daphne Koller, sind KI-Generalisten oder haben sich auf probabilistische Verfahren konzentriert. In der letzten Gruppe ist besonders Judea Pearl hervorzuheben, dem 2012 der Turing Award verliehen wurde – die höchste Auszeichnung in der Informatik, vergleichbar dem Nobelpreis –, hauptsächlich für seine Arbeit an probabilistischen Ansätzen in der KI und beim Machine Learning. Von dieser sehr groben Unterscheidung aufgrund ihrer Haltung zum Deep Learning einmal abgesehen, haben sich einige der Forscher, mit denen ich gesprochen habe, auf bestimmte Bereiche konzentriert. Rodney Brooks, Daniela Rus und Cynthia Breazeal sind in der Robotik die führenden Köpfe. Breazeal und Rana el Kaliouby sind bei der Entwicklung von Systemen führend, die Emotionen verstehen und darauf reagieren können und dadurch in der Lage sind, auf sozialer Ebene mit Menschen zu interagieren. Bryan Johnson hat ein Start-up namens Kernel gegründet und hofft, früher oder später Technologie zur Verbesserung der kognitiven Fähigkeiten des menschlichen Gehirns einzusetzen.

Es gibt drei allgemeine Themenbereiche, die ich für so interessant halte, dass ich sie in allen Interviews angesprochen habe. Da wären zunächst einmal die potenziellen Auswirkungen von KI und Robotik auf den Arbeitsmarkt und die Wirt-

schaft. Ich persönlich glaube Folgendes: Wenn KI allmählich unter Beweis stellt, dass sie alle sich wiederholenden, vorhersagbaren Aufgaben automatisiert lösen kann (wobei es keine Rolle spielt, ob es sich um Tätigkeiten von Angestellten oder Arbeitern handelt), wird es unweigerlich zu Ungerechtigkeiten und womöglich sogar zu Arbeitslosigkeit kommen, zumindest bei bestimmten Gruppen von Arbeitern. In meinem Buch *Rise of the Robots: Technology and the Threat of a Jobless Future* aus dem Jahr 2015 (deutscher Titel: *Aufstieg der Roboter*) habe ich das ausführlich dargelegt.

Meine Gesprächspartner hatten vielfältige Ansichten zu dieser potenziellen wirtschaftlichen Umwälzung und den möglichen politischen Gegenmaßnahmen. Ich habe mich an James Manyika gewendet, den Vorsitzenden des McKinsey Global Institute, um diesem Thema auf den Grund zu gehen. Als erfahrener KI- und Robotik-Forscher, der sich seit Kurzem mit den Auswirkungen dieser Technologien auf Organisationen und Arbeitsplätze befasst, hat Manyika eine einzigartige Sichtweise. Das McKinsey Global Institute ist bei Forschungen auf diesem Gebiet führend, und das Interview enthält viele wichtige Erkenntnisse über die Art der sich anbahnenden Umwälzungen.

Die zweite Frage, die ich allen Gesprächspartnern gestellt habe, betrifft die KI auf menschlichem Niveau, die auch als *Artificial General Intelligence* (AGI) bezeichnet wird. AGI ist von Anfang an der Heilige Gral des Fachgebiets KI gewesen. Ich wollte von meinen Gesprächspartner wissen, was sie von der Aussicht auf eine »denkende Maschine« halten, welche Hürden es zu überwinden gilt und wie lange es dauern wird, bis dieses Ziel erreicht ist. Hier hatten alle wichtige Einsichten beizutragen, aber drei Gespräche waren besonders interessant: Demis Hassabis erläutert die gegenwärtig bei DeepMind unternommenen Anstrengungen, der größten und am besten finanzierten Initiative, die darauf ausgerichtet ist, AGI zu erreichen. David Ferrucci, der das Team leitete, das IBMs Watson entwickelt hat, ist jetzt Geschäftsführer bei Elemental Cognition, einem Start-up, das allgemeine Intelligenz durch das Verständnis von Sprache erreichen möchte. Ray Kurzweil, der jetzt bei Google ein Forschungsprojekt über natürliche Sprache leitet, hat zu diesem Thema wichtige Ideen beigetragen (wie so viele andere auch). Kurzweil ist insbesondere durch sein Buch *The Singularity is Near* aus dem Jahr 2005 (deutscher Titel: *Menschheit 2.0: Die Singularität naht*) bekannt. 2012 veröffentlichte er ein Buch über maschinelle Intelligenz, *How to Create a Mind* (deutscher Titel: *Das Geheimnis des menschlichen Denkens*), das die Aufmerksamkeit von Larry Page erlangte, was zu seiner Anstellung bei Google führte.

Ich hatte bei den Gesprächen die Möglichkeit, diese außerordentlich qualifizierte Gruppe von KI-Forschern um eine Einschätzung zu bitten, wann AGI zur Realität wird. Ich stellte die Frage: »In welchem Jahr wird Ihrer Meinung nach mit einer

Wahrscheinlichkeit von 50% KI auf menschlichem Niveau erreicht?«. Die meisten Teilnehmer zogen es vor, ihre Einschätzung anonym abzugeben. Ich habe das Ergebnis dieser formlosen Umfrage in einem Abschnitt am Ende des Buchs zusammengefasst. Zwei Teilnehmer waren bereit, ihre Einschätzung nicht anonym abzugeben, und diese lassen schon erahnen, wie sehr die Meinungen voneinander abweichen. Ray Kurzweil glaubt, wie er zuvor schon des Öfteren gesagt hat, dass KI auf menschlichem Niveau etwa im Jahr 2029 erreicht wird, also in rund zehn Jahren (vom jetzigen Zeitpunkt aus betrachtet). Rodney Brooks hingegen tippt auf das Jahr 2200 – also erst in 180 Jahren. Zu den faszinierendsten Aspekten der Gespräche gehören sicherlich die äußerst unterschiedlichen Ansichten zu einem breiten Spektrum wichtiger Themen.

Das dritte Diskussionsthema betrifft die verschiedenen Risiken, die bei der Weiterentwicklung der KI auftreten werden, sowohl in naher Zukunft als auch langfristig. Die Anfälligkeit von mit dem Internet verbundenen autonomen Systemen für Cyberangriffe oder Hacking gehört zu den Bedrohungen, die schon heute offensichtlich sind. Da KI immer stärker in Wirtschaft und Gesellschaft integriert wird, ist die Lösung dieses Problems eine der wichtigsten Herausforderungen, denen wir gegenüberstehen. Zu den schon jetzt bekannten Problemen gehört auch die Verzerrung (oder Bias) mancher Machine-Learning-Algorithmen, beispielsweise bezüglich der Rasse oder des Geschlechts. Viele meiner Gesprächspartner betonten, wie wichtig es sei, dieses Problem in Angriff zu nehmen und erzählten mir, dass entsprechende Forschungsarbeiten bereits im Gange sind. Einige waren auch durchaus optimistisch, weil sich KI eines Tages als leistungsstarkes Werkzeug im Kampf gegen systemische Verzerrung oder Diskriminierung erweisen könnte.

Eine Gefahr, die viele Forscher thematisieren, ist das Schreckgespenst vollständig autonomer Waffen. Viele Mitglieder der KI-Gemeinschaft glauben, dass KI-gestützte Roboter und Drohnen mit der Fähigkeit, ohne eine Freigabe durch einen Menschen zu töten, früher oder später genauso gefährlich und destabilisierend sind wie biologische oder chemische Waffen. Im Juli 2018 haben mehr als 160 KI-Unternehmen und 2.400 Forscher aus aller Welt, zu denen auch viele meiner Gesprächspartner gehören, eine öffentlich Erklärung unterzeichnet, in der sie versprechen, niemals derartige Waffen zu entwickeln (siehe <https://futureoflife.org/lethal-autonomous-weapons-pledge/>). Die von bewaffneter KI ausgehenden Gefahren kommen in mehreren der Interviews zur Sprache.

Eine sehr viel futuristischere und spekulativere Gefahr ist das sogenannte »Kontrollproblem«. Dabei handelt es sich um die Befürchtung, dass eine intelligente oder vielleicht superintelligente Maschine sich unserer Kontrolle entzieht oder Entscheidungen trifft, die negative Folgen für die Menschheit haben. Diese Be-

fürchtung ist es offenbar, die zu so übertriebenen Aussagen wie der von Elon Musk führen. Fast alle Gesprächspartner hatten zu diesem Thema etwas zu sagen. Um zu gewährleisten, dass ich diesen Bedenken angemessen und ausgewogen Raum gebe, habe ich mit Nick Bostrom vom Future of Humanity Institute der University of Oxford gesprochen. Bostrom ist der Autor des Bestsellers *Superintelligence: Paths, Dangers, Strategies* (deutscher Titel: *Superintelligenz: Szenarien einer kommenden Revolution*), das sich sorgfältig mit den potenziellen Risiken auseinandersetzt, die mit Maschinen verbunden sind, die womöglich viel schlauer als irgendein Mensch sind.

Alle Gespräche wurden zwischen Februar und August 2018 geführt und haben jeweils mindestens eine Stunde gedauert, manchmal auch erheblich länger. Sie wurden aufgezeichnet, transkribiert und von Verlagsmitarbeitern überarbeitet. Der Text wurde schließlich meinen Gesprächspartnern vorgelegt, um ihnen Gelegenheit zu geben, ihn zu korrigieren oder zu ergänzen. Deshalb habe ich vollstes Vertrauen, dass die hier wiedergegebenen Worte die Ansichten der von mir interviewten Personen korrekt widerspiegeln.

Die KI-Experten, mit denen ich gesprochen habe, sind von ganz unterschiedlicher Herkunft und den verschiedensten Organisationen zugehörig. Schon ein kurzes Durchsehen des Buchs zeigt den übergroßen Einfluss, den Google auf die KI-Gemeinschaft hat. Von den 23 Gesprächspartnern sind oder waren sieben bei Google oder dem Mutterkonzern Alphabet tätig. Weitere Schwerpunkte der KI-Forschung sind das MIT und Stanford. Geoff Hinton und Yoshua Bengio sind an der University of Toronto bzw. an der University of Montreal tätig, und die kanadische Regierung hat den guten Ruf ihrer Forschungsinstitute genutzt, um einen strategischen Schwerpunkt auf Deep Learning zu legen. 19 meiner 23 Gesprächspartner arbeiten in den USA. Von diesen 19 wurden allerdings mehr als die Hälfte nicht in den USA geboren. Sie stammen aus Australien, China, Ägypten, Frankreich, Israel, Rhodesien (jetzt Simbabwe), Rumänien und Großbritannien. Ich würde sagen, dass es sich hier um einen ziemlich eindeutigen Beweis dafür handelt, wie groß die Bedeutung der Einwanderung für die technologische Führung der USA ist.

Bei der Durchführung der Interviews hatte ich immer die Vielfalt potenzieller Leser vor Augen, vom professionellen Informatiker über Manager und Investoren bis hin zu praktisch jedem, der ein Interesse an KI und ihrer Auswirkung auf die Gesellschaft hat. Ein besonders wichtiger Teil der Leserschaft besteht aus jungen Menschen, die in Betracht ziehen, eine berufliche Laufbahn im Bereich der KI einzuschlagen. Es fehlt derzeit an qualifiziertem Nachwuchs, insbesondere was die Fähigkeiten bezüglich des Deep Learnings betrifft. Eine berufliche Laufbahn im Bereich KI oder Machine Learning dürfte spannend, lukrativ und bedeutend sein.

Da die Branche daran arbeitet, mehr qualifizierte Menschen für das Fachgebiet zu interessieren, wird zunehmend deutlich, dass mehr unternommen werden muss, um die Diversität dieser Menschen zu gewährleisten. Wenn KI unserer Welt tatsächlich ein neues Gesicht geben soll, dann ist es unverzichtbar, dass diejenigen, die sich am besten mit der Technologie auskennen – und somit auch am besten beeinflussen können, in welche Richtung sie sich bewegt –, die Gesellschaft als Ganzes repräsentieren.

Etwa ein Viertel der von mir interviewten Personen sind Frauen – und dieser Wert ist vermutlich schon höher als der Wert, den man im gesamten Fachgebiet der KI oder des Machine Learnings ermitteln würde. Eine kürzlich durchgeführte Untersuchung ergab, dass nur etwa 12% der führenden Forscher im Bereich Machine Learning Frauen sind (siehe <https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance>). Viele meiner Gesprächspartner betonten die Notwendigkeit, dass sowohl Frauen als auch Mitglieder von Minderheiten stärker repräsentiert werden.

Wie Sie dem Interview mit ihr entnehmen können, liegt einer der führenden im Bereich KI tätigen Frauen die Erhöhung der Diversität besonders am Herzen. Fei-Fei Li von der Stanford University ist Mitbegründerin einer Organisation, die jetzt den Namen AI4ALL trägt (<http://ai-4-all.org/>). Sie veranstaltet Ferienlager, die sich speziell an unterrepräsentierte Highschool-Schüler richten. AI4ALL hat von der Branche beträchtliche Unterstützung erhalten, wie beispielsweise von Google. Mittlerweile gibt es quer durch die USA sechs Universitäten, die Sommerlager veranstalten. Es ist zwar noch viel zu tun, aber es gibt gute Gründe, optimistisch zu sein, dass sich die Diversität der KI-Forscher in den kommenden Jahren und Jahrzehnten beträchtlich erhöhen wird.

Dieses Buch setzt zwar keine technischen Vorkenntnisse voraus, Sie werden jedoch einigen Konzepten und Begriffen des Fachgebiets begegnen. Wenn Sie vorher noch nie mit KI zu tun hatten, bietet sich eine gute Gelegenheit, direkt von den führenden Köpfen des Fachgebiets etwas über die Technologie zu lernen. Um den weniger erfahreneren Lesern den Einstieg zu erleichtern, folgt dieser Einleitung eine kurze Übersicht über das in der KI gebräuchliche Vokabular. Ich empfehle Ihnen, sich die Zeit zu nehmen, diesen Abschnitt zu lesen, bevor Sie mit der Lektüre der Interviews fortfahren. Darüber hinaus enthält das Interview mit Stuart Russell, der Koautor des führenden KI-Lehrbuchs *Artificial Intelligence: A Modern Approach* (deutscher Titel: *Künstliche Intelligenz. Ein moderner Ansatz*) ist, eine Erklärung vieler wichtiger Konzepte des Fachgebiets.

Es war mir eine besondere Ehre, die in diesem Buch wiedergegebenen Gespräche führen zu dürfen. Ich denke, Sie werden feststellen, dass alle meine Gesprächspartner sich wohlüberlegt und verständlich ausdrücken und sich sehr dafür engagieren, zu gewährleisten, dass die Technologie, an der sie arbeiten, zum Vorteil

der Menschheit eingesetzt wird. Eher selten werden Sie umfassenden Konsens vorfinden. Das Buch enthält eine Vielzahl von Erkenntnissen, Meinungen und Vorhersagen, die oftmals miteinander in Konflikt stehen. Das muss Ihnen klar sein: KI ist ein umfassendes, offenes und breit gefächertes Fachgebiet. Die Art kommender Innovationen, die Geschwindigkeit, mit der sie erreicht werden und ihre Anwendungsgebiete liegen noch völlig im Dunkeln. Diese Kombination aus potenziellen Umwälzungen und der grundsätzlichen Ungewissheit macht es unausweichlich, dass wir eine sinnvolle und umfassende Debatte über die Zukunft der KI führen und welchen Einfluss sie auf unser Leben hat. Ich kann mir nur wünschen, dass dieses Buch einen Beitrag zu dieser Debatte leisten wird.

Übersicht über das in der KI gebräuchliche Vokabular

Die Gespräche umfassen ein breites Themenspektrum und gehen manchmal auf bestimmte Verfahren ein, die in der KI verwendet werden. Für das Verständnis sind keine technischen Vorkenntnisse erforderlich, aber hin und wieder werden Sie auf gebräuchliche Fachbegriffe stoßen. Im Folgenden finden Sie eine kurze Übersicht der wichtigsten Begriffe, die Ihnen in den Interviews begegnen werden. Wenn Sie sich die Zeit nehmen, den folgenden Abschnitt zu lesen, sind Sie bestens für die Lektüre des Buchs gewappnet. Wenn Ihnen ein Abschnitt zu detailliert oder zu technisch erscheint, schlage ich vor, ihn einfach zu überspringen.

Machine Learning ist ein Teilgebiet der KI, das sich damit befasst, Algorithmen zu erstellen, die aus Daten lernen können. Man könnte auch sagen, dass Machine-Learning-Algorithmen Programme sind, die sich im Wesentlichen selbst programmieren, indem sie auf Informationen zurückgreifen. Es heißt oft, »Computer erledigen nur die Aufgaben, für die sie programmiert werden«, aber durch den Aufstieg des Machine Learnings trifft diese Aussage immer weniger zu. Es gibt viele verschiedene Arten von Machine-Learning-Algorithmen, aber Deep Learning hat zu den größten Umwälzungen geführt (und die größte Aufmerksamkeit der Medien auf sich gezogen).

Deep Learning ist ein Machine-Learning-Verfahren, das tiefe (aus vielen Schichten bestehende) **künstliche neuronale Netze** verwendet – also Software, die, vereinfacht gesagt, die Funktionsweise von Neuronen im Gehirn emuliert. Für die Revolution der KI, die wir im vergangenen Jahrzehnt erlebt haben, ist vor allem Deep Learning verantwortlich.

Es gibt noch einige andere Begriffe, die weniger technisch interessierte Leser einfach als »irgendetwas hinter den Kulissen des Deep Learnings« interpretieren können. Es steht Ihnen völlig frei, einen Blick hinter die Kulissen zu werfen und die Fachbegriffe genauer in Augenschein zu nehmen:

Backpropagation (oder kurz **Backprop**) ist der Lernalgorithmus, den Deep-Learning-Systeme verwenden. Beim Trainieren eines neuronalen Netzes (siehe nachfolgend »überwachtes Lernen«) breiten sich Informationen rückwärts durch die Neuronenschichten aus, die das Netz bilden und bewirken eine Neukalibrierung der Gewichte der einzelnen Neuronen. Das führt dazu, dass das Netz insgesamt der richtigen Lösung allmählich näher kommt. Geoff Hinton ist Koautor des wegweisenden Papers über Backpropagation aus dem Jahr 1986. In seinem Interview geht er ausführlicher auf Backpropagation ein.

Der Begriff **Gradientenabstiegsverfahren** ist noch undurchsichtiger. Er bezieht sich auf ein spezielles mathematisches Verfahren, das der Backpropagation-Algorithmus beim Training des Netzes verwendet, um den Fehler zu verringern.

Möglicherweise begegnen Ihnen auch Begriffe, die verschiedene Arten bezeichnen, wie etwa **rekurrente neuronale Netze (RNNs)** und **Convolutional Neural Networks (CNNs)** oder **Boltzmann-Maschinen**. Die Unterschiede betreffen im Allgemeinen die Art und Weise, wie die Neuronen miteinander verknüpft sind. Die Details sind sehr technisch und gehen über den Rahmen dieses Buchs hinaus. Ich habe Yann LeCun, der als Erfinder der CNNs gilt, die bei Anwendungen der Computer Vision häufig zum Einsatz kommen, dennoch gebeten, das Konzept kurz zu erklären.

Wenn ein Begriff den Namen **Bayes** enthält, geht es im Allgemeinen um Wahrscheinlichkeiten oder um die Anwendung der Regeln der Wahrscheinlichkeitsrechnung. Ihnen könnten beispielsweise Begriffe wie Bayes'sches Machine Learning oder Bayes-Netze begegnen, die Algorithmen bezeichnen, die auf den Regeln der Wahrscheinlichkeitsrechnung beruhen. Die Bezeichnungen gehen auf den Namen des Pfarrers Thomas Bayes (1701–1761) zurück, der eine Methode entwickelte, die Wahrscheinlichkeit eines Ereignisses anhand neu vorliegender Informationen vorherzusagen. Bayes'sche Verfahren erfreuen sich bei Informatikern und Forschern, die versuchen, die menschlichen kognitiven Fähigkeiten zu modellieren, großer Beliebtheit. Judea Pearl, der auch zu meinen Gesprächspartnern gehörte, erhielt die höchste Auszeichnung in der Informatik, den Turing Award, zum Teil für seine Arbeit über Bayes'sche Verfahren.

Wie KI-Systeme lernen

Es gibt verschiedene Möglichkeiten, Machine-Learning-Systeme zu trainieren. Innovationen auf diesem Gebiet, also neue Methoden zum Trainieren von KI-Systemen zu entwickeln, ist für zukünftige Fortschritte unverzichtbar.

Überwachtes Lernen bedeutet, einem Lernalgorithmus sorgfältig vorbereitete Trainingsdaten bereitzustellen, die kategorisiert oder gekennzeichnet sind. Sie könnten beispielsweise einem Deep-Learning-System beibringen, einen Hund auf

einem Foto zu erkennen, indem Sie es mit vielen Tausend (oder sogar Millionen) Hundefotos füttern. Diese Bilder wären als »Hund« gekennzeichnet. Darüber hinaus müssten Sie auch eine Menge Bilder bereitstellen, auf denen kein Hund zu sehen ist und diese mit »Kein Hund« kennzeichnen. Nachdem das System trainiert wurde, können Sie völlig neue Bilder als Eingabe verwenden, und es wird als Ausgabe »Hund« oder »Kein Hund« liefern. Dabei kann es eine Leistungsstärke erreichen, die diejenige eines typischen Menschen übertrifft.

Bei aktuellen KI-Systemen ist überwachtes Lernen das weitaus am häufigsten verwendete Verfahren. Es wird in rund 95% der Anwendungen eingesetzt, etwa bei der Übersetzung von Fremdsprachen (das Training erfolgt mit Millionen Dokumenten, die zweisprachig vorliegen) oder bei KI-gestützten Radiologiesystemen (das Training erfolgt mit Millionen medizinischer Bilder, die mit »Krebs« oder »Kein Krebs« gekennzeichnet sind). Beim überwachten Lernen gibt es das Problem, dass Unmengen gekennzeichnete Daten benötigt werden. Das erklärt, weshalb Unternehmen wie Google, Amazon oder Facebook, die über gigantische Datenmengen verfügen, bei der Deep-Learning-Technologie eine so dominierende Stellung einnehmen.

Reinforcement Learning (dt. auch **bestärkendes oder verstärkendes Lernen**) bedeutet im Wesentlichen, durch Üben zu lernen oder durch die Trial-and-Error-Methode. Anstatt einen Algorithmus durch die Bereitstellung korrekt gekennzeichnete Daten zu trainieren, überlässt man es dem System, selbst eine Lösung zu finden, und wenn es erfolgreich ist, gibt es eine »Belohnung«. Stellen Sie sich vor, dass Sie Ihrem Hund beibringen wollen, auf das Kommando »Sitz!« zu hören. Wenn er gehorcht, gibt es ein Leckerli. Reinforcement Learning hat sich bei der Entwicklung von KI-Systemen, die Spiele spielen, als besonders leistungsstark erwiesen. Aus dem Interview mit Demis Hassabis werden Sie erfahren, dass DeepMind ein klarer Befürworter des Reinforcement Learnings ist und es bei der Entwicklung des AlphaGo-Systems einsetzte.

Beim Reinforcement Learning gibt es das Problem, dass der Algorithmus sehr viele Übungsdurchgänge benötigt, bis er erfolgreich ist. Aus diesem Grund wird er vornehmlich für Spiele verwendet oder für Aufgaben, die auf einem Computer sehr schnell simuliert werden können. Auch bei der Entwicklung selbstfahrender Autos kann Reinforcement Learning eingesetzt werden, aber natürlich nicht, indem echte Autos auf der Straße üben, sondern durch das Trainieren virtueller Autos in simulierten Umgebungen. Nachdem die Software trainiert wurde, kann sie in echten Autos eingesetzt werden.

Unüberwachtes Lernen bedeutet, dass Maschinen direkt aus unstrukturierten Daten lernen, die ihrer Umgebung entstammen. Auf diese Weise lernen Menschen. Kleine Kinder lernen beispielsweise das Sprechen hauptsächlich dadurch, dass sie ihren Eltern zuhören. Überwachtes Lernen und Reinforcement Learning

spielen auch eine Rolle, aber das menschliche Gehirn verfügt über die erstaunliche Fähigkeit, allein durch Beobachtung und unüberwachte Interaktion mit der Umgebung zu lernen.

Unüberwachtes Lernen stellt eines der vielversprechendsten Verfahren dar, um in der KI Fortschritte zu erzielen. So sind etwa Systeme denkbar, die selbstständig lernen, ohne dass große Mengen gekennzeichnete Trainingsdaten erforderlich sind. Gleichzeitig ist unüberwachtes Lernen allerdings auch eine der schwierigsten Herausforderungen. Ein Durchbruch, der es ermöglichen würde, dass Maschinen tatsächlich unüberwacht effizient lernen, wäre eines der bedeutendsten Ereignisse in der Geschichte der KI und ein Meilenstein auf dem Weg zur KI auf menschlichem Niveau.

Artificial General Intelligence (AGI) bezeichnet eine denkende Maschine. AGI wird im Allgemeinen mehr oder weniger als Synonym für die Begriffe **KI auf menschlichem Niveau**, **generelle/allgemeine (Künstliche) Intelligenz** oder **starke KI** betrachtet. Sie kennen sicherlich verschiedene Beispiele für eine AGI – die aber alle der Science Fiction entstammen: HAL aus dem Film *2001: A Space Odyssey* (*2001: Odyssee im Weltraum*), der Hauptcomputer der Enterprise (oder Mr. Data) aus *Star Trek*, C3PO aus *Star Wars* (*Krieg der Sterne*) oder Agent Smith aus *Matrix*. Diese erdachten Systeme wären in der Lage, den **Turing-Test** zu bestehen – sie könnten also ein Gespräch führen, das sich nicht von einem Gespräch mit einem Menschen unterscheiden lässt. Alan Turing schlug diesen Test in seinem Papier *Computing Machinery and Intelligence* aus dem Jahr 1950 vor, das wohl KI als Forschungsgebiet begründet hat. Mit anderen Worten: AGI war von Anfang an das Ziel.

Wenn wir es eines Tages geschafft haben, eine AGI zu erreichen, wird das smarte System wahrscheinlich schnell noch smarter werden. Wir erleben dann das Entstehen einer **Superintelligenz**, einer Maschine, deren allgemeinen intellektuellen Fähigkeiten diejenigen irgendeines Menschen übertreffen. Das könnte sich einfach durch leistungsfähigere Hardware ergeben, das Ganze würde sich aber deutlich beschleunigen, wenn eine intelligente Maschine sich darauf konzentriert, noch smartere Versionen von sich selbst zu entwickeln. Das könnte zu dem führen, was als »rekursiver Verbesserungszyklus« oder »schneller Intelligenz-Takeoff« bezeichnet wurde. Dieses Szenario hat zu den Bedenken bezüglich des Außer-Kontrolle-Geratens und zum KI-Ausrichtungsproblem geführt – dass ein superintelligentes System auf eine Weise handelt, die nicht im Interesse der Menschheit liegt.

Ich bin zu dem Schluss gekommen, dass der Weg zur AGI und die Aussicht auf eine Superintelligenz so interessante Themen sind, dass ich sie mit allen Gesprächspartnern diskutiert habe.

Martin Ford ist Zukunftsforscher und Autor zweier Bücher, dem Bestseller der *New York Times* *Rise of the Robots: Technology and the Threat of a Jobless Future* (Gewinner des *Financial Times/McKinsey Business Book of the Year Award 2015*, übersetzt in mehr als 20 Sprachen, deutscher Titel: *Aufstieg der Roboter*) und *The Lights in the Tunnel: Automation, Accelerating Technology and the Economy of the Future*. Er ist außerdem Gründer einer im Silicon Valley ansässigen Softwareentwicklungsfirma. Seinen TED-Talk über die Auswirkungen von KI und Robotik auf Wirtschaft und Gesellschaft haben sich mehr als 2 Millionen Zuschauer angesehen.

Ford ist beratender KI-Experte des neuen »*Rise of the Robots Index*« der *Société Générale*, die zur *Lyxor Robotics & AI ETF* gehört, die in Unternehmen investiert, die bei der KI- und Robotik-Revolution eine bedeutende Rolle spielen. Er besitzt einen Abschluss in Informatik der *University of Michigan, Ann Arbor*, und einen betriebswirtschaftlichen Studienabschluss der *University of California, Los Angeles*.

Er befasst sich mit zukünftigen Technologien und ihren Auswirkungen und hat schon für *The New York Times*, *Fortune*, *Forbes*, *The Atlantic*, *The Washington Post*, *Harvard Business Review*, *The Guardian* und *The Financial Times* geschrieben. Er ist in zahllosen Radio- und Fernsehsendungen zu Gast gewesen, unter anderem bei *NPR*, *CNBC*, *CNN*, *MSNBC* und *PBS*. Ford hält regelmäßig Vorträge über die zunehmenden Fortschritte der Robotik und der KI und was diese Fortschritte für die Wirtschaft, den Arbeitsmarkt und die Gesellschaft der Zukunft bedeuten.

Ford ist weiter als Unternehmer tätig und engagiert sich aktiv als Vorstandsmitglied von *Genesis Systems*, einem Start-up, das eine revolutionäre Technologie zur Gewinnung von Wasser aus der Atmosphäre (*Atmospheric Water Generator, AWG*) entwickelt hat. *Genesis* beabsichtigt, schon bald in den trockensten Regionen der Erde automatisierte Systeme mit eigener Energieversorgung einzurichten, die in industriellem Maßstab Wasser aus der Luft gewinnen.



“ *In dem Szenario, das mir vorschwebt, werden wir medizinische Nanoroboter in unsere Blutbahn injizieren. [...] Diese Roboter werden auch das Gehirn bevölkern. Virtual Reality und Augmented Reality werden nicht durch an die Gliedmaßen unseres Körpers angeschlossene Geräte bereitgestellt, sondern aus dem Inneren des Nervensystems heraus.*

RAY KURZWEIL

LEITER ENGINEERING BEI GOOGLE

Ray Kurzweil gehört zu den führenden Erfindern, Denkern und Zukunftsforschern. Ihm wurden 21 Ehrendokortitel verliehen, und er wurde von drei US-Präsidenten ausgezeichnet. Ihm wurde der MIT-Lemelson-Innovationspreis verliehen, und 1999 erhielt er von Präsident Clinton die National Medal of Technology, die höchste Auszeichnung im Bereich Technologie. Kurzweil ist auch ein produktiver Autor und hat fünf Bestseller verfasst. 2012 kam er zu Google, wo er ein Forschungsteam leitet, das an Maschinenintelligenz und dem Verstehen natürlicher Sprache arbeitet. Anfang 2019 erschien Kurzweils erster Roman, Danielle, Chronicles of a Superheroine. Ein weiteres Buch von Kurzweil, The Singularity is Nearer, wird voraussichtlich Ende 2019 erscheinen.

MARTIN FORD: Wie sind Sie zur KI gekommen?

RAY KURZWEIL: Ich kam 1962 erstmals mit KI in Berührung, das war nur sechs Jahre nachdem der Ausdruck auf der Dartmouth-Konferenz 1956 in Hanover, New Hampshire, von Marvin Minsky und John McCarthy geprägt wurde.

Das Fachgebiet KI war damals schon in zwei sich bekämpfende Lager aufgespalten: die Verfechter symbolischer Modelle und die Anhänger konnektionistischer Modelle. Erstere hatten damals eine Vormachtstellung, und Marvin Minsky galt als ihr wichtigster Vertreter. Die Konnektionisten waren die Emporkömmlinge, und einer davon war Frank Rosenblatt von der Cornell University, der das erste neuronale Netz bekannt gemacht hatte, das als Perzeptron bezeichnet wurde. Ich schrieb beiden Briefe, und beide luden mich ein, also besuchte ich zunächst Minsky, der den ganzen Tag mit mir verbrachte, und wir bauten eine persönliche Beziehung auf, die 55 Jahre währen sollte. Wir sprachen über KI, die damals ein ziemlich undurchsichtiges Fachgebiet war, dem eigentlich niemand Beachtung schenkte. Er fragte mich, wen ich als nächsten besuchen würde, und als ich Dr. Rosenblatt erwähnte, meinte er, die Mühe könne ich mir sparen.

Dann traf ich Dr. Rosenblatt, der dieses einschichtige neuronale Netz namens Perzeptron entwickelt hatte. Dabei handelte es sich um ein Gerät mit einer Kamera. Ich hatte einige ausgedruckte Buchstaben zu meinem Treffen mit Dr. Rosenblatt mitgebracht, die sein Gerät perfekt erkannte, sofern sie in der Schriftart Courier 10 gedruckt waren.

Andere Schriftarten funktionierten nicht so gut, und er sagte: »Kein Problem, ich kann die Ausgabe des Perzeptrons als Eingabe eines zweiten Perzeptrons verwenden, und dessen Ausgabe können wir in eine dritte Schicht einspeisen. Und wenn wir weitere Schichten hinzufügen, wird es intelligenter werden und kann generalisieren und all diese bemerkenswerten Dinge leisten«. Ich fragte ihn: »Haben Sie das ausprobiert?«, und er sagte: »Noch nicht, aber es steht auf unserer Forschungsagenda ganz oben«.

In den 1960er-Jahren ging alles nicht so schnell wie heute, und leider verstarb er 1971, neun Jahre später, ohne diese Idee ausprobiert zu haben. Diese Idee war allerdings bemerkenswert vorausschauend. Die ganze Aufregung um neuronale Netze wurde von diesen tiefen neuronalen Netzen mit vielen Schichten ausgelöst. Das war eine wirklich erstaunliche Erkenntnis, denn es war tatsächlich überhaupt nicht klar, ob die Idee funktionieren würde.

1969 schrieb Minsky zusammen mit seinem Kollegen Seymour Papert sein Buch *Perceptrons*. Im Wesentlichen bewies das Buch ein Theorem, das besagt, dass ein Perzeptron keine Antworten liefern kann, die es erfordern, die Logikfunktion XOR zu verwenden, und dass es das Problem zusammenhängender Räume nicht lösen

kann. Auf dem Buchumschlag sind zwei labyrinthartige Muster abgebildet, und wenn man genau hinsieht, kann man erkennen, dass eins zusammenhängend ist und das andere nicht. Diese Klassifizierung ist das Problem zusammenhängender Räume. Das Theorem bewies, dass ein Perzeptron diese Aufgabe nicht lösen kann. Das Buch sorgte dafür, dass es die folgenden 25 Jahre keine Fördergelder für die Erforschung des Konnektionismus gab, was Minsky bedauerte, denn kurz vor seinem Tod sagte er mir, dass er die Leistungsfähigkeit tiefer neuronaler Netze mittlerweile anerkennt.

MARTIN FORD: Hat Marvin Minsky in den 1950er-Jahren nicht selbst an den ersten konnektionistischen neuronalen Netzen gearbeitet?

RAY KURZWEIL: Das stimmt, aber in den 1960er-Jahren war er davon enttäuscht und wollte die Leistungsfähigkeit mehrschichtiger neuronaler Netze nicht anerkennen. Sie wurde erst Jahrzehnte später deutlich, als dreischichtige neuronale Netze ausprobiert wurden, die etwas besser funktionierten. Die Verwendung mehrerer Schichten bereitete Schwierigkeiten wegen des sogenannten Problems der verschwindenden und explodierenden Gradienten (engl. *vanishing and exploding gradient problem*), das im Wesentlichen dafür verantwortlich ist, dass sich der dynamische Wertebereich der Koeffizienten verschlechtert, weil die Zahlen zu groß oder zu klein werden.

Geoffrey Hinton und eine Gruppe von Mathematikern haben dieses Problem gelöst, und nun können wir beliebig viele Schichten verwenden. Die Lösung besteht darin, dass man die Informationen in jeder Schicht neu kalibriert, damit der repräsentierbare Wertebereich nicht überschritten wird. Solche 100-schichtigen neuronalen Netze waren äußerst erfolgreich. Es gibt allerdings noch immer ein Problem, das sich unter dem Motto »Das Leben beginnt bei einer Milliarde Beispielen« zusammenfassen lässt.

Dass ich bei Google bin, hat unter anderem den Grund, dass hier Milliarden von Beispielen zur Verfügung stehen, wie etwa Bilder von Hunden und Katzen, oder anderen Bildkategorien, die gekennzeichnet sind. Für viele Objekte gibt es aber auch keine Milliarden Beispiele. Wir verfügen über viele Sprachbeispiele, die jedoch nicht mit ihrer Bedeutung gekennzeichnet sind, und wie sollen wir sie in der Sprache kennzeichnen, die wir von vornherein nicht verstehen? Bei einer bestimmten Kategorie von Aufgaben können wir das Problem umgehen, und dafür ist Go-Spielen ein gutes Beispiel. Das DeepMind-System wurde mit allen online verfügbaren Spielzügen trainiert, deren Anzahl in der Größenordnung von einer Million liegt, nicht einer Milliarde. Das System spielte so gut wie ein mittelmäßiger Amateur, also wurden weitere 999 Millionen Beispiele benötigt, aber woher soll man die nehmen?

MARTIN FORD: Sie wollen darauf hinaus, dass Deep Learning und überwachtes Lernen sehr stark von gekennzeichneten Daten abhängig sind.

RAY KURZWEIL: Richtig. Man kann das Problem umgehen, wenn man die Welt, in der man sich befindet, simulieren kann. Dann kann man seine eigenen Trainingsdaten erzeugen. Und genau das hat DeepMind gemacht, indem es gegen sich selbst gespielt hat. Die Kennzeichnung der Spielzüge konnte durch klassische Methoden erfolgen. Nachfolgend hat AlphaZero ein neuronales Netz trainiert, um die Kennzeichnung der Spielzüge zu verbessern, und war so in der Lage, ohne vom Menschen stammende Trainingsdaten alle 100 Spiele gegen AlphaGo zu gewinnen.

Die Frage ist, in welchen Situationen das möglich ist. Mathematik ist ein weiterer Anwendungsbereich, in der das funktioniert, denn Mathematik können wir simulieren. Die Axiome der Zahlentheorie sind nicht komplizierter als die Spielregeln von Go.

Ein weiteres Beispiel sind selbstfahrende Autos, obwohl Autofahren viel komplizierter ist als ein Brettspiel oder mathematische Axiome. Waymo hat ein ziemlich gutes System entwickelt, das mehrere Methoden kombiniert, und ließ Fahrzeuge Millionen von Kilometern fahren, wobei Menschen an Bord waren, die jederzeit das Steuer übernehmen konnten. Auf diese Weise wurden genug Daten erzeugt, um einen präzisen Fahrsimulator zu erstellen. Sie haben im Simulator mehr als eine Milliarde Kilometer mit simulierten Fahrzeugen zurückgelegt und damit Trainingsdaten für ein tiefes neuronales Netz generiert, das dafür ausgelegt ist, den Algorithmus zu verbessern. Das hat funktioniert, obwohl die Fahrsimulation sehr viel komplexer als ein Brettspiel ist.

Der nächste spannende Bereich, den man zu simulieren versucht, ist die Welt der Biologie und Medizin. Wenn wir die Biologie simulieren könnten – und das ist nicht unmöglich –, dann wären wir in der Lage, klinische Studien in Stunden, statt in Jahren durchzuführen, und wir könnten wie bei selbstfahrenden Autos, Brettspielen oder Mathematik, unsere eigenen Daten erzeugen.

Das ist aber nicht der einzige Ansatz zur Lösung des Problems, genügend Trainingsdaten bereitzustellen. Menschen können aus sehr viel weniger Daten lernen, weil wir Transfer Learning betreiben. Wir verwenden Wissen, das wir in Situationen erlernt haben, die sich von dem, was wir gerade zu erlernen versuchen, deutlich unterscheiden. Ich habe ein anderes Lernmodell entwickelt, das auf einer groben Vorstellung der Funktionsweise des Neokortex beruht. 1962 habe ich eine These über die Arbeitsweise des menschlichen Gehirns aufgestellt – und ich habe die letzten 50 Jahre über das Denken nachgedacht. Mein Modell ist kein großes neuronales Netz, sondern ist aus vielen kleinen Modulen zusammengesetzt, die jeweils ein Muster erkennen können. In meinem Buch *How to Create a Mind* be-

schreibe ich, dass der Neokortex im Wesentlichen aus 300 Millionen solcher Module besteht, die jeweils ein sequenzielles Muster erkennen können und dabei einen gewissen Grad an Variabilität zulassen. Die Module sind in einer Hierarchie organisiert, die durch ihr eigenes Denken entsteht. Das System entwickelt seine eigene Hierarchie.

Dieses hierarchische Modell des Neokortex kann aus sehr viel weniger Daten lernen. Beim Menschen verhält es sich ebenso. Wir können aus kleinen Datenmengen lernen, weil wir das Wissen aus einem Bereich auf einen anderen generalisieren können.

Larry Page, einer der Gründer von Google, gefiel meine These in *How to Create a Mind*, und er stellte mich bei Google ein, damit ich diese Ideen auf das Verstehen von Sprache anwende.

MARTIN FORD: Gibt es ein praktisches Beispiel für die Anwendung dieser Konzepte auf ein Google-Produkt?

RAY KURZWEIL: Gmails Smart Reply, das jeweils drei Vorschläge für die Beantwortung einer E-Mail unterbreitet, ist eine von meinem Team entwickelte Anwendung, die dieses hierarchische System verwendet. Wir haben gerade Talk to Books (<https://books.google.com/talktobooks/>) vorgestellt. Sie stellen eine Frage in natürlicher Sprache, und das System liest in einer halben Sekunde 100.000 Bücher – das entspricht 600 Millionen Sätzen – und liefert die besten Antworten zurück, die es in den 600 Millionen Sätzen finden konnte. Das Ganze beruht auf semantischem Verstehen, nicht auf Schlüsselwörtern.

Bei Google machen wir Fortschritte beim Verstehen natürlicher Sprache, und Sprache entstand ursprünglich im Neokortex. Sprache ist hierarchisch; wir können die hierarchischen Vorstellungen in unserem Neokortex miteinander teilen, indem wir die Hierarchie der Sprache verwenden. Ich denke, es war sehr vorausschauend von Alan Turing, dass er den Turing-Test auf Sprache aufbaute, denn ich glaube, dass es des gesamten Spektrums menschlichen Denkens und menschlicher Intelligenz bedarf, um Sprache auf menschlichem Niveau zu erschaffen und zu verstehen.

MARTIN FORD: Haben Sie letztendlich zum Ziel, diese Ideen auszubauen, um tatsächlich eine Maschine zu entwickeln, die den Turing-Test besteht?

RAY KURZWEIL: Nicht alle sind dieser Meinung, aber ich halte den Turing-Test, sofern er richtig durchgeführt wird, tatsächlich für einen sehr guten Test für Intelligenz auf menschlichem Niveau. Problematisch ist, dass es in der kurzen Arbeit, die Turing 1950 verfasst hat, nur in einigen wenigen Absätzen um den Turing-

Test geht, und dass er entscheidende Elemente weggelassen hat. Beispielsweise beschreibt er nicht, wie bei der Durchführung des Tests tatsächlich verfahren werden soll. Die Vorschriften des Tests sind ziemlich kompliziert, wenn man ihn in der Praxis durchführt, aber wenn ein Computer ihn bestehen soll, dann muss er, wie ich glaube, den vollen Umfang der menschlichen Intelligenz besitzen. Das eigentliche Ziel ist das Verständnis von Sprache auf menschlichem Niveau. Wenn eine KI dazu in der Lage ist, kann sie alle Dokumente und alle Bücher lesen und alles andere erlernen. Wir kommen dem ganz allmählich und schrittweise näher. Wir können beispielsweise genug Semantik verstehen, um mit Talk to Books vernünftige Antworten auf die Fragen zu geben, aber noch nicht auf menschlichem Niveau. Mitch Kapur und ich haben eine langfristige Wette um 20.000 Dollar laufen, die der Verlierer einer wohltätigen Organisation nach Wahl des Gewinners spenden muss. Ich behaupte, dass eine KI den Turing-Test bis 2029 besteht, er hält das nicht für möglich.

MARTIN FORD: Stimmen Sie zu, dass es beim Turing-Test überhaupt keine zeitliche Beschränkung geben sollte, damit er Intelligenz zuverlässig erkennt? Jemanden 15 Minuten lang hinters Licht zu führen, erscheint ein wenig wie Spielerei.

RAY KURZWEIL: Ganz genau. Und wenn Sie sich die Regeln ansehen, die Mitch Kapur und ich vereinbart haben, werden Sie feststellen, dass wir mehrere Stunden vorgesehen haben, und selbst das ist vielleicht nicht genug Zeit. Wenn eine KI Sie wirklich davon überzeugt, menschlich zu sein, dann besteht sie den Test, das ist das Entscheidende. Wir können darüber diskutieren, wie lange das dauern muss – vielleicht mehrere Stunden lang, wenn der Schiedsrichter anspruchsvoll ist –, aber ich stimme zu, dass man mit einfachen Tricks davonkommen könnte, wenn die Zeit zu knapp bemessen ist.

MARTIN FORD: Ich glaube, es ist einfach, sich einen intelligenten Computer vorzustellen, dem es nicht besonders gut gelingt, sich als Mensch auszugeben, weil es sich um eine fremde Intelligenz handelt. Deshalb erscheint es wahrscheinlich, dass es einen Test geben könnte, bei dem alle übereinstimmen, dass eine Maschine intelligent ist, obwohl sie offenbar nicht menschlich ist. Wir würden das vermutlich auch als einen angemessenen Test anerkennen.

RAY KURZWEIL: Wale und Tintenfische besitzen große Gehirne und zeigen intelligentes Verhalten, sind aber offensichtlich nicht in der Lage, den Turing-Test zu bestehen. Ein Chinese, der nicht Englisch, sondern Mandarin spricht, würde den englischen Turing-Test nicht bestehen, es gibt also viele Möglichkeiten, intelligent zu sein, ohne den Turing-Test zu bestehen. Die entscheidende Aussage ist die Umkehrung: Um den Test bestehen zu können, muss man intelligent sein.

MARTIN FORD: Glauben Sie, dass Deep Learning in Kombination mit Ihrem hierarchischen Ansatz der richtige Weg ist, oder denken Sie, dass die Notwendigkeit für weitere Paradigmenwechsel besteht, um eine AGI zu erreichen?

RAY KURZWEIL: Nein, ich denke, dass Menschen diesen hierarchischen Ansatz nutzen. Jedes einzelne der Module kann lernen, und ich weise in meinem Buch sogar darauf hin, dass in den Modulen des Gehirns kein Deep Learning stattfindet, sondern etwas, das einer Markow-Kette ähnelt, aber es ist tatsächlich besser, Deep Learning zu verwenden.

Unsere Systeme bei Google verwenden Deep Learning, um Vektoren zu erzeugen, die das Muster in den Modulen repräsentieren, und es gibt eine Hierarchie, die über das Deep-Learning-Paradigma hinaus geht. Ich denke, das ist für eine AGI schon ausreichend. Meiner Ansicht nach nutzt das Gehirn den hierarchischen Ansatz und die Projekte, die versuchen, ein Reverse Engineering des Gehirns vorzunehmen, haben mittlerweile viele Hinweise darauf geliefert.

Es gibt die These, dass menschliche Gehirne ein regelbasiertes System nutzen, kein konnektionistisches. In diesem Zusammenhang wird darauf hingewiesen, dass Menschen in der Lage sind, sehr genaue Unterscheidungen vorzunehmen, und dass wir logisch denken können. Ein entscheidender Punkt ist, dass der Konnektionismus einen regelbasierten Ansatz emulieren kann. Ein konnektionistisches System kann sich in einer bestimmten Situation seiner Beurteilung so sicher sein, dass es sich wie ein regelbasiertes System verhält, ist aber dennoch in der Lage, mit seltenen Ausnahmen und den Nuancen der scheinbaren Regeln zurechtzukommen.

Ein regelbasiertes System kann jedoch kein konnektionistisches System emulieren, die Umkehrung der Aussage trifft also nicht zu. Doug Lenats »Cyc« ist ein beeindruckendes Projekt, aber ich glaube, dass es die Beschränkungen eines regelbasierten Systems belegt. Man stößt irgendwann an eine Komplexitätsgrenze, und die Regeln werden so kompliziert, dass der Versuch, einen Fehler zu beheben, dazu führt, dass drei andere Dinge nicht mehr funktionieren.

MARTIN FORD: Cyc ist das Projekt, bei dem versucht wird, manuell logische Regeln für einen gesunden Menschenverstand zu erstellen?

RAY KURZWEIL: Richtig. Ich kann keine Zahl nennen, aber es gibt eine gewaltige Anzahl von Regeln. Es gab einen Modus, bei dem das System die einem Verhalten zugrunde liegenden Überlegungen ausgeben konnte, und die Erklärungen waren viele Seiten lang und ließen sich nur schwer nachvollziehen. Eine zweifellos beeindruckende Arbeit, aber sie zeigt, dass dies nicht der richtige Ansatz ist, jedenfalls nicht für sich allein genommen, und dass Menschen Intelligenz nicht auf

diese Weise erzielen. Wir durchlaufen keine Kaskaden von Regeln, wir verwenden den hierarchischen selbstorganisierten Ansatz.

Ein hierarchischer, aber konnektionistischer Ansatz hat darüber hinaus den Vorteil, dass er sich besser selbst erklärt, weil man sich die Module in der Hierarchie ansehen und erkennen kann, welches Modul welche Entscheidung beeinflusst. Wenn es sich um riesige 100-schichtige neuronale Netze handelt, verhalten sie sich wie eine große Blackbox. Es ist sehr schwierig, die Vorgänge nachzuvollziehen, wenngleich es einige Versuche gab, das zu tun. Ich halte die hierarchische Erweiterung eines konnektionistischen Ansatzes für gut geeignet, und ich glaube, dass Menschen auf diese Weise denken.

MARTIN FORD: Aber es gibt doch einige Strukturen im menschlichen Gehirn, die schon bei der Geburt vorhanden sind. Babys können beispielsweise Gesichter erkennen.

RAY KURZWEIL: Bestimmte Merkmale können wir selbst erzeugen. In unserem Gehirn gibt es beispielsweise ein Modul in Form einer Gehirnwindung namens Gyrus fusiformis, das spezielle neuronale Verknüpfungen zum Berechnen von Verhältnissen enthält, wie etwa den Abstand zwischen Nasenspitze und Nasenwurzel oder den Augenabstand. Es gibt vielleicht ein Dutzend ziemlich einfacher Merkmale, und Experimente haben Folgendes gezeigt: Wenn man anhand dieser Merkmale eines Bilds ein neues Bild mit den gleichen Merkmalen – den gleichen Verhältnissen – erzeugt, erkennen Menschen auf dem Bild sofort die gleiche Person, obwohl sich andere Details des Bilds ziemlich stark verändert haben. Es gibt verschiedene solcher Merkmalsgeneratoren, einige erzeugen Audioinformationen wie Tonverhältnisse und erkennen Obertöne. Diese Merkmale werden dem hierarchischen konnektionistischen System zugeführt. Es ist also wichtig, diese Merkmalsgeneratoren zu verstehen, und gibt es einige sehr spezielle Merkmale, auf die Babys bei der Erkennung von Gesichtern zurückgreifen.

MARTIN FORD: Ich möchte auf den Weg zu einer AGI zu sprechen kommen und auf den Zeitpunkt, zu dem sie erreicht wird. Ich gehe davon aus, dass AGI und Intelligenz auf menschlichem Niveau gleichwertige Bezeichnungen sind.

RAY KURZWEIL: Es sind Synonyme, allerdings gefällt mir die Bezeichnung AGI nicht, weil ich glaube, dass sie eine implizite Kritik an KI darstellt. Es war schon immer das Ziel der KI, immer größere Intelligenz zu erzielen und schließlich eine Intelligenz auf menschlichem Niveau zu erreichen. Während wir Fortschritte gemacht haben, wurden eigene Fachgebiete abgespalten. Nachdem wir beispielsweise die Erkennung von Buchstaben gemeistert hatten, wurde daraus das eigene Fachgebiet OCR (*Optical Character Recognition*, optische Zeichenerkennung). Bei

der Spracherkennung und der Robotik verhielt es sich gleichermaßen, und man bekam den Eindruck, dass sich das übergreifende Fachgebiet KI nicht mehr auf eine allgemeine Intelligenz konzentriert. Ich war schon immer der Ansicht, dass wir eine allgemeine Intelligenz schrittweise erreichen werden, indem wir ein Problem nach dem anderen lösen.

Ein weiterer Aspekt ist, dass die menschliche Leistung bei einer bestimmten Aufgabe einen sehr großen Bereich umfasst. Wie hoch ist das menschliche Leistungsniveau beim Go-Spielen? Es ist ein breites Spektrum, das von einem Kind, das zum ersten Mal spielt, bis zum Weltmeister reicht. Wir haben schon erlebt, dass ein Computer, sobald er menschliches Niveau auch nur am unteren Ende des Spektrums erreicht, sehr schnell das menschliche Leistungsvermögen übertrifft. Es ist nur wenig länger als ein Jahr her, dass Computer Go nur auf einem niedrigen Niveau spielen konnten, dann aber schnell immer besser wurden. Und vor Kurzem hat AlphaZero nach einigen wenigen Stunden Training AlphaGo überholt und in 100 Spielen 100 Mal geschlagen.

Computer haben sich auch beim Verstehen von Sprache verbessert, allerdings nicht im gleichen Tempo, weil sie noch nicht über genügend Allgemeinwissen verfügen. Heutige Computer können nicht besonders gut mehrstufige Schlussfolgerungen ziehen, wie sie beim Schließen aus mehreren Aussagen gleichzeitig das Allgemeinwissen berücksichtigen müssen. Bei einem Sprachverständnistest für Drittklässler hat ein Computer beispielsweise nicht verstanden, dass ein Junge vermutlich deswegen schmutzige Schuhe hat, weil er durch eine schlammige Pfütze gelaufen ist, und dass seine Mutter sauer wird, wenn er durch die Küche läuft und den Fußboden verschmutzt. Für uns Menschen mag das alles auf der Hand liegen, weil wir so etwas schon erlebt haben, aber für eine KI ist das nicht offensichtlich.

Von der Leistung eines durchschnittlichen Erwachsenen, die Computer bei manchen Sprachverständnistests heute erreichen, zu einer übermenschlichen Leistung zu gelangen, wird nicht so schnell gelingen wie beim Go, denn ich glaube, dass dafür noch weitere grundlegende Probleme gelöst werden müssen. Dessen ungeachtet umfasst die menschliche Leistung ein breites Spektrum, und sobald Computer in diesen Bereich vordringen, werden sie ihn letztendlich überschreiten und übermenschliche Leistungen erbringen. Die Tatsache, dass sie beim Sprachverständnis überhaupt eine Leistung auf dem Niveau eines Erwachsenen zeigen, ist schon sehr beeindruckend, denn ich bin davon überzeugt, dass Sprachverständnis die gesamte Bandbreite menschlicher Intelligenz erfordert, und dass Sprache das gesamte Spektrum menschlicher Mehrdeutigkeit und hierarchischen Denkens umfasst. Um es zusammenzufassen: Die KI macht sehr schnell Fortschritte, und bei alledem werden konnektionistische Ansätze verwendet.

Ich habe kürzlich mit meinem Team diskutiert, was wir zusätzlich zu dem, was wir schon getan haben, unternehmen müssen, um den Turing-Test zu bestehen. Ein gewisses Sprachverständnis haben wir bereits erreicht. Eine entscheidende Anforderung ist das mehrstufige Schlussfolgern – in der Lage zu sein, die Folgen und Auswirkungen von Konzepten zu berücksichtigen, ist von hoher Priorität. Das ist ein Gebiet, auf dem Chatbots regelmäßig scheitern.

Wenn ich Ihnen erzähle, dass ich mir Sorgen um die Leistung meiner Tochter in der Vorschule mache, werden Sie mich kaum ein paar Augenblicke später fragen, ob ich Kinder habe. Chatbots unterlaufen solche Fehler, weil sie nicht alle Folgen berücksichtigen, die das schon Gesagte hat. Wie erwähnt, gibt es auch noch das Problem des Allgemeinwissens, aber wenn wir alle mit der Sprache einhergehenden Auswirkungen verstehen könnten, wäre es möglich, das Allgemeinwissen durch das Lesen und Verstehen der vielen online verfügbaren Dokumente aufzubauen. Ich glaube, wir haben ein paar sehr gute Ideen, wie wir diese Aufgaben erledigen können, und wir haben dafür genügend Zeit.

MARTIN FORD: Sie haben schon seit einiger Zeit keinen Zweifel daran gelassen, dass Sie glauben, dass wir 2029 eine KI auf menschlichem Niveau erreichen. Ist das noch immer der Fall?

RAY KURZWEIL: Ja. In meinem Buch *The Age of Intelligent Machines*, das 1989 erschien, habe ich einen Zeitraum um 2029 genannt, vielleicht ein Jahrzehnt früher oder später. 1999 habe ich *The Age of Spiritual Machines* veröffentlicht und darin die genauere Vorhersage 2029 getroffen. An der Stanford University gab es eine Konferenz von KI-Experten, die sich mit dieser offenbar aufsehenerregenden Prognose befassten. Damals gab es noch keine Wahlmaschinen, wir haben also durch Handzeichen abgestimmt. Die übereinstimmende Meinung war, dass es noch Hunderte von Jahren dauern wird, wobei rund ein Viertel der Gruppe der Ansicht war, dass es niemals geschehen wird.

2006 gab es anlässlich des 50. Jahrestags der vorhin erwähnten Dartmouth-Konferenz 1956 am Dartmouth College ein Treffen, und dort gab es Wahlmaschinen. Der Konsens war, dass es noch etwa 50 Jahre dauern wird. 12 Jahre später, im Jahr 2018, ist die übereinstimmende Meinung, dass es in 20 bis 30 Jahren so weit sein wird, also irgendwann zwischen 2038 und 2048, ich bin also noch immer optimistischer als die meisten KI-Experten, aber nur ein wenig. Meine Vorhersage und der Konsens der KI-Experten nähern sich einander an, aber nicht, weil ich meine Meinung geändert habe. Es gibt eine zunehmende Zahl von Leuten, die meinen, ich sei zu konservativ.

MARTIN FORD: 2029 ist schon in 11 Jahren, also in gar nicht allzu ferner Zukunft. Das wird mir klar, wenn ich an meine 11 Jahre alte Tochter denke.

RAY KURZWEIL: Der Fortschritt nimmt exponentiell zu – denken sie nur an den allein im letzten Jahr erzielten atemberaubenden Fortschritt. Bei selbstfahrenden Autos, beim Verstehen von Sprache, beim Go-Spielen und in vielen anderen Bereichen haben wir dramatische Verbesserungen erreicht. Das Tempo ist rasant, sowohl bei Hardware als auch bei Software. Bei der Hardware findet der exponentielle Fortschritt sogar noch schneller statt als bei der normalen Rechenleistung. In den letzten paar Jahren haben wir die für Deep Learning verfügbare Rechenleistung alle drei Monate verdoppelt, während die normale Rechenleistung sich nur alle 12 Monate verdoppelt.

MARTIN FORD: Einige sehr kluge Köpfe mit profunden KI-Kenntnissen sagen dennoch vorher, dass es noch mindestens 100 Jahre dauert. Glauben Sie, das liegt daran, dass sie in die Falle des linearen Denkens geraten sind?

RAY KURZWEIL: Zum einen denken sie linear, und zum anderen fallen sie dem zum Opfer, was ich als Pessimismus des Entwicklers bezeichne, also sich so sehr auf ein Problem zu konzentrieren und das Gefühl zu haben, dass es wirklich schwierig ist, weil man es noch nicht gelöst hat, und zu extrapolieren, dass man das Problem allein in dem Tempo lösen muss, in dem man daran arbeitet. Die Sache sieht völlig anders aus, wenn man das Tempo des Fortschritts eines Fachgebiets berücksichtigt und das Phänomen betrachtet, dass Ideen sich gegenseitig beeinflussen. Manche Menschen sind einfach nicht in der Lage, den exponentiellen Charakter des Fortschritts zu begreifen, insbesondere dann nicht, wenn es um Informationstechnologie geht.

In der ersten Hälfte des Humangenomprojekts hatte man nach 7 Jahren 1% des Genoms entschlüsselt, und die Kritiker meinten: »Wir haben doch gleich gesagt, dass es nicht klappt. 1 Prozent in 7 Jahren bedeutet, dass es 700 Jahre dauern wird«. Meine Reaktion war: »Wir haben 1 Prozent erledigt – dann sind wir so gut wie fertig. Wir werden jedes Jahr eine Verdopplung erreichen. 1 Prozent muss nur sieben Mal verdoppelt werden, um 100 Prozent zu erreichen«. Und tatsächlich war das Projekt sieben Jahre später abgeschlossen.

Die entscheidende Frage ist: Warum verstehen manche Menschen das problemlos und andere nicht? Das hängt mit Sicherheit nicht von den Fähigkeiten oder von der Intelligenz ab. Manche Menschen, die beruflich gar nichts damit zu tun haben, verstehen es mit Leichtigkeit, weil sie den Fortschritt an ihren Smartphones erleben. Andere, sehr verdienstvolle Menschen, die an der Spitze ihres Fachgebiets stehen, bleiben hartnäckig bei der linearen Denkweise. Ich kann diese Frage also tatsächlich nicht beantworten.

MARTIN FORD: Würden Sie sagen, dass es nicht nur um den exponentiellen Fortschritt bei der Rechenleistung und der Speicherkapazität geht? Es müssen zwei-

felsohne noch einige grundlegende konzeptuelle Durchbrüche erfolgen, um Computer dazu zu bringen, so wie wir Menschen in Echtzeit aus unstrukturierten Daten zu lernen, oder zu schlussfolgern und sich eine Vorstellung von etwas zu bilden.

RAY KURZWEIL: Der Fortschritt bei der Software erfolgt ebenfalls exponentiell, wenngleich es die unvorhersehbaren Aspekte gibt, auf die Sie anspielen. Es gibt eine gegenseitige Bereicherung durch Ideen, die ihrer Natur nach exponentiell ist, und sobald wir bei der Leistung eine bestimmte Stufe erreicht haben, wird es neue Ideen geben, die zur nächsten Stufe führen.

Der wissenschaftliche Beirat der Obama-Regierung hat zu dieser Frage eine Studie durchgeführt, die den Fortschritt von Hardware und Software miteinander vergleicht. Dazu wurde ein Dutzend klassischer naturwissenschaftlich/technischer Aufgaben genutzt, um die erzielten Fortschritte quantitativ daraufhin zu überprüfen, wie viel auf Hardware zurückzuführen ist. In den letzten 10 Jahren betrug das Verhältnis von Hardware zu Software im Allgemeinen etwa 1.000 zu 1, was mit der jährlichen Verdopplung der Rechenleistung in Einklang steht. Wie vielleicht zu erwarten, variierten die Fortschritte bei der Software deutlich, die relative Verbesserung war jedoch in allen Fällen größer als bei der Hardware. Fortschritte sind tendenziell exponentiell. Wenn man bei der Software Fortschritte erzielt, sind sie also auch nicht linear, sondern exponentiell. Insgesamt ergibt sich der Fortschritt aus der Kombination des Fortschritts bei Hardware und Software.

MARTIN FORD: Sie haben für das Jahr 2045 eine weitere Vorhersage getroffen, nämlich dass dann das eintritt, was Sie als Singularität bezeichnen. Ich denke, die meisten Menschen stellen sich darunter eine Intelligenzexplosion oder das Auftreten einer echten Superintelligenz vor. Kann man das so sagen?

RAY KURZWEIL: Bezüglich der Singularität gibt es tatsächlich zwei Denkrichtungen. Das eine Lager glaubt an einen schnellen Takeoff, das andere an einen gemäßigten Takeoff. Ich zähle mich zu der zweiten Gruppe, die der Ansicht ist, dass wir weiterhin exponentielle Fortschritte machen werden, was schon beängstigend genug ist. Die Vorstellung einer Intelligenzexplosion besagt, dass es einen magischen Moment gibt, ab dem ein Computer auf sein eigenes Design zugreifen und es modifizieren kann, um so verbesserte Versionen von sich selbst zu erschaffen, und das in einer sehr schnellen iterativen Schleife wiederholt, sodass seine Intelligenz explosionsartig zunimmt.

Tatsächlich glaube ich, dass wir das schon seit Tausenden von Jahren so machen. Wir sind durch unsere Technologie zweifellos smarter geworden. Unsere Smartphones sind eine Art Erweiterung des Gehirns, und sie machen uns tatsächlich

smarter. Es handelt sich um einen exponentiellen Vorgang. Vor tausend Jahren haben Paradigmenwechsel und Fortschritte Jahrhunderte gedauert, und man hatte den Eindruck, dass überhaupt nichts geschieht. Ihre Großeltern haben das gleiche Leben wie Sie gelebt, und Sie erwarten, dass Ihre Enkel das gleiche Leben wie Sie leben werden. Aber heute sind Veränderungen schon nach einem Jahr oder sogar noch schneller erkennbar. Das Ganze ist exponentiell, und das führt zu einem beschleunigten Fortschritt, der aber in diesem Sinn nicht explosionsartig verläuft.

Ich glaube, dass wir Intelligenz auf menschlichem Niveau im Jahr 2029 erreichen werden, und sie wird nahezu sofort übermenschlich sein. Betrachten Sie als Beispiel Talk to Books. Sie stellen eine Frage, und das System liest 600 Millionen Sätze, 100.000 Bücher, in einer halben Sekunde. Ich persönlich brauche schon ein paar Stunden, um 100.000 Bücher zu lesen!

Ihr Smartphone ist schon heute in der Lage, eine Suche anhand von Schlagwörtern und anderen Verfahren durchzuführen. Googles Suche geht schon über eine Schlagwortsuche hinaus und verfügt über einige semantische Fähigkeiten. Das semantische Verständnis ist noch nicht auf menschlichem Niveau, aber es ist eine Milliarde Mal schneller als menschliches Denken. Und sowohl Software als auch Hardware werden sich in exponentiellem Tempo weiter verbessern.

MARTIN FORD: Sie sind auch dafür bekannt, dass Sie sich darüber Gedanken gemacht haben, Technologie zur Verlängerung des menschlichen Lebens einzusetzen. Können Sie mir dazu mehr erzählen?

RAY KURZWEIL: Eine meiner Thesen lautet, dass wir mit der intelligenten Technologie verschmelzen werden, die wir erschaffen. In dem Szenario, das mir vorschwebt, werden wir medizinische Nanoroboter in unsere Blutbahn injizieren. Eine Anwendung dieser medizinischen Nanoroboter wird die Erweiterung unseres Immunsystems sein. Ich bezeichne das als die dritte Brücke zur radikalen Lebensverlängerung. Die erste Brücke ist das, was wir schon heute tun können, und die zweite Brücke ist die Perfektionierung der Biotechnologie und die Umprogrammierung der Software des Lebens. Der dritte Brücke besteht schließlich darin, dass diese medizinischen Nanoroboter das Immunsystem perfektionieren. Diese Roboter werden auch das Gehirn bevölkern. Virtual Reality und Augmented Reality werden nicht durch an die Gliedmaße unseres Körpers angeschlossene Geräte bereitgestellt, sondern aus dem Inneren des Nervensystems. Die wichtigste Anwendung der medizinischen Nanoroboter wird sein, dass wir die obersten Schichten unseres Neokortex mit einem synthetischen Neokortex in der Cloud verbinden.

MARTIN FORD: Arbeiten Sie bei Google an so etwas?

RAY KURZWEIL: Die Projekte, die ich hier bei Google mit meinem Team durchgeführt habe, verwenden das, was ich als eine grobe Simulation des Neokortex bezeichnen würde. Wir verstehen den Neokortex noch nicht richtig, aber wir versuchen, ihn mit dem vorhandenen Wissen annähernd nachzuahmen. Wir haben schon einige interessante Sprachanwendungen entwickelt, aber Anfang der 2030er-Jahre werden wir über eine sehr gute Simulation des Neokortex verfügen.

So wie ein Smartphone eine Million Mal smarter wird, indem es auf die Cloud zugreift, werden wir das direkt mit unserem Gehirn tun. Durch unsere Smartphones tun wir das schon heute, auch wenn sie kein Teil unserer Körper und Gehirne sind, was eigentlich eine willkürliche Unterscheidung ist. Wir benutzen unsere Finger und unsere Augen und Ohren auch als Erweiterungen unseres Gehirns. Zukünftig werden wir in der Lage sein, das direkt mit unserem Gehirn zu erledigen, aber nicht nur, um Suchen und Sprachübersetzungen direkt mit unserem Gehirn durchzuführen, sondern um die obersten Schichten unseres Neokortex tatsächlich mit einem synthetischen Neokortex in der Cloud zu verbinden.

Vor zwei Millionen Jahren hatten wir noch keine hohe Stirn, aber als wir uns weiterentwickelten, bekamen wir mehr Raum, um einen größeren Neokortex beherbergen zu können. Und was haben wir daraus gemacht? Wir haben ihn in der neokortikalen Hierarchie an erster Stelle platziert. Wir haben als Primaten schon sehr gute Arbeit geleistet, aber jetzt können wir auf noch abstrakterer Ebene darüber nachdenken.

Das war der entscheidende Faktor, der uns dazu befähigte, Technologie, Wissenschaft, Sprache und Musik zu erfinden. Alle uns bekannten menschlichen Kulturen machen Musik, aber nicht eine der Kulturen von Primaten. Die Vergrößerung des für den Neokortex verfügbaren Raums war ein einmaliger Vorgang, denn er konnte nicht weiterwachsen, weil die Geburt, die nach der Vergrößerung des Gehirns vor zwei Millionen Jahren schon schwierig genug war, dadurch unmöglich geworden wäre.

Die Erweiterung unseres Gehirns in den 2030er-Jahren wird kein einmaliger Vorgang sein. Die Leistungsfähigkeit der Cloud verdoppelt sich jedes Jahr und ist nicht durch räumliche Grenzen beschränkt, also wird der nicht-biologische Teil unseres Denkens weiterwachsen. Wir werden unsere Intelligenz bis 2045 um ein Milliardenfaches steigern, und das ist eine so tiefgreifende Umwälzung, dass es schwierig ist, über diesen Ereignishorizont hinaus zu blicken. Wir verwenden hier die aus der Physik entlehene Metapher des Ereignishorizonts, die auch die Schwierigkeit beinhaltet, darüber hinaus zu blicken.

Technologien wie Googles Suche oder Talk to Books sind mindestens eine Milliarde Mal schneller als Menschen. Die Intelligenz bewegt sich noch nicht auf menschli-

chem Niveau, aber sobald wir diesen Punkt erreichen, wird die KI den schon vorhandenen enormen Geschwindigkeitsvorteil und die weiter exponentiell wachsenden Kapazitäten und Fähigkeiten zu ihrem Vorteil nutzen. Das ist die Bedeutung der Singularität: Es ist ein gemäßigter Takeoff, dennoch sind die exponentiellen Entwicklungen beängstigend genug. Wenn man einen Wert 30 Mal verdoppelt, entspricht das einer Multiplikation mit einer Milliarde.

MARTIN FORD: Im Bereich Medizin, insbesondere bezüglich der Langlebigkeit des Menschen, haben Sie sich ausführlich dazu geäußert, welche Auswirkungen die Singularität haben wird, und Sie sind dafür auch kritisiert worden. Ich habe letztes Jahr am MIT einen Vortrag von Ihnen besucht, in dem Sie davon sprachen, dass die meisten Menschen innerhalb der nächsten 10 Jahre in der Lage sein könnten, das zu erreichen, was Sie als »Fluchtgeschwindigkeit der Langlebigkeit« bezeichnen. Darüber hinaus sagten Sie, dass Sie persönlich das vielleicht schon erreicht hätten. Glauben Sie tatsächlich, dass es schon so bald geschehen wird?

RAY KURZWEIL: Wir stehen in der Biotechnologie vor einem Wendepunkt. Die Menschen gehen davon aus, dass die Medizin wie in der Vergangenheit weiter in einem gemächlichen Tempo vor sich hindümpelt. Erfolge der medizinischen Forschung waren im Wesentlichen Zufallstreffer. Arzneimittelhersteller gehen eine Liste mit zigtausend chemischen Verbindungen durch und hoffen, etwas Wirksames zu finden, anstatt zu versuchen, das Ganze zu verstehen und die Software des Lebens systematisch umzuprogrammieren.

Zu behaupten, dass unsere genetischen Prozesse Software sind, ist mehr als nur eine Metapher. Es handelt sich um eine Verkettung von Daten, die sich in einem Zeitalter entwickelt hat, in dem es nicht im Interesse unserer Spezies lag, dass einzelne Menschen sehr lange leben, weil es begrenzte Ressourcen wie etwa Nahrung gab. Wir erleben einen Übergang von einem Zeitalter der Knappheit zu einem Zeitalter des Überflusses.

Die Leistungsfähigkeit der informationsverarbeitenden Prozesse in der Biologie hat sich jedes Jahr verdoppelt. Betrachten Sie beispielsweise die genetische Sequenzierung. Das erste Genom kostete eine Milliarde Dollar, und heute betragen die Kosten um die 1.000 Dollar. Aber unsere Fähigkeit, diesen rohen Objektcode des Lebens nicht nur zu erfassen, sondern auch zu verstehen, zu modellieren, zu simulieren und vor allem neu zu programmieren, verdoppelt sich auch jedes Jahr in ihrer Leistung.

Es gibt jetzt erste klinische Anwendungen – noch ist es ein Rinnsal, das im kommenden Jahrzehnt jedoch zu einer Flut anwachsen wird. Es gibt Hunderte von Anträgen für tiefgehende Eingriffe, die auf eine Zulassung der Regulierungsbe-

hörden warten. Wir sind jetzt in der Lage, ein durch einen Herzanfall angegriffenes Herz zu heilen, also es durch die Injektion von umprogrammierten adulten Stammzellen zu verjüngen. Wir können Organe züchten, die wir erfolgreich Primaten implantiert haben. Die Immuntherapie ist im Wesentlichen eine Umprogrammierung des Immunsystems. Das Immunsystem geht nicht eigenständig gegen Krebs vor, weil es nicht dafür entwickelt wurde, Krankheiten zu bekämpfen, von denen wir tendenziell erst in höherem Alter befallen werden. Tatsächlich können wir es umprogrammieren, sodass es Krebs erkennt und ihn bekämpft. Das ist ein großer Lichtblick in der Krebstherapie, und es gibt bemerkenswerte Studien, in denen sich der Zustand praktisch aller Patienten von Stufe 4 (Krebs im Endstadium) verbesserte und die Rückbildung einsetzte.

Die Medizin in zehn Jahren wird sich grundlegend von der heutigen unterscheiden. Wenn man sich bemüht, wird man die Fluchtgeschwindigkeit der Langlebigkeit erreichen können, was bedeutet, dass man die Lebenszeit verlängert, also nicht nur die Lebenserwartung zum Zeitpunkt der Geburt, sondern die noch verbleibende Lebenserwartung. Dafür gibt es natürlich keine Garantie, weil man ja immer noch von dem sprichwörtlichen Bus überfahren werden kann und die Lebenserwartung tatsächlich ein kompliziertes statistisches Konzept ist, aber die Lebensuhr wird langsamer ticken als heute. Und ein weiteres Jahrzehnt später werden wir auch in der Lage sein, Alterungsprozesse umzukehren.

MARTIN FORD: Ich möchte über die Nachteile und Risiken der KI sprechen. Ich würde sagen, dass Sie manchmal zu Unrecht dafür kritisiert werden, bei diesen Fragen zu optimistisch zu sein, vielleicht sogar übertrieben optimistisch. Gibt es bei diesen Entwicklungen irgendetwas, über das wir uns Sorgen machen sollten?

RAY KURZWEIL: Ich habe mehr über die Nachteile geschrieben als irgendjemand sonst, und das war Jahrzehnte bevor Stephen Hawking und Elon Musk ihre Bedenken geäußert haben. Auf die Nachteile von GNR (Genetik, Nanotechnologie und Robotik, womit KI gemeint ist) bin ich ausführlich eingegangen – in meinem Buch *The Age of Spiritual Machines*, das 1999 erschien und Bill Joy dazu veranlasste, im Januar 2000 seine berühmte Wired-Titelgeschichte *Why the Future Doesn't Need Us* (Weshalb die Zukunft uns nicht braucht) zu schreiben.

MARTIN FORD: Das beruhte auf einem Zitat von Ted Kaczynski, dem Unabomber¹, richtig?

1 Anm. des Übersetzers: Unabomber (*university and airline bomber*) nannte das FBI Ted Kaczynski, bevor seine Identität bekannt wurde. Er verschickte insgesamt 16 Briefbomben innerhalb der USA, hauptsächlich an Universitätsprofessoren und Vorstandsmitglieder von Fluggesellschaften.

RAY KURZWEIL: Auf einer Seite steht ein Zitat von ihm, das sehr nach einem wohlüberlegten Ausdruck der Besorgnis klingt, und dann blättert man eine Seite weiter und stellt fest, dass es dem Unabomber-Manifest entstammt. In dem Buch erläutere ich die mit GNR verbundenen existenziellen Risiken. In meinem Buch *The Singularity is Near* (deutscher Titel: *Menschheit 2.0: Die Singularität naht*) gehe ich sehr ausführlich auf das Thema Risiken von GNR ein. Kapitel 8 trägt den Titel »The Deeply Intertwined Promise versus Peril of GNR« (Die eng verwobenen Nutzen und Gefahren von GNR).

Ich bin optimistisch, dass wir als Spezies das überstehen werden. Technologie bereitet uns sehr viel mehr Nutzen als Schaden, aber man muss nicht lange suchen, um den Schaden zu erkennen, der angerichtet wurde, beispielsweise all die Zerstörung, die im 20. Jahrhundert stattgefunden hat – obwohl es das bis dahin friedlichste Jahrhundert war, und wir heute sogar in einer noch friedlicheren Zeit leben. Die Welt wird grundsätzlich besser, beispielsweise hat sich in den letzten 200 Jahren die Armut um 95 Prozent verringert, und die Alphabetisierungsrate ist von weniger als 10 Prozent auf mehr als 90 Prozent gestiegen.

Um zu beurteilen, ob die Welt besser oder schlechter geworden ist, fragen sich Menschen: »Wie oft höre ich gute und wie oft schlechte Nachrichten?«, und das ist kein besonders guter Algorithmus. Es gab eine Umfrage unter 24.000 Teilnehmern aus 26 Ländern, denen die Frage gestellt wurde: »Hat sich die Armut weltweit in den letzten 20 Jahren verringert oder vergrößert?«. 87 Prozent antworteten fälschlicherweise, sie habe sich vergrößert. Nur 1 Prozent konnte korrekt angeben, dass sie sich in den letzten 20 Jahren mehr als halbiert hat. Menschen bevorzugen aus evolutionären Gründen schlechte Nachrichten, denn vor 10.000 Jahren war es von großer Bedeutung, schlechten Nachrichten Beachtung zu schenken, beispielsweise dem leisen Rascheln im Gebüsch, das von einem Raubtier stammen könnte. Dem Aufmerksamkeit zu schenken, war wichtiger als festzustellen, dass die Gemüseernte ein halbes Prozent besser ist als im letzten Jahr, und wir geben noch immer schlechten Nachrichten den Vorzug.

MARTIN FORD: Aber es gibt doch einen drastischen Unterschied zwischen solchen Gefahren und existenziellen Risiken.

RAY KURZWEIL: Wir sind auch mit existenziellen Risiken durch Informationstechnologie ganz gut zurechtgekommen. Vor 40 Jahren gab es eine Gruppe visionärer Wissenschaftler, die sowohl den Nutzen als auch die Gefahren der Biotechnologie erkannten, die damals noch in ferner Zukunft lag, und sie veranstalteten die erste Asilomar-Konferenz zur Sicherheit in der Molekularbiologie. Die ethischen Standards und die Strategien wurden regelmäßig aktualisiert, und das hat sehr gut funktioniert. Die Anzahl der Personen, die durch absichtlichen oder unabsichtli-

chen Missbrauch der Biotechnologie oder durch Unfälle zu Schaden gekommen sind, liegt praktisch bei null. Wir kommen allmählich in den Genuss der Vorteile, auf die ich angespielt hatte und die im kommenden Jahrzehnt zu einer Flut anwachsen werden.

Das ist ein Erfolg für diesen Ansatz, umfassende ethische Standards und technische Strategien zur Bewahrung der Sicherheit festzulegen, und vieles davon ist inzwischen gesetzlich geregelt. Das soll aber nicht heißen, dass wir Gefahren durch Biotechnologie von unserer Liste der Bedenken streichen können. Wir entwickeln immer leistungsfähigere Technologien, wie etwa CRISPR, und wir müssen die Standards dementsprechend immer wieder anpassen.

Wir haben vor rund 18 Monaten erstmals an einer Asilomar-Konferenz teilgenommen und eine Reihe ethischer Standards vorgestellt. Ich denke, sie müssen noch weiterentwickelt werden, aber der Ansatz als solcher kann funktionieren. Wir müssen ihm hohe Priorität einräumen.

MARTIN FORD: Eine Befürchtung, die gerade sehr viel Aufmerksamkeit findet, ist das sogenannte Kontrollproblem, bei dem eine Superintelligenz Ziele verfolgen könnte, die nicht damit in Einklang stehen, was für die Menschheit am besten ist. Nehmen Sie das ernst, und sollte daran gearbeitet werden?

RAY KURZWEIL: Die Ziele von Menschen stehen auch nicht alle miteinander in Einklang, und das ist eigentlich der entscheidende Punkt. Es ist ein Missverständnis, KI als eine eigene Zivilisation zu betrachten, so als ob es sich um eine Invasion von Marsmenschen handeln würde. Wir entwickeln Tools, um unseren Einflussbereich zu erweitern. Vor 10.000 Jahren konnten wir die Früchte nicht erreichen, die an den oberen Zweigen hingen, also entwickelten wir Werkzeuge, die unsere Reichweite vergrößerten. Wir können Wolkenkratzer nicht mit bloßen Händen errichten, also setzen wir statt unserer Muskeln Maschinen ein. Und ein Kind in Afrika, das ein Smartphone besitzt, kann durch ein paar Tastendrucke auf das gesamte Wissen der Menschheit zugreifen.

Das ist die Rolle, die Technologie einnimmt: Sie ermöglicht es uns, über unsere Grenzen hinauszuwachsen, und dafür nutzen wir jetzt und zukünftig KI. Es geht anders als in vielen dystopischen Science-Fiction-Filmen nicht um Menschheit gegen KI. Wir werden vielmehr mit ihr verschmelzen. Damit haben wir schon angefangen. Dass ein Smartphone kein physischer Bestandteil des Körpers und des Gehirns ist, macht im Grunde keinen Unterschied, weil es ebenso gut so sein könnte. Ohne Smartphone gehen wir nicht aus dem Haus, weil wir uns sonst irgendwie unvollständig fühlen. Ohne Smartphone kann heutzutage niemand mehr seine Arbeit erledigen, seine Ausbildung erhalten oder Beziehungen pflegen, und diese Bindung wird immer enger.

Ich bin zum MIT gegangen, weil es 1965 so fortschrittlich war, dass es über Computer verfügte. Wenn ich einen benutzen wollte, musste ich mit dem Fahrrad quer über den Campus fahren und meinen Ausweis vorzeigen, um das Gebäude betreten zu dürfen. Und heute, ein halbes Jahrhundert später, haben wir Computer in der Hosentasche und verwenden sie ständig. Sie sind zu einem Bestandteil unseres Lebens geworden und werden letztendlich zum Bestandteil unserer Körper und Gehirne werden.

Die Konflikte und Kriege der letzten Jahrtausende wurden stets durch Meinungsverschiedenheiten unter Menschen ausgelöst. Ich glaube, dass Technologie tendenziell zu mehr Eintracht, mehr Frieden und mehr Demokratisierung führt. Das Aufkommen der Demokratisierung ist eng mit Verbesserungen der Kommunikationsmittel verknüpft. Vor zwei Jahrhunderten gab es weltweit nur eine Demokratie. Vor einem Jahrhundert gab es ein halbes Dutzend. Heute sind 123 der anerkannten 192 Staaten Demokratien, das sind 64 Prozent. Die Welt ist keine perfekte Demokratie, aber diese Staatsform hat sich als Standard etabliert. Wir leben in der friedlichsten Zeit in der Geschichte der Menschheit, und alle Aspekte des Lebens verbessern sich, und das verdanken wir den Auswirkungen der Technologie, die zunehmend intelligenter wird und die zu uns einfach dazugehört.

Die heutigen Konflikte zwischen Bevölkerungsgruppen werden durch die jeweils verfügbaren Technologien verstärkt. Das wird auch weiterhin der Fall sein. Allerdings glaube ich, hier sollte berücksichtigt werden, dass durch bessere Kommunikationstechnologien auch unsere kurzfristige Empathie zum Tragen kommt. Wir zeigen von Natur aus Empathie für kleine Bevölkerungsgruppen, und das wird durch unsere Fähigkeit, tatsächlich mitzuerleben, was Menschen am anderen Ende der Welt widerfährt, weiter verstärkt. Ich glaube, das ist der entscheidende Punkt. Wir müssen nach wie vor mit menschlichen Beziehungen umgehen, wenn wir unsere persönlichen Fähigkeiten durch Technologie erweitern.

MARTIN FORD: Lassen Sie uns über das Potenzial für Umwälzungen der Wirtschaft und des Arbeitsmarkts sprechen. Ich persönlich glaube, dass es ein großes Potenzial für den Verlust oder die Entqualifizierung von Arbeitsplätzen und eine stark wachsende Ungleichheit gibt. Tatsächlich glaube ich, dass es Umwälzungen in der Größenordnung einer neuen industriellen Revolution geben könnte.

RAY KURZWEIL: Dann frage ich Sie: Wie ist die letzte industrielle Revolution ausgegangen? Vor 200 Jahren gab es die Webergilde, die seit Hunderten von Jahren von einer Generation zur nächsten fortgeführt wurde. Das Geschäftsmodell der Weber wurde auf den Kopf gestellt und zerschlagen, als plötzlich überall Garn und Kleidung produzierende Webmaschinen auftauchten, die ihre Existenzgrundlage

bedrohten. Es wurde vorhergesagt, dass weitere Maschinen kommen, und dass die meisten Menschen ihre Jobs verlieren werden. Beschäftigung würde es nur noch für eine Elite geben. Ein Teil dieser Vorhersage erwies sich als richtig – es wurden weitere Maschinen eingeführt, und viele Fähigkeiten und Jobs wurden nicht mehr benötigt. Die Beschäftigungsrate allerdings sank nicht, sondern stieg an, weil die Gesellschaft wohlhabender wurde.

Wenn ich ein vorausschauender Zukunftsforscher im Jahr 1900 wäre, würde ich darauf hinweisen, dass 38 Prozent der Bevölkerung in der Landwirtschaft und 25 Prozent in Fabriken arbeiten. Ich würde vorhersagen, dass in 115 Jahren, also 2015, nur noch 2 Prozent in der Landwirtschaft und 9 Prozent in Fabriken arbeiten werden. Darauf würden alle mit »Meine Güte, ich werde keine Arbeit haben« reagieren, und ich würde entgegnen: »Keine Sorge, die Arbeitsplätze, die wegfallen, gehören zur niedrigsten Qualifizierungsstufe, und wir werden eine viel größere Zahl an Jobs schaffen, die zur höchsten Qualifizierungsstufe gehören«. Die Leute würden mich fragen: »Oh tatsächlich, was sind das für Jobs?«, und ich würde antworten: »Das weiß ich nicht, wir haben sie noch nicht erfunden«.

Es wird behauptet, dass wir mehr Jobs vernichtet als neu geschaffen hätten, aber das stimmt nicht. 1900 gab es 24 Millionen Arbeitsplätze, und heute sind es 142 Millionen. Als prozentualer Anteil der Bevölkerung entspricht das einem Anstieg von 31 Prozent auf 44 Prozent. Und wie sehen die Jobs im Vergleich miteinander aus? Zum einen ist der durchschnittliche Stundenlohn heute 11 Mal so hoch wie 1900 (inflationbereinigt). Wir haben das Arbeitsjahr von etwa 3.000 auf 1.800 Stunden verkürzt, aber die Leute verdienen immer noch sechs Mal so viel pro Jahr, und die Jobs sind viel interessanter geworden. Ich glaube, dass wird auch weiterhin so bleiben, auch nach der nächsten industriellen Revolution.

MARTIN FORD: Die eigentliche Frage ist, ob es dieses Mal anders verläuft. Was Sie über die früheren Vorgänge gesagt haben, ist zweifellos richtig, aber den meisten Schätzungen zufolge verrichtet etwa die Hälfte der Arbeitnehmerschaft Tätigkeiten, die grundsätzlich vorhersehbar und ziemlich monoton sind, und all diese Jobs werden potenziell durch Machine Learning bedroht sein. Für die Automatisierung der meisten dieser vorhersehbaren Jobs ist keine KI auf menschlichem Niveau erforderlich.

Es wird vermutlich neue Jobs für Roboterentwickler, Deep-Learning-Forscher und dergleichen geben, aber Sie können realistischerweise nicht erwarten, dass all die Menschen, die jetzt Hamburger braten oder Taxi fahren, diese Stellen besetzen, selbst wenn man annimmt, dass es genügend davon geben wird. Wir sprechen hier von einer Technologie, die Wahrnehmung und Intellekt von Menschen ersetzen kann, und sie wird ein außerordentlich breites Spektrum abdecken.

RAY KURZWEIL: Diese Vorhersagen beruhen auf einer »Wir-gegen-sie«-Mentalität und auf dem, was die Menschen gegen die Maschinen unternehmen können. Wir sind schon smarter geworden, um besser für diese anspruchsvolleren Jobs geeignet zu sein. Noch nicht durch Dinge, die direkt mit unserem Gehirn verbunden sind, sondern durch intelligente Geräte. Niemand kann seine Aufgaben ohne diese Gehirnerweiterungen erledigen, und sie werden unser Gehirn sogar noch umfassender erweitern und noch stärker in unser Leben eingebunden sein.

Zwecks Verbesserung unserer Fähigkeiten haben wir das Bildungswesen ausgebaut. 1870 gab es 68.000 Hochschulstudenten, und heute sind es 15 Millionen. Wenn man sie und alle, die sie betreuen, wie Angestellte der Fakultät und anderes Personal, zusammenrechnet, kommt man auf etwa 20 Prozent der gesamten Erwerbsbevölkerung, die allein an der Hochschulbildung beteiligt ist. Und wir schaffen ständig neue Aufgaben. Vor sechs Jahren gab es den Wirtschaftszweig für Apps noch nicht, der heute einen bedeutenden Anteil der Wirtschaft darstellt. Wir werden uns immer smarter machen.

Bei den Überlegungen zu dieser Frage muss noch etwas ganz anderes berücksichtigt werden, nämlich die vorhin erwähnte These, dass wir vor einem Zeitalter des Überflusses stehen. Beim jährlichen Treffen des Internationalen Währungsfonds habe ich an einer Podiumsdiskussion mit der geschäftsführenden Direktorin Christine Lagarde teilgenommen, bei der sie fragte: »Woher soll hier das Wirtschaftswachstum kommen? Die Digitalwirtschaft bietet all diese fantastischen Dinge, aber man kann Informationstechnologie nicht essen, nicht anziehen und nicht darin wohnen«. Ich habe geantwortet: »All das wird sich ändern. Alle Arten von nominell physischen Produkten werden zu einer Informationstechnologie werden. Wir werden in KI-gesteuerten Gebäuden in vertikalen Hydrokulturen Obst und Gemüse anbauen und die In-vitro-Kultivierung von Muskelgewebe zur Fleischproduktion betreiben. Das wird uns ohne den Einsatz von Chemie und ohne leidende Tiere sehr hochwertige Nahrungsmittel zu geringen Kosten liefern. Die Deflationsrate von Informationstechnologie beträgt 50 Prozent. Die Rechenleistung, die Kommunikationsmöglichkeiten und die genetischen Sequenzierungen, die man vor einem Jahr kaufen konnte, kosten heute nur noch die Hälfte, und diese ausgeprägte Deflation wird sich auf die herkömmlichen physischen Produkte auswirken.«

MARTIN FORD: Sie glauben also, dass Technologien wie 3D-Druck, Fabrikroboter und KI-gestützte Agrarwirtschaft die Kosten für nahezu alle Produkte senken können?

RAY KURZWEIL: Genau. In den 2020er-Jahren werden 3D-Drucker Kleidung drucken. Wir sind aus verschiedenen Gründen noch nicht ganz so weit, aber alles bewegt sich in die richtige Richtung. Wir werden mit 3D-Druckern auch modulare

Bauteile herstellen, die sich in wenigen Tagen zu einem Gebäude zusammensetzen lassen. Die KI-gesteuerten Informationstechnologien werden letztendlich die Herstellung aller physischen Objekte ermöglichen, die wir benötigen.

Die Nutzung von Sonnenenergie wird erleichtert, indem durch Deep-Learning-Verfahren verbesserte Materialien entwickelt werden, wodurch die Kosten für Energiegewinnung und -speicherung schnell sinken. Die Gesamtmenge der gewonnenen Sonnenenergie verdoppelt sich alle zwei Jahre, und bei Windenergie gibt es den gleichen Trend. Die Nutzung von erneuerbaren Energiequellen muss sich nur noch fünf Mal alle zwei Jahre verdoppeln, um 100 Prozent unseres Energiebedarfs zu decken. Wir werden dann ein Tausendstel der Energie nutzen, die Sonne und Wind liefern können.

Christine Lagarde hat gesagt: »Okay, aber es gibt eine Ressource, die niemals zu einer Informationstechnologie werden kann, und das ist Land. Wir leben schon zusammengedrängt.« Ich habe geantwortet: »Das liegt nur daran, dass wir beschlossen haben, so zu leben und Städte gebaut haben, um Beruf und Freizeit verbinden zu können«. Die Menschen haben schon damit angefangen, sich weiträumiger zu verteilen, weil unsere virtuelle Kommunikation stabiler geworden ist. Unternehmen Sie irgendwo auf der Welt eine Reise mit der Bahn, und Sie werden feststellen, dass 95 Prozent der Landfläche nicht genutzt wird.

Wenn wir die 2030er-Jahre erreichen, werden wir in der Lage sein, der gesamten menschlichen Bevölkerung eine sehr hohe Lebensqualität zu bieten, die über das hinausgeht, was wir heutzutage als hohen Lebensstandard betrachten. In einem TED-Talk habe ich die Vorhersage getroffen, dass es ein bedingungsloses Grundeinkommen geben wird, wenn wir die 2030er-Jahre erreichen, das tatsächlich gar nicht besonders hoch sein muss, um einen sehr hohen Lebensstandard zu ermöglichen.

MARTIN FORD: Sie sind also ein Befürworter des Grundeinkommens? Glauben Sie, dass nicht jeder einen Job finden oder vielleicht gar keinen Job brauchen wird, und dass es andere Einkommensquellen wie ein bedingungsloses Grundeinkommen geben wird?

RAY KURZWEIL: Wir nehmen immer an, dass ein Job ein Weg zum Glück ist. Ich denke, dass hier Bedeutung und Zweck entscheidend sind. Die Menschen werden weiter miteinander wetteifern, um sich einbringen zu können und um Erfüllung zu finden.

MARTIN FORD: Aber man wird nicht unbedingt dafür bezahlt, dass man das tut, was man für sinnvoll hält?

RAY KURZWEIL: Ich denke, wir werden das Wirtschaftsmodell verändern und dass wir damit bereits angefangen haben. Student in einem College zu sein, wird als erstrebenswert betrachtet. Das ist zwar keine Arbeit, wird aber dennoch als lohnende Tätigkeit betrachtet. Man wird kein Arbeitseinkommen benötigen, um in Bezug auf die physischen Bedürfnisse des Lebens einen sehr hohen Lebensstandard zu erreichen, und wir werden uns in Maslows Bedürfnishierarchie weiterhin nach oben bewegen. Das haben wir schon getan, Sie brauchen nur heute mit dem Jahr 1900 zu vergleichen.

MARTIN FORD: Was denken Sie über den vermeintlichen Wettbewerb mit China, eine fortgeschrittene KI zu erreichen? China ist im Vorteil, weil es dort weniger Regulierung beispielsweise hinsichtlich der Privatsphäre gibt. Zudem ist die Bevölkerung erheblich größer, wodurch mehr Daten erzeugt werden, und es bedeutet auch, dass China potenziell mehr junge Turings oder von Neumanns hervorbringen könnte.

RAY KURZWEIL: Ich glaube nicht, dass es sich hier um ein Nullsummenspiel handelt. Wenn einem Forscher in China ein Durchbruch bei der Sonnenenergie oder beim Deep Learning gelingt, ist das für uns alle von Nutzen. China veröffentlicht, wie auch die USA, sehr viel, und die Informationen werden tatsächlich ziemlich freizügig geteilt. Denken Sie nur an Google, das TensorFlow, ein Deep-Learning-Framework, zu einem Public-Domain-Projekt gemacht hat. Und die Technologie, die Talk to Books und Smart Reply zugrunde liegt, haben wir als Open Source veröffentlicht.

Ich persönlich begrüße es, dass China Wert auf wirtschaftliche Entwicklung und Unternehmertum legt. Als ich kürzlich in China war, konnte man den enormen Unternehmertegeist überall spüren. Ich kann China nur empfehlen, sich weiter in Richtung freier Informationsaustausch zu bewegen. Ich glaube, das ist für diese Art Fortschritt von grundlegender Bedeutung. Auf der ganzen Welt wird das Silicon Valley als motivierendes Vorbild betrachtet. Tatsächlich ist Silicon Valley nur eine Metapher für einen Unternehmertegeist, der Experimentieren feiert und Fehlschläge als Erfahrung bezeichnet. Ich halte das für eine gute Sache und betrachte das Ganze nicht als internationalen Wettbewerb.

MARTIN FORD: Machen Sie sich Sorgen darüber, dass China ein autoritärer Staat ist, und dass es für diese Technologien beispielsweise auch militärische Anwendungen gibt? Unternehmen wie Google und natürlich DeepMind in London haben sehr deutlich gemacht, dass sie nicht wollen, dass ihre Technologie für etwas eingesetzt wird, das auch nur im Entferntesten mit Militär zu tun hat. Unternehmen wie Tencent und Baidu in China haben hier eigentlich keine Wahl.

Sollten wir uns darüber Sorgen machen, dass es hier eine weiterhin bestehende Asymmetrie gibt?

RAY KURZWEIL: Militärische Anwendung ist ein anderes Problem als die autoritäre Regierung. Die autoritäre Haltung der chinesischen Regierung macht mir Sorgen, und ich kann nur dazu raten, mehr Informationsfreiheit zuzulassen und demokratischer zu regieren. Ich denke, damit wäre China und allen anderen wirtschaftlich geholfen.

Ich denke, die politischen, sozialen und philosophischen Fragen bleiben sehr wichtig. Meine Befürchtung ist nicht, dass die KI verrücktspielt und eigenständig handelt, weil ich den Eindruck habe, dass sie umfassend mit uns verbunden ist. Ich mache mir Sorgen um die Zukunft der menschlichen Bevölkerung, die schon eine technologische Zivilisation ist. Wir werden damit fortfahren, uns selbst durch Technologie voranzubringen, deshalb ist es am vernünftigsten, die Sicherheit der KI dadurch zu gewährleisten, dass wir darauf achten, wie wir Menschen über uns selbst herrschen.

Ray Kurzweil ist weltweit als einer der führenden Vordenker und Zukunftsforscher anerkannt. Er erhielt seinen Ingenieursabschluss vom MIT, wo er von Marvin Minsky betreut wurde, einem der Urväter des Fachgebiets KI. Er ist Haupterfinder des ersten CCD-Flachbettscanners, der ersten optischen Zeichenerkennung für mehrere Schriften, des ersten Vorlesegeräts für Blinde, des ersten Sprachsynthesizers, des ersten Musiksynthesizers, der einen Flügel und andere Orchesterinstrumente nachbilden konnte, und der ersten kommerziell vermarkteten Spracherkennung.

Neben vielen anderen Auszeichnungen erhielt Kurzweil einen Grammy für außergewöhnliche Leistungen auf dem Gebiet der Musiktechnologie. Ihm wurde die National Medal of Technology (die höchste Auszeichnung im Bereich Technology) verliehen und er wurde in die National Inventors Hall of Fame aufgenommen. Er trägt 21 Ehrendokortitel und wurde von drei US-Präsidenten ausgezeichnet.

Kurzweil hat fünf Bestseller verfasst, unter anderem *The Singularity is Near* (2005, deutscher Titel: *Menschheit 2.0: Die Singularität naht*) und *How to Create a Mind* (2012). Er ist Mitbegründer und Rektor der Singularity University und Leiter Engineering bei Google, wo er ein Forschungsteam leitet, das an Maschinenintelligenz und dem Verstehen natürlicher Sprache arbeitet.

Kurzweil ist durch seine Arbeiten über exponentiellen Fortschritt in der Technologie bekannt, die er durch das »Gesetz der sich beschleunigenden Gewinne« (Law of Accelerating Returns) formalisiert hat. Im Verlauf der Jahrzehnte hat er eine Vielzahl bedeutender Vorhersagen getroffen, die sich oft als richtig erwiesen.

Anfang 2019 erschien Kurzweils erster Roman, *Danielle, Chronicles of a Superheroine*. Ein weiteres Buch von Kurzweil, *The Singularity is Nearer*, wird voraussichtlich Ende 2019 erscheinen.



“ Wir haben nichts vorzuweisen, was einem Insekt auch nur nahekommt, deshalb mache ich mir keine Sorgen, dass es in absehbarer Zeit eine Superintelligenz geben wird.

RODNEY BROOKS

VORSTAND RETHINK ROBOTICS

Rodney Brooks ist weltweit anerkannter Robotikingenieur. Brooks ist Mitbegründer von iRobot Corporation, dem Branchenführer sowohl bei Endkundenprodukten (bekannt ist vor allem der Staubsaugerroboter Roomba) als auch bei militärischen Robotern, die beispielsweise im Irakkrieg zum Entschärfen von Bomben eingesetzt wurden (die militärische Abteilung von iRobot wurde 2016 veräußert). 2008 war Brooks Mitbegründer eines neuen Unternehmens, Rethink Robotics, das sich auf flexible Fertigungsroboter konzentriert, die mit menschlichen Arbeitern auf sichere Weise zusammenarbeiten können.

MARTIN FORD: Während Sie am MIT waren, haben Sie das Unternehmen iRobot gegründet, das jetzt einer der weltweit größten Hersteller von kommerziellen Robotern ist. Wie kam es dazu?

RODNEY BROOKS: Ich habe iRobot zusammen mit Colin Angle und Helen Greiner gegründet, das war 1990. Bei iRobot hatten wir eine Serie von 14 erfolglosen kommerziellen Modellen, erfolgreiche gab es erst 2002. Wir haben damals in einem Jahr zwei erfolgreiche Modelle entwickelt. Das eine war ein Roboter für das Militär, der in Afghanistan eingesetzt wurde, um in Höhlen zu fahren und Aufklärung zu betreiben. Während des Afghanistan- und des Irak-Konflikts wurden rund 6.500 dieser Modelle dazu eingesetzt, an Straßenrändern versteckte Sprengsätze zu beseitigen.

Zur gleichen Zeit, also 2002, brachten wir den *Roomba* auf den Markt, einen Staubsaugerroboter. 2017 konnte das Unternehmen einen Jahresumsatz von 884 Millionen Dollar erwirtschaften. Seit der Gründung wurden mehr als 20 Millionen Geräte ausgeliefert. Man kann also sagen, dass der *Roomba* hinsichtlich der Verkaufszahlen der erfolgreichste Roboter aller Zeiten ist. Das Gerät basiert auf einer Intelligenz auf Insekten-Niveau, die ich um 1984 am MIT entwickelt hatte.

Als ich das MIT 2010 verließ, habe ich mich von dort vollkommen zurückgezogen und ein neues Unternehmen gegründet, Rethink Robotics, wo wir Roboter entwickeln, die weltweit in Fabriken eingesetzt werden. Wir haben bis heute Tausende davon ausgeliefert. Sie unterscheiden sich von normalen Industrierobotern dadurch, dass man sich völlig gefahrlos in ihrer Nähe aufhalten kann. Sie müssen nicht durch ein Gitter abgeschirmt werden, und man kann ihnen zeigen, welche Aufgabe sie erledigen sollen.

Bei der neuesten Version der Software, die wir verwenden, *Intera 5*, schreibt der Roboter tatsächlich ein Programm, wenn man ihm zeigt, was er tun soll. Dabei handelt es sich um ein grafisches Programm, das einen Verhaltensbaum repräsentiert, den man bei Bedarf modifizieren kann, aber das ist nicht unbedingt erforderlich. Seitdem es diese Roboter gibt, wollten andere Unternehmen in der Lage sein, auf die Robotersteuerung zuzugreifen und genau abzustimmen, was der Roboter macht, nachdem ihm gezeigt worden ist, was er tun soll, ohne dass man die zugrunde liegenden Repräsentationen kennen muss. Die Roboter verwenden Kraftsensoren und Kameras und können in normalen Umgebungen, in denen sich Menschen aufhalten, betrieben werden – rund um die Uhr, sieben Tage die Woche, 365 Tage im Jahr, überall auf der Welt. Ich bin mir sicher, dass es die fortschrittlichsten KI-Roboter sind, die in Serie gefertigt werden.

MARTIN FORD: Wie sind Sie zur Speerspitze der Robotik und der KI gelangt? Wo fängt Ihre Geschichte an?

RODNEY BROOKS: Ich bin in Adelaide in Südaustralien aufgewachsen. 1962 entdeckte meine Mutter zwei Exemplare der amerikanischen *How and Why Wonder Books* (dem US-Pendant zu der Sachbuchreihe *Was ist was*) mit den Titeln *Electricity* (Elektrizität) und *Robots and Electronic Brains* (Roboter und Elektronengehirne). Ich wurde regelrecht süchtig danach. Den Rest meiner Kindheit habe ich damit verbracht, das, was ich aus den Büchern erfahren hatte, genauer zu erkunden, und ich habe versucht, intelligente Computer zu basteln und schließlich Roboter zu entwickeln.

Ich habe dann einen Abschluss im Grundstudium in Mathematik gemacht und angefangen, eine Doktorarbeit über Künstliche Intelligenz zu schreiben. Allerdings musste ich feststellen, dass es in Australien gar keine Fakultät für Informatik und keine KI-Forscher gab. Also bewarb ich mich an den Instituten, von denen ich wusste, dass sie KI-Forschung betreiben: MIT (Massachusetts Institute of Technology), Carnegie Mellon (Pittsburgh, USA) und Stanford University. Das MIT sagte ab, aber von Carnegie Mellon und Stanford erhielt ich 1977 Zusagen. Ich habe mich für Stanford entschieden, weil es näher an Australien liegt.

Meine Doktorarbeit in Stanford habe ich bei Tom Binford über Computer Vision geschrieben. Anschließend war ich als Postdoc erst an der Carnegie Mellon University und später am MIT tätig. 1983 schließlich ging ich wieder nach Stanford und wurde dort Fakultätsmitglied mit Aussicht auf eine Festanstellung. 1984 ging ich zurück ans MIT, wo ich 26 Jahre bleiben sollte.

Als Postdoc am MIT hatte ich angefangen, mehr an intelligenten Robotern zu arbeiten. Als ich dann 1984 ans MIT zurückkehrte, fiel mir auf, wie wenige Fortschritte wir bei der Modellierung der Wahrnehmung von Robotern gemacht hatten. Ich habe mich von Insekten inspirieren lassen, die mit Hunderttausend Neuronen um Größenordnungen besser waren als jeder Roboter. Ich habe dann versucht, Intelligenz anhand von Insekten zu modellieren. Damit habe ich mich die nächsten paar Jahre befasst.

Dann habe ich das von Marvin Minsky gegründete KI-Lab am MIT geleitet. Das wurde später mit dem Informatik-Lab vereint und wurde zum CSAIL, dem Computer Science and Artificial Intelligence Lab, das auch heute noch das größte Lab am MIT ist.

MARTIN FORD: Wenn Sie auf Ihre Laufbahn mit Robotern und KI zurückblicken, was war der Höhepunkt?

RODNEY BROOKS: Ich bin besonders stolz auf das, was kurz nach dem Erdbeben in Japan mit dem nachfolgendem Tsunami, der die Nuklearkatastrophe von Fukushima auslöste, im März 2011 geschehen ist. Etwa eine Woche nach dem

Vorfall hörten wir davon, dass die japanischen Behörden große Schwierigkeiten hatten, mit Robotern in das Kraftwerk vorzudringen, um herauszufinden, was dort vor sich geht. Ich war damals noch bei iRobot, und wir lieferten innerhalb von 48 Stunden sechs Roboter nach Fukushima und wiesen das Technikerteam der Betreiberfirma in die Verwendung der Roboter ein. Sie wussten es sehr zu schätzen, dass unsere Roboter das Herunterfahren des Reaktors ermöglichten, weil sie Dinge erledigen konnten, die sie allein nicht hätten vollbringen können.

MARTIN FORD: An diese Geschichte aus Japan kann ich mich erinnern. Ich war etwas überrascht, weil Japan eigentlich in der Robotik als führend gilt, aber dennoch mussten sie sich an Sie wenden, um gut funktionierende Roboter zu bekommen.

RODNEY BROOKS: Ich glaube, hieraus kann man eine Lehre ziehen, nämlich dass die Presse die Roboter als viel leistungsfähiger dargestellt hat, als sie es tatsächlich sind. Alle dachten, dass Japan über enorme Fähigkeiten im Bereich Robotik verfügt, dabei taten sich insbesondere ein oder zwei Autohersteller hervor, die zwar tatsächlich tolle Videos vorweisen konnten, die allerdings nichts mit der Realität zu tun hatten.

Unsere Roboter hatte man bereits seit neun Jahren täglich tausendfach in Kriegsgeländen eingesetzt. Sie waren nicht spektakulär, und die KI-Fähigkeiten wurden als praktisch nicht vorhanden abgetan, aber das ist die Realität dessen, was heutzutage real und anwendbar ist. Ich habe den Leuten immer wieder gesagt, dass sie sich täuschen lassen, wenn sie Videos sehen und glauben, dass fantastische Dinge praktisch vor der Tür stehen, oder dass es morgen Massenarbeitslosigkeit gibt, weil Roboter alle unsere Jobs übernehmen.

Bei Rethink Robotics sagen wir: Wenn es vor 30 Jahren noch keinen Prototypen gab, dann ist es zu früh, zu glauben, dass wir jetzt ein marktreifes Produkt daraus machen können. So lange dauert es, aus einem Prototypen ein serienreifes Produkt zu machen. Das trifft mit Sicherheit auf autonomes Fahren zu – selbstfahrende Autos sind in aller Munde. Aber die Menschen übersehen, dass das erste selbstfahrende Auto schon 1987 in der Nähe von München auf einer Autobahn 20 Kilometer weit mit einer Geschwindigkeit von 100 Stundenkilometern gefahren ist. 1995 fuhr das erste Mal ein Auto ohne Hand am Steuer und ohne Bedienung der Pedale quer durch die USA (*No Hands across America*). Wird es also morgen serienmäßig produzierte selbstfahrende Autos geben? Nein. Es dauert sehr, sehr lange so etwas zu entwickeln, und ich denke, die Menschen überschätzen noch immer, wie schnell diese neue Technologie zum Einsatz kommen wird.

MARTIN FORD: Für mich hört sich das so an, als ob Sie Ray Kurzweils »Gesetz der sich beschleunigenden Gewinne« keinen Glauben schenken, also der Vorstel-

lung, dass sich alles immer schneller weiterentwickelt. Ich habe den Eindruck, dass Sie glauben, dass sich alles in gleichmäßigem Tempo entwickelt?

RODNEY BROOKS: Deep Learning hat Fantastisches geleistet, und Menschen, die dem Fachgebiet fremd sind, sehen sich das an und sagen »Wow!«. Wir sind an exponentielles Wachstum gewöhnt, weil es das Moore'sche Gesetz gab, aber das gilt allmählich nicht mehr, weil es nicht länger möglich ist, die Integrationsdichte zu verdoppeln. Das wird zu einer Renaissance der Computerarchitektur führen. Vor 50 Jahren konnte man es sich nicht erlauben, etwas Außergewöhnliches auszuprobieren, weil man nur wegen des Moore'schen Gesetzes sofort überholt wurde. Jetzt blüht die Computerarchitektur wieder auf, und ich glaube, sie wird ein goldenes Zeitalter erleben, weil das Moore'sche Gesetz nicht mehr gilt. Das wird Ray Kurzweil und Leute, die exponentielles Wachstum erlebt haben und glauben, dass alles exponentiell wächst, nicht freuen.

Manche Dinge entwickeln sich exponentiell, aber nicht alle. In Gordon Moores Arbeit *The Future of Integrated Electronics* aus dem Jahr 1965, aus der das Moore'sche Gesetz stammt, geht es im letzten Teil darum, wofür das Gesetz nicht gilt. Moore schreibt, es sei beispielsweise nicht auf Energiespeicher anwendbar, weil es dabei nicht um die Informationsabstraktion von Nullen und Einsen geht, sondern um physische Eigenschaften.

Betrachten Sie als Beispiel Greentech (umweltfreundliche Technologie). Vor einem Jahrzehnt haben sich Risikokapitalgeber im Silicon Valley die Finger verbrannt, weil sie dachten, Moores Gesetz ist allgemeingültig und auch auf Greentech anwendbar. Aber so funktioniert es nicht. Greentech beruht auf Masse und auf Energie; man kann das nicht einfach physisch halbieren und erwarten, dass der gleiche Informationsgehalt vorhanden ist.

Ich komme zurück auf Deep Learning. Weil es dann und wann Verbesserungen gab, glauben die Leute, dass es immer so weiter geht. Der für Deep Learning grundlegende Backpropagation-Algorithmus wurde in den 1980er-Jahren entwickelt. Und die damaligen Entwickler haben es nach 30 Jahren Arbeit geschafft, dass er fantastisch funktioniert. In den 1980er- und 1990er-Jahren hatte man Backpropagation schon weitgehend beschrieben, weil es keinen Fortschritt gab, aber gleichzeitig wurden auch Hundert andere Dinge beschrieben. Niemand hat vorhergesagt, welches der Hundert Dinge wiederaufleben würde. Es sollte eben Backpropagation sein, zusammen mit einigen anderen Dingen, wie Clamping (das Festsetzen von einzelnen Neuronenwerten), mehr Schichten und sehr viel mehr Berechnungen, und es stellte sich als großartig heraus. Man hätte niemals voraussagen können, dass Backpropagation und nicht eins der anderen 99 Dinge das Rennen macht. Das war in keiner Weise zwangsläufig.

Deep Learning hat großen Erfolg gehabt und wird weiter Erfolg haben, aber es wird nicht immer so weitergehen, dass die Erfolge größer werden. Es gibt Grenzen. Ray Kurzweil wird sein Bewusstsein in absehbarer Zeit nicht in die Cloud hochladen. So funktionieren biologische Systeme nicht. Deep Learning wird einiges leisten, aber biologische Systeme beruhen nicht nur auf einem Algorithmus, sondern auf Hunderten. Wir werden Hunderte weitere Algorithmen benötigen, bevor wir einen solchen Fortschritt erzielen, und wir können nicht vorhersagen, wann es soweit sein wird. Jedes Mal, wenn ich Kurzweil treffe, erinnere ich ihn daran, dass er sterben wird.

MARTIN FORD: Das ist gemein.

RODNEY BROOKS: Ich werde auch sterben. Daran habe ich keinen Zweifel, aber er mag es nicht, wenn man darauf hinweist, weil er einer dieser Anhänger der »Religion der Technologie« ist. Davon gibt es verschiedene Varianten. Es gibt diese Unternehmen, die das Leben verlängern wollen, die von irgendwelchen Milliardären aus dem Silicon Valley gegründet wurden. Dann gibt es die Leute, die ihr Bewusstsein auf einem Computer speichern wollen, so wie Ray Kurzweil. Ich glaube, dass wir noch einige Jahrhunderte lang sterblich sein werden.

MARTIN FORD: Dem stimme ich tendenziell zu. Sie haben selbstfahrende Autos erwähnt, deshalb möchte ich Sie fragen, wie bald Sie das kommen sehen. Google betreibt in Arizona angeblich schon Autos ohne Fahrer.

RODNEY BROOKS: Dazu sind mir keine Einzelheiten bekannt, aber es hat sehr viel länger gedauert, als alle angenommen haben. Mountain View (in Kalifornien) und Phoenix (in Arizona) sind Städte, die sich zu sehr von den übrigen USA unterscheiden. Vielleicht werden wir dort ein paar Demos sehen, aber es wird noch einige Jahre dauern, bis es einen automatischen Fahrdienst gibt, der auch nur im Entferntesten profitabel ist. Mit profitabel meine ich, Geld ungefähr genauso schnell zu verdienen, wie Uber Geld verliert – das waren im letzten Jahr 4,5 Milliarden Dollar.

MARTIN FORD: Weil Uber mit jeder Fahrt Geld verliert, ist die allgemeine Meinung, dass dieses Geschäftsmodell nicht nachhaltig ist, wenn Uber nicht auf selbstfahrende Autos umstellt.

RODNEY BROOKS: Ich habe erst heute Morgen einen Bericht gesehen, in dem es hieß, dass der mittlere Stundenlohn eines Uber-Fahrers 3,57 Dollar beträgt. Sie verlieren also immer noch Geld. Das ist keine große Gewinnspanne. Die kann man loswerden und die teuren Sensoren einsetzen, die für autonomes Fahren erforderlich sind. Wir haben noch nicht einmal herausgefunden, was die prakti-

sche Lösung für selbstfahrende Autos sein soll. Die Autos von Google haben auf dem Dach einen Haufen kostspieliger Sensoren. Tesla hat versucht, nur mit eingebauten Kameras auszukommen und ist damit gescheitert. Es wird zweifelsohne einige beeindruckende Demonstrationen geben, aber die werden frisiert sein. Wir haben das bei den japanischen Robotern schon erlebt, die Demonstrationen waren geschönt, außerordentlich geschönt.

MARTIN FORD: Meinen Sie damit gefälscht?

RODNEY BROOKS: Nicht gefälscht, aber da gibt es eine Menge Sachen hinter den Kulissen, die man nicht sieht. Man zieht Schlüsse aus dem Gezeigten oder generalisiert, aber es ist schlicht und einfach nicht die Wahrheit. Hinter diesen Demonstrationen steht ein Team, und hinter den Demonstrationen selbstfahrender Autos in Phoenix werden noch lange Teams stehen, deshalb sind sie noch weit von der Realität entfernt.

Außerdem unterscheidet sich ein Ort wie Phoenix von Cambridge, Massachusetts, wo ich wohne und wo es viele verstopfte Einbahnstraßen gibt. Das wirft Fragen auf, etwa wo in meiner Nachbarschaft der Fahrdienst mich abholt. Können Sie mitten auf der Straße zusteigen? Oder benutzt das Fahrzeug eine Busspur? Für gewöhnlich wird es beim Einsteigen die Straße blockieren, deswegen muss es schnell gehen, die anderen Autofahrer werden hupen und so weiter. Es wird noch eine Weile dauern, bis vollständig autonome Systeme in solch einem Umfeld betrieben werden können. Deshalb glaube ich, dass es sogar in Orten wie Phoenix noch lange eigens dafür vorgesehene Haltestellen geben wird, denn die Fahrzeuge werden nicht in der Lage sein, sich so einfach in das vorhandene Straßennetz einzuordnen.

Uber nutzt inzwischen eigens dafür vorgesehene Haltestellen für Fahrdienste. Sie verwenden jetzt ein neues System, das sie in San Francisco und Boston ausprobiert und kürzlich auf sechs weitere Städte ausgedehnt haben, bei dem Sie sich in eine Uber-Warteschlange einreihen und zusammen mit anderen Leuten frierend auf ihr Fahrzeug warten. Wir stellen uns immer vor, dass selbstfahrende Autos genauso sein werden wie heutige Autos, nur ohne Fahrer. Aber nein, bei der Art und Weise der Nutzung wird es einen Wandel geben.

Mit dem Aufkommen des Autos haben sich unsere Städte gewandelt, und wir werden für diese Technologie einen weiteren Wandel benötigen. Es wird nicht so wie heute sein, nur dass es eben keine Fahrer mehr gibt. Das wird noch eine ganze Weile dauern, und es spielt keine Rolle, wie begeistert man vom Silicon Valley ist, es wird nicht so schnell geschehen.

MARTIN FORD: Lassen Sie uns spekulieren. Wie lange wird es dauern, bis es so etwas wie die heutigen Fahrdienste von Uber gibt, also ein Produkt, das es ermög-

licht, Sie irgendwo in Manhattan oder San Francisco abzuholen und zu einem Ort Ihrer Wahl zu befördern?

RODNEY BROOKS: Es wird schrittweise umgesetzt werden. Der erste Schritt könnte sein, dass Sie sich zu einer Haltestelle begeben, wo sich das Fahrzeug befindet. Das würde wie die Nutzung eines heutigen Zipcars (ein amerikanischer Carsharing-Anbieter) funktionieren, für die es eigene Parkplätze gibt. Das wird es früher geben als den Dienst, den ich heute von Uber erhalte, bei dem der Wagen vorfährt und direkt vor meinem Zuhause in zweiter Reihe parkt. Irgendwann – ich weiß nicht, ob ich das noch erleben werde – wird es jede Menge selbstfahrender Autos in unseren Städten geben, aber das kann noch Jahrzehnte dauern, und es wird ein Wandel erforderlich sein, aber wir haben noch nicht ganz herausgefunden, wie genau er aussehen wird.

Wenn es überall selbstfahrende Autos gibt, wie werden sie dann aufgetankt oder aufgeladen? Wo werden sie aufgeladen? Und wer schließt sie an die Ladestation an? Einige Start-ups haben schon Überlegungen angestellt, wie die Verwaltung einer Flotte selbstfahrender elektrischer Autos funktionieren könnte. Irgendjemand muss die Wartung und die täglichen Routineaufgaben erledigen. Dafür muss es eine entsprechende Infrastruktur geben, damit selbstfahrende Autos zu einem Massenprodukt werden können, und das wird eine Weile dauern.

MARTIN FORD: Andere Schätzungen, wann es Dienste geben wird, die grob mit Uber vergleichbar sind, liegen eher im Bereich von fünf Jahren. Ich nehme an, Sie halten das für völlig unrealistisch?

RODNEY BROOKS: Ja, das ist völlig unrealistisch. Vielleicht wird es Dienste geben, die bestimmten Aspekten entsprechen, aber nicht etwas Vergleichbares. Sie werden sich von Uber unterscheiden, und zu ihrer Unterstützung wird es eine ganze Reihe neuer Unternehmen und Dienstleistungen geben müssen, die noch nicht vorhanden sind. Fangen wir bei den Grundlagen an. Wie werden Sie in das Fahrzeug gelangen? Wie teilen Sie dem System mit, dass Sie es sich während der Fahrt anders überlegt haben und zu einem anderen Ziel befördert werden möchten? Vermutlich durch Sprache. Amazons Alexa und Googles Home haben gezeigt, wie gut die Spracherkennung inzwischen ist, deshalb denke ich, dass Sprache funktionieren wird.

Und nun zum Regulierungssystem. Welche Anweisungen können sie dem Fahrzeug geben? Welche Anweisungen können Sie ihm geben, wenn Sie keinen Führerschein besitzen? Welche Anweisungen kann ein 12-Jähriger ihm geben, der von seinen Eltern ins Auto gesetzt wurde, damit es ihn zum Fußballtraining bringt? Nimmt das Fahrzeug Sprachbefehle von 12-Jährigen an, oder hört es nicht zu? Es

gibt unglaublich viele praktische und regulatorische Probleme, über die noch nicht gesprochen wurde, die aber gelöst werden müssen. Heutzutage können Sie einen 12-Jährigen in ein Taxi setzen, das ihn irgendwo hinbringt. Mit selbstfahrenden Autos wird das so bald nicht möglich sein.

MARTIN FORD: Kommen wir auf eine Ihre Bemerkungen über Ihre Forschung mit Insekten zurück. Das ist interessant, denn ich dachte eigentlich, dass Insekten sehr gute biologische Roboter sind. Ich weiß, dass Sie selbst kein Forscher mehr sind, aber ich frage mich, wie der Stand der Dinge bei der Entwicklung eines Roboters oder einer Intelligenz mit Fähigkeiten, die ansatzweise denen eines Insekts entsprechen, aussieht, und wie dadurch die Schritte auf dem Weg zu einer Superintelligenz beeinflusst werden.

RODNEY BROOKS: Vereinfacht gesagt, haben wir nichts vorzuweisen, was einem Insekt auch nur nahekommt, deshalb mache ich mir keine Sorgen, dass es in absehbarer Zeit eine Superintelligenz geben wird. Wir können die Lernfähigkeiten von Insekten nicht anhand nur einiger weniger unüberwachter Beispiele reproduzieren. Wir können die Resilienz, die ein Insekt bei der Anpassung an die Welt zeigt, nicht erreichen. Und die wirklich erstaunliche Mechanik von Insekten können wir ganz sicher nicht reproduzieren. Niemand hat etwas vorzuweisen, das die gezielte Handlungsweise eines Insekts erreicht. Wir verfügen über tolle Modelle, die sich etwas ansehen, es klassifizieren und in bestimmten Fällen sogar mit einer Kennzeichnung versehen können, aber das ist selbst beim Vergleich mit der Intelligenz eines Insekts etwas ganz anderes.

MARTIN FORD: Hat die Robotik seit den 90ern und der Zeit, als Sie iRobot gegründet haben, Ihre Erwartungen erfüllt oder sogar übertroffen, oder sind Sie enttäuscht?

RODNEY BROOKS: Als ich 1977 in die Vereinigten Staaten kam, war ich sehr an Robotern interessiert und arbeitete schließlich an Computer Vision. Damals gab es auf der Welt drei mobile Roboter. Einer dieser Roboter befand sich in Stanford, wo Hans Moravec Experimente durchführte und versuchte, den Roboter dazu zu bringen, in sechs Stunden einen knapp 20 Meter langen Raum zu durchqueren. Ein weiterer Roboter war beim Jet Propulsion Laboratory (JPL) der NASA, und der dritte war beim Laboratory for Analysis and Architecture of Systems (LAAS) in Toulouse in Frankreich.

Auf der ganzen Welt gab es drei mobile Roboter. iRobot liefert pro Jahr Millionen mobiler Roboter aus, also bin ich in dieser Hinsicht ziemlich zufrieden. Wir haben große Anstrengungen unternommen und sind wirklich sehr weit gekommen. Dass den Fortschritten in der Robotik nicht mehr Beachtung geschenkt wurde,

liegt nur daran, dass wir im gleichen Zeitraum von zimmergroßen Mainframe-Computern zu Milliarden von Smartphones überall auf der Welt gelangt sind.

MARTIN FORD: So viel zu den Insekten. Ich weiß, dass Sie an Roboterhänden gearbeitet haben. Es gibt einige verblüffende Videos von Roboterhänden verschiedener Teams. Können Sie etwas dazu erzählen, welche Fortschritte es auf diesem Gebiet gibt?

RODNEY BROOKS: Ja, ich wollte die Arbeit an den kommerziellen mobilen Robotern bei iRobot von dem abgrenzen, was ich am MIT mit meinen Studenten unternahm, deshalb wechselte ich bei der Forschung am MIT von Insekten zu Humanoiden, was dazu führte, dass ich anfangs, dort an Roboterarmen zu arbeiten. Die Arbeit macht nur langsam Fortschritte. Bei den Prototypen gibt es einige spannende Entwicklungen. Allerdings konzentrieren sie sich auf eine bestimmte Aufgabe, die sich sehr von unserem eher allgemeinen Vorgehen unterscheidet,

MARTIN FORD: Ist der langsame Fortschritt einem Hardware- oder einem Software-Problem geschuldet? Liegt es an der Mechanik oder nur an der Steuerung?

RODNEY BROOKS: Es ist alles auf einmal. Man muss hier in mehreren Bereichen gleichzeitig Fortschritte erzielen. Sie müssen Fortschritte bei der Mechanik erzielen, beim Material, aus dem Haut besteht, bei den in der Hand eingebetteten Sensoren und bei den Algorithmen, die sie steuern – und all das gleichzeitig. Man kommt nicht voran, indem man nur einen der Faktoren verbessert, die ändern müssen ebenfalls verbessert werden.

Ich gebe Ihnen ein Beispiel, um das zu verdeutlichen. Sie kennen vielleicht diese Greifer aus Kunststoff, die an einem Ende einen Griff besitzen, den Sie zusammendrücken, um am anderen Ende eine kleine Hand zu öffnen und zu schließen. Sie können sie verwenden, um schwer erreichbare Gegenstände zu ergreifen oder eine Lampe in die Fassung zu drehen, an die Sie sonst nicht ganz herankommen.

Sie können mit dieser wirklich primitiven Hand fantastische Manipulationen vornehmen, die weit über das hinausreichen, was Roboter derzeit können, obwohl es ein erstaunlich primitives Stück Plastikspielzeug ist. Ausschlaggebend ist, dass Sie als Mensch die Manipulation vornehmen. In Videos, in denen Forscher eine neue von ihnen entwickelte Roboterhand vorstellen, sieht man oft, dass ein Mensch die Roboterhand bedient und herumbewegt, um eine Aufgabe zu erledigen. Sie hätten für die gleiche Aufgabe auch das Plastikspielzeug verwenden können, denn es ist der Mensch, der die Aufgabe erledigt. Wenn es so einfach wäre, könnten wir den Plastikgreifer am Ende eines Roboterarms montieren und die Aufgaben erledigen lassen. Einem Menschen gelingt das mit dem Spielzeug am Ende des Arms, aber wieso kann ein Roboter das nicht? Irgendetwas wirklich Wichtiges fehlt.

MARTIN FORD: Ich habe Berichte darüber gesehen, dass Deep Learning und Reinforcement Learning dazu verwendet werden, dass Roboter Dinge durch Übung oder sogar nur durch das Ansehen von YouTube-Videos erlernen. Was halten Sie davon?

RODNEY BROOKS: Denken sie daran, dass es Prototypen sind. Bei DeepMind gibt es eine Gruppe, die unsere Roboter verwendet, die kürzlich eine interessante Arbeit über Kraftsensoren veröffentlicht hat. Die Roboter versuchen, Klammern an Objekten zu befestigen. An jedem der Roboter hat ein Team wirklich guter Forscher monatelang akribisch gearbeitet. Die Leistung ist aber weit von der eines Menschen entfernt. Wenn sie einem Menschen zeigen, wie er eine Aufgabe geschickt erledigen kann, ist er dazu sofort in der Lage. Bei Robotern sind wir noch nicht annähernd so weit.

Ich habe kürzlich ein paar IKEA-Möbel zusammengebaut. Manche Leute meinen, das wäre ein ausgezeichnete Test für Roboter. Stellen Sie einen IKEA-Bausatz und die Bauanleitung bereit, und lassen Sie den Roboter das Möbelstück zusammenbauen. Ich musste beim Zusammenbau des Möbelstücks ungefähr 200 Aufgaben erledigen, die Fingerfertigkeit erfordern. Nehmen wir nun an, wir versuchen, das zu reproduzieren, und zwar mit meinen Robotern, die wir zu Tausenden verkaufen, die auf dem neuesten Stand der Technik sind und mehr Sensoren besitzen, als irgendwelche anderen Roboter, die heutzutage angeboten werden. Wenn wir einige Monate in einer sehr eingeschränkten Umgebung daran arbeiten, würden wir vielleicht eine einfache Demonstration von einer dieser 200 Aufgaben hinbekommen, die ich aus dem Stehgreif erledigen konnte. Hier geht wieder die Fantasie mit einem durch, wenn man glaubt, dass ein Roboter schon bald all diese Aufgaben erledigen kann. Die Realität sieht völlig anders aus.

MARTIN FORD: Was ist die Realität? Was wird es in fünf bis zehn Jahren auf dem Gebiet der Robotik und der KI Neues geben? Welche Art von Durchbrüchen sollten wir realistischerweise erwarten?

RODNEY BROOKS: Man kann Durchbrüche niemals erwarten. Ich erwarte, dass Deep Learning in zehn Jahren nicht mehr so viel Aufmerksamkeit geschenkt wird. Es wird etwas Neues geben, das den Fortschritt vorantreibt.

Deep Learning ist für uns eine wunderbare Technologie gewesen. Sie hat die Spracherkennung von Amazon Echo und Google Home ermöglicht, und das war ein gewaltiger Schritt nach vorn. Ich weiß, dass Deep Learning noch weitere Schritte nach vorn ermöglichen wird, aber irgendwann wird es durch etwas anderes ersetzt werden.

MARTIN FORD: Wenn Sie von Deep Learning sprechen, meinen Sie damit neuronale Netze mit Backpropagation?

RODNEY BROOKS: Ja, aber mit vielen Schichten.

MARTIN FORD: Dann werden vielleicht neuronale Netze mit einem anderen Algorithmus die Rolle übernehmen? Oder Bayes-Netze?

RODNEY BROOKS: Das könnte sein, es könnte aber auch etwas ganz anderes sein, das wissen wir nicht. Ich garantiere Ihnen aber, dass es innerhalb der nächsten 10 Jahre etwas Neues geben wird, das für Anwendungen eingesetzt wird und bestimmten anderen Technologien einen Schub gibt. Ich weiß zwar nicht, was das sein wird, aber bei einem Zeitrahmen von 10 Jahren wird das sicherlich passieren.

Es ist unmöglich, zu prognostizieren, was funktionieren wird und weshalb, aber man kann auf vorhersehbare Weise etwas über den Market Pull sagen, den ein paar Megatrends verursachen werden, die gerade stattfinden.

Das Verhältnis von älteren Menschen zu Menschen im erwerbsfähigen Alter beispielsweise ändert sich drastisch. Je nachdem, wessen Zahlen man betrachtet, wird sich das Verhältnis von etwa neun Menschen im erwerbsfähigen Alter pro Ruhestandler (9:1) auf etwa zwei (2:1) ändern. Es gibt immer mehr ältere Menschen auf der Welt. Das hängt vom jeweiligen Land und anderen Faktoren ab, bedeutet aber, dass es einen Market Pull in Richtung Hilfe für ältere Menschen geben wird, wenn sie gebrechlicher werden. In Japan kann man das auf Roboter messen schon beobachten, bei denen viele Prototypen von Robotern gezeigt werden, die älteren Menschen helfen, einfache Aufgaben zu erledigen, wie etwa aus dem Bett aufzustehen oder ins Badezimmer zu gelangen, eben einfache Dinge des Alltags. Für diese Tätigkeiten wurden früher Menschen benötigt, aber wenn sich das Verhältnis von Menschen im erwerbsfähigen Alter zu den Älteren ändert, wird es nicht genügend Arbeitskräfte geben, um diese Nachfrage zu decken. Dann werden Roboter zum Einsatz kommen, um den älteren Menschen zu helfen.

MARTIN FORD: Ich stimme zu, dass sich der Robotik- und KI-Branche im Bereich der Altenpflege viele Möglichkeiten bieten, aber die Geschicklichkeit, die erforderlich ist, um einem älteren Menschen wirklich zu helfen, allein zurechtzukommen, scheint mir eine ziemliche Herausforderung zu sein.

RODNEY BROOKS: Es sollen nicht einfach nur Menschen durch Robotersysteme ersetzt werden. Aber es wird eine Nachfrage geben, und dementsprechend wird es motivierte Entwickler geben, die versuchen werden, Lösungen zu finden, denn dieser Markt wird enorm sein.

Ich glaube auch, dass es eine erhöhte Nachfrage nach Bauarbeiten geben wird, denn wir urbanisieren die Welt in einem unglaublichen Tempo. Viele der Verfahren, die wir beim Bau einsetzen, sind schon von den Römern erfunden worden. Da ist durchaus Raum für ein paar technologische Neuerungen.

MARTIN FORD: Denken Sie dabei an Bauroboter oder eher an 3D-Druck in großem Maßstab?

RODNEY BROOKS: Für einige Dinge könnte 3D-Druck gut geeignet sein. Wir werden kaum das gesamte Gebäude drucken, aber vielleicht einige vorgeformte Komponenten. Wir werden viele fertige Bauteile verwenden können, die nicht vor Ort hergestellt werden müssen, was wiederum zu Innovationen bei der Lieferung, beim Transport und beim Bewegen dieser Bauteile führen wird. Hier gibt es Spielraum für viele Innovationen.

Die Landwirtschaft ist eine weitere Branche, in der Robotik und KI-Innovationen sich als nützlich erweisen können, insbesondere in Anbetracht des Klimawandels, der unsere Lebensmittelversorgung bedroht. Landwirtschaft in Städten (»Urban Farming«) ist schon in aller Munde und kann die Produktion vom Feld in eine Fabrik verlagern. Hier kann sich Machine Learning als äußerst hilfreich erweisen. Wir verfügen jetzt über die Rechenleistung, um Pflanzen zu überwachen, sie mit genau den richtigen Nährstoffen zu versorgen und ihnen die Bedingungen zu bieten, die sie brauchen, ohne dass wir uns Sorgen über das Wetter machen müssten. Ich denke, dass der Klimawandel die Automatisierung der Landwirtschaft auf andere Weise als bisher vorantreiben wird.

MARTIN FORD: Was ist mit Haushaltsrobotern für Privatkunden? Man hört immer von dem Beispiel, dass ein Roboter Ihnen ein Bier bringt. Es hört sich so an, als ob das auch noch nicht machbar ist.

RODNEY BROOKS: Colin Angle, der Geschäftsführer von iRobot, der das Unternehmen 1990 zusammen mit mir gegründet hat, spricht seit 28 Jahren davon. Ich denke, dass ich noch eine Weile selbst zum Kühlschrank gehen werde.

MARTIN FORD: Glauben Sie, dass es jemals einen allgegenwärtigen Roboter für Privatkunden geben wird, der den Endkundenmarkt abdeckt, indem er etwas leistet, das als unentbehrlich betrachtet wird?

RODNEY BROOKS: Ist Roomba unentbehrlich? Nein, aber er leistet etwas Nützliches zu einem Preis, der so niedrig ist, dass die Leute bereit sind, ihn zu zahlen. Er ist nicht unentbehrlich, es geht mehr um Bequemlichkeit.

MARTIN FORD: Wann wird es einen Roboter geben, der mehr kann, als herumzufahren und den Fußboden zu staubsaugen? Einen Roboter, der geschickt genug ist, um einfache Aufgaben zu erledigen?

RODNEY BROOKS: Das würde ich auch gerne wissen! Aber das weiß wohl niemand. Alle reden davon, dass Roboter die Welt beherrschen werden, aber wir können noch nicht einmal die Frage beantworten, wann sie in der Lage sein werden, uns ein Bier zu holen.

MARTIN FORD: Ich habe kürzlich einen Artikel gelesen, in dem der Geschäftsführer von Boeing, Dennis Muilenburg, gesagt hat, dass es im nächsten Jahrzehnt autonome Drohntaxis geben wird, die Passagiere herumfliegen. Was halten Sie von dieser Prognose?

RODNEY BROOKS: Ich würde das mit der Prognose vergleichen, dass es fliegende Autos geben wird. Fliegende Autos, mit denen man herumfahren und dann einfach abheben kann, sind schon lange ein Wunschtraum, aber ich glaube nicht, dass es sie geben wird.

Ich glaube, es war Travis Kalanick, der ehemalige Geschäftsführer von Uber, der behauptet hat, dass Uber im Jahr 2020 autonome fliegende Autos einsetzen wird. Aber so wird es nicht kommen. Das soll nicht heißen, dass ich nicht daran glaube, dass es irgendeine Form von autonomer persönlicher Beförderung geben wird. Es gibt ja jetzt schon Helikopter und ähnliche Luftfahrzeuge, die zuverlässig von A nach B fliegen können, ohne dass sie jemand steuert. Ich glaube, es ist eher die Wirtschaftlichkeit, die festlegt, wann es soweit sein wird, aber wann es geschehen wird, kann ich nicht sagen.

MARTIN FORD: Und was ist mit einer AGI? Glauben Sie, dass sie erreichbar ist? Und wenn ja: In welchem Zeitrahmen werden wir eine Chance von 50 Prozent haben, sie zu erreichen?

RODNEY BROOKS: Ja, ich glaube, dass eine AGI erreichbar ist. Meine Schätzung ist das Jahr 2200. Aber das ist nur eine Schätzung.

MARTIN FORD: Auf welchem Weg werden wir eine AGI erreichen? Mit welchen Hindernissen müssen wir rechnen?

RODNEY BROOKS: Über das Hindernis der mangelnden Geschicklichkeit haben wir bereits gesprochen. Die Fähigkeit, zu navigieren und die Umgebung zu manipulieren, ist wichtig, um die Welt zu verstehen, aber bei der Welt gibt es einen viel umfassenderen Kontext, als nur den physischen. Roboter und KI-Systeme begreifen beispielsweise nicht, dass heute ein anderer Tag ist als gestern – abgesehen

davon, dass sich die Ziffer des Kalenderdatums geändert hat. Sie haben kein Gedächtnis mit Erfahrungen, kein Bewusstsein dafür, dass sie einen Tag nach dem anderen vorhanden sind und verstehen langfristige Ziele nicht, und dass man inkrementelle Fortschritte erzielen kann, um sie zu erreichen. Jedes heutige KI-Programm ist ein »Idiot savant«, der in einem Meer des Jetzt lebt. Es erhält eine Eingabe und reagiert darauf.

Das Programm AlphaGo und Schachprogramme wissen nicht, was ein Spiel ist, sie wissen nichts über das Spielen eines Spiels, sie wissen nicht, dass Menschen existieren, sie wissen nichts von alledem. Wenn eine AGI einem Menschen ebenbürtig ist, muss sie sich allerdings all dieser Dinge bewusst sein.

Vor 50 Jahren arbeiteten Menschen an Forschungsprojekten über diese Dinge. Es gab eine ganze Community, die in den 1980er- und 1990er-Jahren an der Simulation von adaptivem Verhalten arbeitete. Seitdem haben wir keine großen Fortschritte erzielt, und wir wissen nicht, wie wir es anstellen könnten. Derzeit arbeitet niemand daran, und die Leute, die behaupten, einer AGI näherzukommen, wiederholen nur die gleichen Dinge, von denen John McCarthy in den 1960er-Jahren gesprochen hat, und sie machen ungefähr genauso viele Fortschritte.

Es ist eine schwierige Aufgabe. Es bedeutet nicht, dass man keinen Fortschritt in vielen verschiedenen Technologien erzielt, aber um manche Dinge zu erreichen, sind einfach Hunderte von Jahren erforderlich. Wir glauben, die glückliche Generation am entscheidenden Zeitpunkt zu sein, aber das haben viele Leute schon oft gedacht, und es wird für uns nicht plötzlich wahr, jedenfalls kann ich keine Anzeichen dafür erkennen.

MARTIN FORD: Es gibt Bedenken, dass wir beim Wettlauf um eine fortgeschrittene Künstliche Intelligenz mit China nicht Schritt halten können. China besitzt eine größere Bevölkerung, verfügt somit über mehr Daten und die Datenschutzvorschriften, die beschränken könnten, was sie mit KI anstellen, sind nicht so streng. Glauben Sie, dass uns ein neuer KI-Rüstungswettlauf bevorsteht?

RODNEY BROOKS: Sie haben recht, es wird einen Wettlauf geben. Es gab einen Wettlauf von Unternehmen, und es wird einen Wettlauf von Ländern geben.

MARTIN FORD: Halten Sie es für eine große Gefahr für den Westen, wenn ein Land wie China einen beträchtlichen Vorsprung in der KI hätte?

RODNEY BROOKS: Ich glaube nicht, dass es so einfach ist. Es wird einen uneinheitlichen Einsatz von KI-Technologien geben. Ich glaube, in China können wir das beim Einsatz von Gesichtserkennung schon beobachten, der auf eine Weise erfolgt, die wir hier bei uns in den USA lieber nicht sehen würden. Bei der Ent-

wicklung von KI-Chips kann es sich ein Land wie die USA nicht erlauben, auch nur ein wenig ins Hintertreffen zu geraten. Um allerdings nicht ins Hintertreffen zu geraten, müssten die USA hier die Führungsrolle einnehmen, was derzeit nicht der Fall ist.

Es hat politische Weisungen gegeben, die besagen, dass wir mehr Bergleute brauchen, während wissenschaftliche Budgets wie das des National Institute of Standards and Technology zusammengestrichen werden. Das ist verrückt, das ist wahnhaft, es ist rückschrittliches Denken, und es ist zerstörerisch.

MARTIN FORD: Kommen wir auf die Risiken oder potenziellen Gefahren zu sprechen, die mit KI und Robotik in Verbindung gebracht werden. Viele Menschen glauben, dass wir kurz vor einer großen Umwälzung stehen, die einer neuen industriellen Revolution gleichkommt. Glauben Sie das auch? Wird es große Auswirkungen auf den Arbeitsmarkt und die Wirtschaft geben?

RODNEY BROOKS: Ja, aber nicht auf die Art und Weise, wie die Menschen sich das vorstellen. Ich glaube nicht, dass KI an sich die Ursache ist, sondern die Digitalisierung und die Entwicklung neuer digitaler Schnittstellen. Ich führe als Beispiel gern mautpflichtige Straßen an. In den USA gibt es bei mautpflichtigen Straßen und Brücken kaum noch menschliche Geldeinnehmer. Nicht hauptsächlich, weil sie durch KI ersetzt wurden, sondern weil in unserer Gesellschaft in den letzten 30 Jahren die verschiedensten digitalen Schnittstellen entwickelt worden sind.

Die Geldeinnehmer zu ersetzen, wurde beispielsweise durch die Aufkleber möglich, die Sie an der Windschutzscheibe anbringen und die Ihrem Auto eine digitale Signatur verleihen. Ein weiterer Fortschritt, der es ermöglichte, die Geldeinnehmer zu ersetzen, ist Computer Vision, bei dem ein KI-System mithilfe von Deep Learning ein Foto des Kfz-Kennzeichens aufnimmt und es zuverlässig erkennt. Es betrifft aber nicht nur die Mautstellen. Digitalisierung findet auch in vielen anderen Bereichen statt. Sie können beispielsweise eine Website aufsuchen, den Aufkleber in ihrem Auto und die Ihnen zugeordnete Seriennummer registrieren sowie Ihr Kfz-Kennzeichen angeben, damit es ein Backup gibt.

Dann gibt es noch das digitale Onlinebanking, das es Dritten ermöglicht, Ihre Kreditkarte regelmäßig zu belasten, ohne dass sie physisch vorliegen muss. Früher brauchten Sie die physische Kreditkarte, heutzutage gibt es ein digitales Verfahren. Darüber hinaus gibt es für den Betreiber der Mautstellen noch den Nebeneffekt, dass er das Bargeld nicht mehr per Geldtransporter einsammeln und zur Bank transportieren muss, weil alles digital abgewickelt wird. Es gibt eine ganze Reihe digitaler Verfahren, die zwecks Automatisierung dieser Dienstleistung zusammengekommen sind und den menschlichen Geldeinnehmer ersetzen. KI spielte dabei nur eine kleine, wenngleich unverzichtbare Rolle. Die Menschen wur-

den nicht über Nacht durch KI-Systeme ersetzt. Diese zunehmend vorhandenen digitalen Schnittstellen sind es, die Veränderungen auf dem Arbeitsmarkt verursachen, aber es ist keine einfache Eins-zu-eins-Ersetzung.

MARTIN FORD: Glauben Sie, dass diese digitalen Abläufe bei vielen der einfachen Dienstleistungsjobs Umwälzungen verursachen werden?

RODNEY BROOKS: Digitale Abläufe können vieles leisten, aber nicht alles. Sie sind etwa völlig ungeeignet für Aufgaben, die wir typischerweise kaum zu schätzen wissen, die aber unverzichtbar sind, um unsere Gesellschaft am Laufen zu halten, wie etwa älteren Menschen auf der Toilette oder beim Betreten und Verlassen der Dusche zu helfen. Es sind aber nicht nur Aufgaben dieser Art – denken Sie beispielsweise an Lehrtätigkeiten. In den USA haben wir es nicht geschafft, Lehrern die Anerkennung zu zollen oder die Gehälter zu zahlen, die sie verdienen würden, und ich weiß nicht, wie wir unsere Gesellschaft dazu bewegen können, diese wichtige Arbeit wertzuschätzen und finanziell lohnend zu machen. Wenn einige der Jobs durch Automatisierung verlorengehen, wie können wir dann die anderen Jobs erkennen und in den Vordergrund stellen, bei denen das nicht so ist?

MARTIN FORD: Das hört sich so an, als ob Sie keine Massenarbeitslosigkeit erwarten, aber dass sich die Jobs verändern werden. Ich glaube, dass viele begehrte Jobs verschwinden werden. Denken Sie an all die Angestelltentätigkeiten, bei denen man vor einem Computer sitzt, irgendwelche vorhersehbaren und monotonen Aufgaben erledigt, und immer wieder die gleichen Berichte erstellt. Das sind sehr begehrte und gut bezahlte Jobs, für die Leute ein College besuchen müssen, und diese Jobs werden bedroht sein, aber das Zimmermädchen, das Hotelzimmer putzt, wird auf der sicheren Seite sein.

RODNEY BROOKS: Das kann ich nicht abstreiten, ich bestreite aber sehr wohl, dass die Leute sagen: »Oh, hier sind KI und Roboter im Einsatz!«. Wie ich schon sagte, glaube ich, dass eher die Digitalisierung ausschlaggebend ist.

MARTIN FORD: Dem stimme ich zu, aber es ist auch eine Tatsache, dass auf solch einer Plattform KI eingesetzt wird.

RODNEY BROOKS: Ja, es ist eindeutig einfacher, KI auf solch einer Plattform einzusetzen. Eine weitere Sorge ist natürlich, dass die Plattform aus völlig unsicheren Komponenten aufgebaut wird und jeder sie hacken kann.

MARTIN FORD: Kommen wir zum Thema Sicherheit. Über welche Dinge sollten wir uns, abgesehen von den wirtschaftlichen Umwälzungen, wirklich Sorgen machen? Was sind die wirklichen Risiken, wie etwa Sicherheit, die tatsächlich vorhanden sind und denen wir Beachtung schenken sollten?

RODNEY BROOKS: Am wichtigsten ist die Sicherheit. Ich mache mir Sorgen um die Sicherheit der digitalen Verfahren und um den Datenschutz, den wir alle im Austausch gegen eine gewisse Bequemlichkeit bereitwillig aufgegeben haben. Wir haben schon erlebt, dass soziale Plattformen zu einer Waffe gemacht wurden. Anstatt sich über eine KI mit Selbstbewusstsein Sorgen zu machen, die etwas Böswilliges oder Schlechtes tut, sollten wir etwas anderem Beachtung schenken. Es ist sehr viel wahrscheinlicher, dass menschliche Akteure Böses tun, die herausgefunden haben, wie sie die Schwachstellen der digitalen Verfahren ausnutzen können, seien es nun Nationalstaaten, kriminelle Unternehmen oder sogar der einzelne Hacker im heimischen Schlafzimmer.

MARTIN FORD: Was ist mit der Bewaffnung von Robotern und Drohnen? Stuart Russell, einer meiner Gesprächspartner in diesem Buch, hat einen ziemlich erschreckenden Film mit dem Titel *Slaughterbots* produziert, der sich mit diesem Problem befasst.

RODNEY BROOKS: Ich denke, dass derartige Dinge schon heute durchaus möglich sind, weil dafür keine KI erforderlich ist. *Slaughterbots* war eine reflexartige Reaktion mit der Aussage, dass Roboter und Krieg eine schlechte Kombination sind. Bei mir löst das noch eine andere Reaktion aus. Ich hatte immer den Eindruck, dass Roboter es sich leisten können, als Zweiter zu schießen. Ein 19 Jahre alter Junge, der gerade die Highschool verlassen hat und sich in einem fremden Land im Dunkel der Nacht befindet und um sich herum Schüsse hört, kann es sich nicht leisten, als Zweiter zu schießen.

Manche Menschen meinen, dass KI dem Militär vorenthalten werden sollte, um das Problem zu lösen. Ich bin der Ansicht, man sollte stattdessen lieber darüber nachdenken, was man eigentlich verhindern möchte und das gesetzlich regeln, anstatt die verwendete Technologie zu verbieten. Viele dieser Dinge könnten auch ohne KI entwickelt werden.

Ein Beispiel: Wenn wir das nächste Mal auf dem Mond landen, werden dabei jede Menge KI und Machine Learning zum Einsatz kommen, aber in den 1960ern haben wir es auch ohne KI und Machine Learning dorthin und wieder zurück geschafft. Wir müssen über die Handlung selbst nachdenken, nicht über die zur Ausführung der Handlung eingesetzte Technologie. Es wäre naiv, eine Technologie zu verbieten, ohne die guten Dinge zu berücksichtigen, die sie bewirken kann – wie etwa als Zweiter zu schießen, nicht als Erster.

MARTIN FORD: Was ist mit dem AGI-Kontrollproblem und Elon Musks Bemerkungen über das »Heraufbeschwören eines Dämons«? Sollten wir darüber schon jetzt diskutieren?

RODNEY BROOKS: Als die Menschen in Paris 1789 zum ersten Mal Heißluftballons sahen, machten sie sich Sorgen, dass den Ballonfahrern in der Höhe die Seele aus dem Leib gesaugt wird. Das ist das gleiche Niveau des Verständnisses, das bei einer AGI vorliegt. Wir haben keine Ahnung, wie sie aussehen wird. Ich habe einen Artikel mit dem Titel *The Seven Deadly Sins of Predicting the Future of AI* (Die sieben Todsünden der Vorhersage der Zukunft der KI, siehe <https://rodneybrooks.com/the-seven-deadly-sins-of-predicting-the-future-of-ai/>) geschrieben, der das Ganze zusammenfasst. Die Welt wird eine andere als heute sein, wenn in ihrem Zentrum eine Superintelligenz steht. Sie wird sich im Lauf der Zeit ganz allmählich entwickeln. Wir haben überhaupt keine Ahnung, wie die Welt oder das KI-System aussehen werden. Das Vorhersagen einer KI-Zukunft ist lediglich ein Machtspielchen für vereinzelte Akademiker, die in einem realitätsfernen Elfenbeinturm leben. Das soll nicht heißen, dass es die Technologie nicht geben wird, aber wir werden nicht wissen, wie sie aussieht, bevor es soweit ist.

MARTIN FORD: Wenn diese technologischen Durchbrüche eintreten, ist dann Ihrer Meinung nach eine Regulierung erforderlich?

RODNEY BROOKS: Wie ich vorhin schon gesagt habe: Reguliert werden muss, was diesen Systemen erlaubt ist und was nicht, nicht die Technologie, die ihnen zugrunde liegt. Sollten wir die Forschung an optischen Computern einstellen, weil sie Ihnen viel schnellere Matrizenmultiplikationen ermöglichen, sodass Sie umfassenderes Deep Learning viel zügiger anwenden könnten? Nein, das ist verrückt. Sollte es selbstfahrenden Lieferwagen gestattet sein, in verstopften Bereichen von San Francisco in zweiter Reihe zu parken? Das zu regulieren, wäre sinnvoll, nicht die dazu nötige Technologie.

MARTIN FORD: Wenn ich all dies berücksichtige, nehme ich an, dass Sie insgesamt ein Optimist sind? Sie arbeiten weiterhin daran, also müssen sie glauben, dass die Vorteile die möglichen Risiken wettmachen.

RODNEY BROOKS: Ja, unbedingt. Wir haben die Überbevölkerung der Welt zugelassen, deshalb müssen wir diesen Weg gehen, um zu überleben. Ich mache mir große Sorgen, dass der Lebensstandard sinkt, weil es nicht genügend Arbeit gibt, wenn ich älter werde. Ich mache mir Sorgen um die Sicherheit und um den Datenschutz, um zwei weitere Punkte zu nennen. All das sind echte und gegenwärtige Gefahren, und wir können allmählich erahnen, wie sie aussehen werden.

Hollywoods Vorstellung, dass eine AGI die Weltherrschaft übernehmen könnte, liegt noch in ferner Zukunft, und wir wissen noch nicht einmal, wie wir uns das vorstellen sollen. Wir sollten uns um die echten Gefahren und Risiken Sorgen machen, denen wir jetzt schon gegenüberstehen.

Rodney Brooks ist Robotik-Unternehmer und hat einen Dokortitel in Informatik von der Stanford University. Er ist derzeit Vorstand und Chief Technology Officer von Rethink Robotics. Von 1997 bis 2007 war Brooks Leiter des MIT Artificial Intelligence Laboratory und später des MIT Computer Science and Artificial Intelligence Laboratory (CSAIL). Er ist Fellow bei verschiedenen Organisationen, unter anderem bei der Association for the Advancement of Artificial Intelligence (AAAI), wo er Gründungsmitglied war. Er hat während seiner Laufbahn eine Reihe von Auszeichnungen für seine Arbeit auf dem Fachgebiet erhalten, unter anderem den Computers and Thought Award, den IEEE Inaba Technical Award for Innovation Leading to Production, den Robotics Industry Association's Engelberger Robotics Award for Leadership und den IEEE Robotics and Automation Award.

In dem 1997 erschienenen Film *Fast, Cheap and Out of Control* (Schnell, billig und außer Kontrolle) von Errol Morris spielt er sich selbst. Der Film ist nach dem Titel einer seiner Arbeiten benannt und ist bei Rotten Tomatoes mit 91 Prozent bewertet.