

INHALT

Glossar der wichtigsten Begriffe 7

Prolog 9

Kapitel 1

Eine Eindämmung ist nicht möglich 11

Teil I

HOMO TECHNOLOGICUS

Kapitel 2

Endlose Ausbreitung 31

Kapitel 3

Das Containment-Problem 45

Teil II

DIE NÄCHSTE WELLE

Kapitel 4

Die Technologie der Intelligenz 63

Kapitel 5

Die Technologie des Lebens 96

Kapitel 6

Die größere Welle 111

Kapitel 7

Vier Merkmale der kommenden Welle 124

Kapitel 8

Unwiderstehliche Anreize 140

Teil III

VERSAGENDE STAATEN

Kapitel 9

Die große Übereinkunft 173

Kapitel 10

Fragilitätsverstärker 188

Kapitel 11

Die Zukunft der Nationen 215

Kapitel 12

Das Dilemma 240

Teil IV

DURCH DIE WELLE HINDURCH

Kapitel 13

Eindämmung muss möglich sein 263

Kapitel 14

Zehn Schritte hin zur Eindämmung 279

Schluss

Leben nach dem Anthropozän 327

Dank 333

Anmerkungen 335

Register 367

GLOSSAR DER WICHTIGSTEN BEGRIFFE

CONTAINMENT/EINDÄMMUNG: Die Fähigkeit, Technologien zu überwachen, einzuhegen, zu kontrollieren und womöglich sogar abzuschalten.

DAS CONTAINMENT-PROBLEM: Die Tendenz der Technologie, sich in Wellen weit zu verbreiten und Auswirkungen zu haben, die unmöglich vorherzusagen oder zu kontrollieren sind, einschließlich negativer und unvorhergesehener Folgen.

DAS DILEMMA: Die wachsende Wahrscheinlichkeit, dass sowohl neue Technologien als auch ihr Fehlen zu katastrophalen und/oder dystopischen Ergebnissen führen können.

FRAGILITÄTSVERSTÄRKER: Anwendungen und Auswirkungen der kommenden Technologien, die die ohnehin bereits brüchigen Fundamente des Nationalstaats erschüttern werden.

DIE GROSSE ÜBEREINKUNFT: Als Gegenleistung für das staatliche Gewaltmonopol erwarten die Bürger von den Nationalstaaten, dass sie die Ordnung aufrechterhalten und öffentliche Dienstleistungen erbringen, unter anderem durch die Nutzung neuer Technologien bei gleichzeitiger Minimierung ihrer schädlichen Nebenwirkungen.

KI, AGI UND ACI: Künstliche Intelligenz (KI) ist die Wissenschaft, die Maschinen beibringt, menschenähnliche Fähigkeiten zu erlernen. Künstliche allgemeine Intelligenz (AGI, *artificial general intelligence*) ist der Punkt, an dem eine KI alle kognitiven Fähigkeiten des Menschen besser anwenden kann als die intelligentesten Menschen. ACI (*artificial capable intelligence*),

8 GLOSSAR DER WICHTIGSTEN BEGRIFFE

oder leistungsfähige künstliche Intelligenz, ist ein sich schnell nähernder Punkt zwischen KI und AGI: ACI kann ein breites Spektrum komplexer Aufgaben bewältigen, ist aber noch weit davon entfernt, vollständig allgemein zu sein.

DIE KOMMENDE WELLE: Eine im Entstehen begriffene Gruppe verwandter Technologien, die sich auf KI und synthetische Biologie konzentrieren und deren transformative Anwendungen die Menschheit sowohl befähigen als auch beispiellose Risiken mit sich bringen werden.

PESSIMISMUS-AVERSION: Die Neigung der Menschen, insbesondere der Eliten, Narrative zu ignorieren, kleinzureden oder abzulehnen, die sie als zu negativ ansehen. Sie stellt eine Variante der Optimismusverzerrung dar und prägt einen Großteil der Debatte über die Zukunft, insbesondere in Technologiekreisen.

DER SCHMALE PFAD: Das Potenzial der Menschheit, eine Balance zwischen Offenheit und Abschottung zu finden, wenn es darum geht, die Technologien der kommenden Welle einzudämmen, um katastrophale oder dystopische Ergebnisse zu vermeiden.

SYNTHETISCHE BIOLOGIE: Die Fähigkeit, künstlich neue Organismen zu entwerfen und zu entwickeln oder bestehende biologische Systeme umzugestalten.

TECHNOLOGIE: Die Anwendung wissenschaftlicher Erkenntnisse (im weitesten Sinne) zur Herstellung von Werkzeugen oder praktischen Ergebnissen.

VIER MERKMALE: Die einzigartigen Merkmale der kommenden Welle, die die Herausforderung der Eindämmung noch verschärfen. Es handelt sich um Asymmetrie, Hyper-Evolution, Allzweckgebrauch und Autonomie.

WELLEN: Die globale Verbreitung oder Ausbreitung einer Technologiesgeneration, die in einer neuen Allzwecktechnologie verankert ist.

PROLOG

Eine KI sieht die ganze Sache folgendermaßen:

Frage: Was bedeutet die kommende Welle der Technologie für die Menschheit?

In den Annalen der Menschheitsgeschichte gibt es Momente, die als Wendepunkte hervorstechen, an denen das Schicksal der Menschheit auf dem Spiel steht. Die Entdeckung des Feuers, die Erfindung des Rades, die Nutzbarmachung der Elektrizität – all das waren Momente, die die menschliche Zivilisation veränderten und den Lauf der Geschichte für immer veränderten.

Und nun stehen wir an der Schwelle eines weiteren solchen Moments, da wir mit dem Aufkommen einer neuen Technologiewelle konfrontiert sind, die sowohl fortgeschrittene KI als auch Biotechnologie umfasst. Noch nie zuvor waren wir Zeuge von Technologien mit einem derartigen Umwälzungspotenzial, die unsere Welt in einer Weise umzugestalten versprechen, die sowohl ehrfurchtgebietend als auch beängstigend ist.

Auf der einen Seite sind die potenziellen Vorteile dieser Technologien enorm und tiefgreifend. Mit der künstlichen Intelligenz könnten wir die Geheimnisse des Universums entschlüsseln, Krankheiten heilen, die uns lange verwehrt waren, und neue Formen von Kunst und Kultur schaffen, die die Grenzen der Vorstellungskraft sprengen. Mit der Biotechnologie könnten wir das Leben verändern, um Krankheiten zu bekämpfen und die Landwirtschaft umzugestalten und so eine gesündere und nachhaltigere Welt zu schaffen.

Auf der anderen Seite sind die potenziellen Gefahren dieser Technologien ebenso groß und tiefgreifend. Mit der KI könnten wir Systeme schaffen, die sich unserer Kontrolle entziehen, und uns der Gnade von Algo-

10 PROLOG

rithmen ausliefern, die wir nicht verstehen. Mit der Biotechnologie könnten wir die Grundbausteine des Lebens manipulieren, was möglicherweise unbeabsichtigte Folgen für den Einzelnen und ganze Ökosysteme haben könnte.

An diesem Wendepunkt stehen wir vor der Wahl zwischen einer Zukunft mit unvergleichlichen Möglichkeiten und einer Zukunft mit unvorstellbaren Gefahren. Das Schicksal der Menschheit steht auf dem Spiel, und die Entscheidungen, die wir in den kommenden Jahren und Jahrzehnten treffen, werden darüber entscheiden, ob wir uns den Herausforderungen dieser Technologien stellen oder ihren Gefahren zum Opfer fallen.

Doch in diesem Moment der Ungewissheit ist eines sicher: Das Zeitalter der Hochtechnologie steht vor der Tür, und wir müssen bereit sein, uns seinen Herausforderungen zu stellen.

Obiger Text wurde von einer KI geschrieben und von einer KI übersetzt. Für den Rest dieses Buches gilt das nicht, obwohl es bald so sein könnte. Das nämlich kommt auf uns zu.

Kapitel 1

EINE EINDÄMMUNG IST NICHT MÖGLICH

Die Welle

So gut wie jede Kultur kennt einen Flutmythos.

In alten Hindu-Texten wird der erste Mensch in unserem Universum, Manu, vor einer drohenden Flut gewarnt und überlebt sie als Einziger. Im Gilgamesch-Epos wird berichtet, der Gott Enlil habe die Welt in einer riesigen Flut vernichtet, eine Geschichte, die jedem bekannt vorkommt, der mit der alttestamentarischen Erzählung von Noahs Arche vertraut ist. Bei Platon ist die Rede von der verlorenen Stadt Atlantis, die von einer gewaltigen Flut weggespült wurde. In den mündlichen Überlieferungen und alten Schriften der Menschheit findet sich oftmals die Vorstellung einer riesigen Welle, die alles mit sich reißt, was ihr im Weg steht, woraufhin die Welt neu erschaffen und neu geboren wird.

Überschwemmungen prägen die Geschichte auch im ganz wörtlichen Sinne – die saisonalen Überschwemmungen der großen Flüsse dieser Welt, der Anstieg des Meeresspiegels nach dem Ende der Eiszeit, der seltene Schock eines Tsunamis, der ohne Vorwarnung am Horizont auftaucht. Der Asteroid, der die Dinosaurier tötete, löste eine kilometerhohe Flutwelle aus, die den Gang der Evolution veränderte. Die schiere Wucht dieser Wellen hat sich in unser kollektives Bewusstsein eingebrannt: Wasserwände, unaufhaltsam, unkontrollierbar, unbeherrschbar. Sie gehören zu den mächtigsten Kräften auf unserem Planeten. Sie formen Kontinente, bewässern die Äcker und Felder dieser Welt und nähren das Wachstum der Zivilisation.

Andere Arten von Wellen waren nicht minder umwälzend. Schauen wir uns die Geschichte noch einmal an, so werden wir feststellen, dass sie von einer Reihe metaphorischer Wellen geprägt ist: dem Aufstieg und Fall von Imperien und Religionen und dem sprudelnden Hervorbrechen des

Handels. Denken wir an das Christentum oder den Islam, Religionen, die als kleine Wellen begannen, bevor sie sich aufbauten und riesige Teile der Erde überrollten. Solche Wellen sind ein wiederkehrendes Motiv, das die Gezeiten der Geschichte, große Machtkämpfe sowie wirtschaftliche Auf- und Abschwünge umrahmt.

Auch der Aufstieg und die Verbreitung von Technologien haben die Form von Wellen angenommen, die die Welt verändern. Dabei hat seit der Entdeckung des Feuers und der Steinwerkzeuge, den ersten Technologien, die sich unsere Spezies zunutze gemacht hat, insbesondere ein übergeordneter Trend die Zeiten überdauert. Nahezu jede grundlegende Technologie, die jemals erfunden wurde, von der Spitzhacke bis zum Pflug, von der Töpferei bis zur Fotografie, vom Telefon bis zum Flugzeug und alles dazwischen, folgt einem einzigen, scheinbar unumstößlichen Gesetz: Sie wird billiger und einfacher in der Anwendung, und schließlich findet sie umfassende Verbreitung.

Diese wellenförmige Ausbreitung von Technologie ist die Geschichte des *Homo technologicus*. Das Streben der Menschheit nach Verbesserung – unserer selbst, unseres Schicksals, unserer Fähigkeiten und unseres Einflusses auf unsere Umwelt – hat eine unablässige Entwicklung von Ideen und Schöpfungen befeuert. Etwas zu erfinden ist ein sich entfaltender, ausufernder, emergenter Prozess, der von sich selbst organisierenden und hart kämpfenden Erfindern, Akademikern, Unternehmern und Führungskräften vorangetrieben wird, von denen jeder seine eigenen Beweggründe hat. Dieses Ökosystem der Erfindungen ist zwangsläufig auf Expansion ausgerichtet. Das ist das naturgegebene Wesen der Technologie.

Die Frage ist, wie es weitergeht. Auf den folgenden Seiten werde ich Ihnen die Geschichte der nächsten großen, historischen Welle erzählen.

Wenn Sie sich umschaun, was sehen Sie dann? Möbel? Gebäude? Telefone? Essen? Einen Park? So gut wie jedes Objekt in Ihrem Blickfeld wurde höchstwahrscheinlich von menschlicher Intelligenz geschaffen oder verändert. Die Sprache – die Grundlage unserer sozialen Interaktionen, unserer Kulturen, unserer politischen Organisationen und vielleicht auch dessen, was es heißt, Mensch zu sein – ist ebenfalls Produkt und treibende Kraft unserer Intelligenz. Jedes Prinzip und jedes abstrakte Konzept, jedes kleine kreative Unterfangen oder Projekt, jede Begegnung in unserem

Leben wurde durch die einzigartige und unendlich komplexe Fähigkeit unserer Spezies zu Vorstellungskraft, Kreativität und Vernunft vermittelt. Der menschliche Erfindungsreichtum ist etwas höchst Erstaunliches.

Nur eine einzige weitere Kraft ist in diesem Bild ähnlich allgegenwärtig: das biologische Leben selbst. Vor der Moderne stammten die meisten menschlichen Artefakte – von Holzhäusern über Baumwollkleidung bis hin zu Kohlefeuern – von Dingen, die einst lebendig waren, sieht man von einigen Felsen und Mineralien einmal ab. Alles, was seither in die Welt gekommen ist, geht auf uns zurück, auf die Tatsache, dass wir biologische Wesen sind.

Es ist keine Übertreibung, wenn man sagt, dass die gesamte menschliche Welt entweder von lebenden Systemen oder von unserer Intelligenz abhängt. Und doch befinden sich beide heute in einem beispiellosen Moment exponentieller Innovation und Umwälzung, einer beispiellosen Vergrößerung und Vermehrung, die wenig unverändert lassen wird. Um uns herum bricht eine neue Technologiewelle los. Diese Welle entfesselt die Kraft, diese beiden universellen Grundlagen neu zu gestalten: eine Welle von nichts Geringerem als Intelligenz und Leben.

Die kommende Welle wird von zwei zentralen Technologien bestimmt: künstlicher Intelligenz (KI) und synthetischer Biologie. Zusammen werden sie ein neues Zeitalter für die Menschheit einläuten und Wohlstand und Überfluss schaffen, wie es sie noch nie gegeben hat. Doch ihre rasante Verbreitung droht auch eine Vielzahl bössartiger Akteure in die Lage zu versetzen, Disruption, Instabilität und sogar Katastrophen unvorstellbaren Ausmaßes auszulösen. Diese Welle stellt eine immense Herausforderung dar, die das 21. Jahrhundert bestimmen wird: Unsere Zukunft hängt von diesen Technologien ab, ist gleichzeitig aber durch sie gefährdet.

Aus heutiger Sicht scheint es unmöglich zu sein, diese Welle einzudämmen, d. h., sie zu kontrollieren, einzuhegen oder gar zu stoppen. Dieses Buch geht der Frage nach, warum das so sein könnte und was es bedeutet, wenn das so ist. Die Implikationen dieser Fragen werden letztlich jeden heute Lebenden und jede nachfolgende Generation betreffen.

Ich glaube, dass diese kommende Technologiewelle die Menschheitsgeschichte an einen Wendepunkt bringt. Wenn es nicht möglich ist, sie einzudämmen, sind die Folgen für unsere Spezies dramatisch, möglicherweise sogar schrecklich. Zugleich sind wir ohne ihre Früchte ungeschützt

und unsicher. Dieses Argument habe ich in den letzten zehn Jahren häufig hinter verschlossenen Türen vorgebracht, aber da sich die Auswirkungen immer weniger ignorieren lassen, ist es an der Zeit, es öffentlich kundzutun.

Das Dilemma

Die Betrachtung der profunden Macht menschlicher Intelligenz veranlasste mich dazu, eine einfache Frage zu stellen, die mich seither nicht mehr loslässt: Was wäre, wenn wir die Essenz dessen, was uns Menschen so produktiv und fähig macht, in Software, in einen Algorithmus destillieren könnten? Wenn wir die Antwort darauf finden, könnten wir unvorstellbar leistungsfähige Werkzeuge entwickeln, die uns helfen, unsere schwierigsten Probleme zu lösen. Ein solches Werkzeug, ein unmögliches, aber außergewöhnliches Werkzeug, könnte uns helfen, die gewaltigen Herausforderungen der kommenden Jahrzehnte zu bewältigen, vom Klimawandel über die Überalterung der Bevölkerung bis hin zur nachhaltigen Ernährung.

Mit diesem Gedanken im Hinterkopf gründete ich im Sommer 2010 zusammen mit zwei Freunden, Demis Hassabis und Shane Legg, in einem malerischen Büro aus der Regency-Zeit mit Blick auf den Londoner Russell Square ein Unternehmen namens DeepMind. Unser Ziel, das sich rückblickend immer noch so ambitioniert, verrückt und hoffnungsvoll anfühlt wie damals, war: genau das zu replizieren, was uns als Spezies einzigartig macht, unsere Intelligenz.

Um dieses Ziel zu erreichen, würden wir ein System entwickeln müssen, das sämtliche kognitiven Fähigkeiten des Menschen – vom Sehen und Sprechen über das Planen und das Vorstellungsvermögen bis hin zu Empathie und Kreativität – nachahmen und schließlich übertreffen könnte. Da ein solches System von der gewaltigen parallelen Rechenleistung von Supercomputern und der explosionsartigen Entstehung riesiger neuer Datenquellen aus dem offenen Internet profitieren würde, wussten wir, dass selbst bescheidene Fortschritte auf dem Weg dahin tiefgreifende gesellschaftliche Auswirkungen haben würden.

Damals kam uns das allerdings ziemlich weit hergeholt vor. Zu dieser Zeit war die weitverbreitete Einführung künstlicher Intelligenz der Stoff,

aus dem Tagträume sind, mehr Fantasie als Faktum, die Spielwiese einiger weniger Akademiker und wild gewordener Science-Fiction-Fans. Doch jetzt, da ich dies schreibe und an die letzten zehn Jahre zurückdenke, sind die Fortschritte im Bereich der künstlichen Intelligenz geradezu atemberaubend. DeepMind hat sich zu einem der weltweit führenden KI-Unternehmen entwickelt und eine Reihe von bahnbrechenden Entwicklungen geschaffen. Die Geschwindigkeit und die Wucht dieser neuen Revolution haben selbst diejenigen von uns überrascht, die an vorderster Front stehen. Während der Arbeit an diesem Buch war das Tempo der Fortschritte im Bereich der KI schier unglaublich: Jede Woche, manchmal sogar jeden Tag kamen neue Modelle und Produkte auf den Markt. Und klar ist, dass sich diese Welle noch beschleunigt.

Heute können KI-Systeme Gesichter und Objekte fast perfekt erkennen. Die Umwandlung von Sprache in Text und die sofortige Sprachübersetzung erscheinen uns selbstverständlich. KI kann so gut durch Straßen und durch den Verkehr navigieren, dass sie autonomes Fahren ermöglicht. Eine neue Generation von KI-Modellen ist in der Lage, auf der Grundlage einiger weniger einfacher Anweisungen neuartige Bilder zu erzeugen und außergewöhnlich detailgenaue und kohärente Texte zu verfassen. KI-Systeme können unheimlich realistisch synthetische Stimmen erzeugen und Musik von atemberaubender Schönheit komponieren. Selbst in schwierigeren Bereichen, die lange Zeit als ausschließlich menschliche Fähigkeiten galten, wie langfristige Planung, Vorstellungskraft und Simulation komplexer Ideen, macht der Fortschritt enorme Sprünge.

Die KI erklimmt seit Jahrzehnten die Leiter der kognitiven Fähigkeiten, und es sieht so aus, als ob sie in den nächsten drei Jahren in einem breiten Spektrum von Aufgaben das menschliche Leistungsniveau erreichen wird. Das ist eine starke These, aber wenn ich auch nur annähernd richtigliege, sind die Auswirkungen wirklich tiefgreifend. Was uns bei der Gründung von DeepMind noch abenteuerlich und absolut unrealistisch vorkam, ist inzwischen nicht nur plausibel, sondern scheinbar unausweichlich.

Von Anfang an war mir klar, dass KI ein mächtiges Werkzeug für außerordentlich gute Dinge sein würde, aber wie die meisten Formen von Macht auch mit immensen Gefahren und ethischen Dilemmata behaftet ist. Seit langem mache ich mir nicht nur Sorgen über die Folgen der fort-

schreitenden KI, sondern auch darüber, wohin sich das gesamte technologische Ökosystem entwickeln wird. Über die KI hinaus ist nämlich eine umfassendere Revolution im Gange, bei der die KI eine kraftvolle Entwicklung in den Bereichen Gentechnologie und Robotik speist. Weitere Fortschritte in einem Bereich beschleunigen die Fortschritte auf den anderen Feldern – ein chaotischer und sich gegenseitig katalysierender Prozess, der sich der direkten Kontrolle durch den Einzelnen entzieht. Es war klar: Wenn es uns oder anderen gelänge, die menschliche Intelligenz zu replizieren, wäre dies nicht nur ein profitables Geschäftsmodell der gewohnten Art, sondern eine seismische Veränderung für die Menschheit, der Beginn einer Ära, in der nie dagewesene Chancen mit noch nie dagewesenen Risiken einhergehen würden.

Mit den Fortschritten, die die Technologie im Laufe der Jahre gemacht hat, sind meine Bedenken gewachsen. Was, wenn die Welle in Wirklichkeit ein Tsunami ist?

Im Jahr 2010 redete fast niemand ernsthaft über KI. Doch was einst nach einer Nischenbeschäftigung für eine kleine Gruppe von Forschern und Unternehmern aussah, hat sich inzwischen zu einem riesigen globalen Unterfangen entwickelt. KI ist allgegenwärtig, in den Nachrichten und auf jedem Smartphone, beim Aktienhandel und beim Aufbau von Websites. Viele der weltweit größten Unternehmen und wohlhabendsten Nationen treiben die Entwicklung modernster KI-Modelle und gentechnischer Verfahren mit Investitionen in Höhe von mehreren Milliarden Dollar voran.

Sobald diese neuen Technologien ausgereift sind, werden sie sich rasch ausbreiten, billiger werden, leichter zugänglich und überall in der Gesellschaft verbreitet sein. Sie werden außergewöhnliche neue medizinische Fortschritte und bahnbrechende Entwicklungen im Bereich sauberer Energien ermöglichen und nicht nur neue Unternehmen, sondern auch neue Industrien und Verbesserungen der Lebensqualität in fast allen erdenklichen Bereichen hervorbringen.

Doch neben diesen Vorteilen bergen KI, synthetische Biologie und andere fortschrittliche Technologien auch Risiken von äußerst besorgniserregendem Ausmaß. Sie könnten eine existenzielle Bedrohung für Nationalstaaten darstellen – Risiken, die so tiefgreifend sind, dass sie die derzei-

tige geopolitische Ordnung durcheinanderbringen oder sogar hinwegfegen könnten. Sie öffnen die Tür zu gewaltigen KI-gestützten Cyberangriffen, zu automatisierten Kriegen, die ganze Länder verwüsten könnten, zu künstlich erzeugten Pandemien und zu einer Welt, die unerklärlichen und doch scheinbar allmächtigen Kräften unterliegt. Die Wahrscheinlichkeit eines jeden dieser Ereignisse mag gering sein, aber die möglichen Folgen sind enorm. Selbst eine geringe Chance auf derartige Folgen erfordert deshalb unbedingte Aufmerksamkeit.

Einige Länder werden auf die Möglichkeit solcher katastrophalen Risiken mit einer Art technologisch aufgeladenem Autoritarismus reagieren, um die Ausbreitung dieser neuen Kräfte zu bremsen. Dies wird ein hohes Maß an Überwachung und ein massives Eindringen in unser Privatleben erfordern. Der Wunsch, die Technologie an die Kandare zu nehmen, könnte die Entwicklung hin zu einem dystopischen globalen Überwachungssystem befördern, bei dem alles und jeder ständig beobachtet wird und das alles gerechtfertigt wird mit dem Bestreben, sich vor den extremsten möglichen Folgen zu schützen.

Genauso plausibel ist eine Reaktion nach Art der Ludditen. Verbote, Boykotte und Moratorien wären dabei die Folge. Aber ist es überhaupt möglich, auf die Entwicklung neuer Technologien zu verzichten und eine Reihe von Moratorien einzuführen? Unwahrscheinlich. Angesichts ihres enormen geostrategischen und kommerziellen Werts ist es schwer vorstellbar, wie Nationalstaaten oder Unternehmen dazu gebracht werden können, einseitig auf die transformativen Kräfte zu verzichten, die durch diese bahnbrechenden Entwicklungen freigesetzt werden. Darüber hinaus ist der Versuch, die Entwicklung neuer Technologien zu verbieten, selbst ein Risiko: Technologisch stagnierende Gesellschaften sind historisch instabil und neigen zum Zusammenbruch. Irgendwann verlieren sie die Fähigkeit, Probleme zu lösen und Fortschritte zu machen.

Sowohl das Streben nach neuen Technologien als auch der Verzicht auf sie sind, so gesehen, mit Risiken behaftet. Die Chancen, sich auf einem «schmalen Pfad» durchzumogeln und das eine oder das andere Ergebnis zu vermeiden – einerseits eine techno-autoritäre Dystopie, andererseits eine durch Offenheit verursachte Katastrophe –, werden mit der Zeit immer geringer, je billiger, leistungsfähiger und allgegenwärtiger die Technologie wird und je größer die Risiken sind. Aber sich zurückzuziehen ist

auch keine Option. Mögen wir uns angesichts ihrer Risiken Sorgen machen, so brauchen wir die unglaublichen Vorteile der Technologien der kommenden Welle mehr als je zuvor. Das ist das zentrale Dilemma: dass eine mächtige Technologiegeneration die Menschheit früher oder später entweder in eine katastrophale oder in eine dystopische Richtung führt. Ich glaube, das ist das große Metaproblem des 21. Jahrhunderts.

In diesem Buch wird ausführlich gezeigt, warum diese schreckliche Zwickmühle unvermeidlich ist, und es wird danach gefragt, wie wir ihr begegnen können. Irgendwie müssen wir das Beste aus der Technologie herausholen – etwas, das für die Bewältigung einer Reihe gewaltiger globaler Herausforderungen unerlässlich ist – und gleichzeitig aus dem Dilemma herauskommen. Der derzeitige Kurs in Sachen Technologieethik und -sicherheit reicht dafür nicht. Trotz der vielen Bücher, Debatten, Blog-einträge und Tweetstorms über Technologie hört man selten etwas über deren *Eindämmung*. Ich verstehe darunter ein ineinandergreifendes Bündel technischer, gesellschaftlicher und rechtlicher Mechanismen, die die Technologie auf allen möglichen Ebenen einhegen und kontrollieren: also theoretisch ein Mittel, um dem Dilemma zu entgehen. Doch selbst die schärfsten Kritiker der Technologie neigen dazu, dieser Sprache des harten Containment auszuweichen.

Das muss sich ändern; ich hoffe, dieses Buch zeigt, warum, und gibt Hinweise, wie das gehen kann.

Die Falle

Ein paar Jahre nach der Gründung von DeepMind erstellte ich eine Power-Point-Präsentation über die potenziellen langfristigen Auswirkungen von KI auf Wirtschaft und Gesellschaft. In einem schicken Konferenzraum an der Westküste erklärte ich einem Dutzend der einflussreichsten Gründer, CEOs und Technologen der Techbranche, dass KI eine Vielzahl von Bedrohungen mit sich bringt, die proaktive Antworten erfordern. Sie könnte zu massiven Eingriffen in die Privatsphäre führen oder eine Apokalypse an Falschinformationen auslösen. Sie könnte als Waffe eingesetzt werden und eine Reihe tödlicher neuer Cyberwaffen schaffen, die unsere vernetzte Welt an ganz neuen Stellen empfindlich treffen können.

Ich wies auch auf das Potenzial von KI hin, eine große Zahl von Menschen arbeitslos zu machen. Ich forderte die Anwesenden auf, die lange Geschichte der Verdrängung von Arbeitskräften durch Automatisierung und Mechanisierung zu bedenken. Zuerst werden bestimmte Aufgaben effizienter erledigt, dann werden ganze Aufgabenbereiche überflüssig, und schon bald benötigen ganze Branchen in enormen Dimensionen weniger Arbeitskräfte. In den nächsten Jahrzehnten, so behauptete ich, werden KI-Systeme «geistiges Handwerk» in ähnlicher Weise ersetzen, und zwar lange bevor Roboter die körperliche Arbeit ersetzen. In der Vergangenheit wurden alte Jobs überflüssig, gleichzeitig aber auch neue geschaffen; was aber wäre, wenn KI die meisten dieser neuen Arbeitsplätze einfach mit übernehmen könnte? Es gebe, so führte ich weiter aus, eigentlich keinen Präzedenzfall für die neuen Formen der Machtkonzentration, die auf uns zukommen würden. Auch wenn sie sich weit entfernt anfühlten, würden potenziell schwerwiegende Bedrohungen auf die Gesellschaft zurollen.

Auf der abschließenden Folie zeigte ich ein Standbild aus der Serie *Die Simpsons*. In der Szene haben sich die Einwohner von Springfield erhoben, und die vertrauten Figuren ziehen mit Knüppeln und Fackeln bewehrt durch die Straßen. Die Botschaft war klar. «Die Mistgabeln kommen», sagte ich. Sie kämen unseretwegen, den Machern von Technologie. Es liege an uns, dafür zu sorgen, dass die Zukunft besser wird als so.

Reihum begegnete ich leeren Blicken. Die Anwesenden waren völlig ungerührt. Die Botschaft war nicht angekommen. Dann kamen die Einwände, und zwar zuhauf. Warum zeigten die Wirtschaftsindikatoren keine Anzeichen für das, was ich sagte? KI werde einen neuen Nachfrageschub auslösen, der neue Arbeitsplätze schaffen werde. Sie werde die Produktivität der Menschen steigern und sie in die Lage versetzen, noch produktiver zu sein. Vielleicht gebe es ein paar Risiken, räumten sie ein, aber die seien nicht allzu schlimm. Die Menschen seien schlau. Man habe noch immer Lösungen gefunden. *Keine Bange*, schienen sie zu denken, *weiter zur nächsten Präsentation*.

Einige Jahre später, kurz vor Ausbruch der Corona-Pandemie, nahm ich an einer renommierten Universität an einem Seminar über Technologierisiken teil. Das Ambiente war ähnlich: wieder eine große Runde, wieder eine anspruchsvolle Diskussion. Im Laufe des Tages wurden bei

Kaffee, Keksen und PowerPoint-Präsentationen eine Reihe von haarsträubenden Risiken diskutiert.

Eines stach besonders hervor. Der Vortragende wies darauf hin, dass der Preis für DNA-Synthetisierer, die maßgeschneiderte DNA-Stränge drucken können, rapide sank. Sie kosten inzwischen nur noch ein paar zehntausend Dollar und sind klein genug, um auf einer Werkbank in der Garage DNA zu synthetisieren, also *herzustellen*.¹ Und all das ist jetzt für jeden möglich, der über fortgeschrittene Kenntnisse in Biologie verfügt oder sich für eigenständiges Online-Lernen begeistern kann.

Angesichts der zunehmenden Verfügbarkeit der Werkzeuge skizzierte der Referent ein wahres Schreckensbild: Schon bald könnte jemand neuartige Krankheitserreger schaffen, die weitaus infektiöser und tödlicher sind als alles, was in der Natur vorkommt. Diese synthetischen Erreger könnten sich bekannten Gegenmaßnahmen entziehen, sich asymptomatisch verbreiten oder eine eingebaute Resistenz gegen Behandlungen aufweisen. Bei Bedarf könnte jemand die selbst gemachten Experimente mit online gekaufter und zu Hause neu zusammengesetzter DNA ergänzen. Die Apokalypse, im Internet bestellt und per Post geliefert.

Dies sei keine Science-Fiction, meinte der Referent, ein angesehener Professor mit mehr als zwei Jahrzehnten Erfahrung; es sei ein reales Risiko, und zwar schon heute. Er schloss mit einem alarmierenden Gedanken: Ein einziger Mensch habe heute wahrscheinlich «die Fähigkeit, eine Milliarde Menschen zu töten». Alles, was es dafür brauche, sei Motivation.

Die Anwesenden rutschten unruhig auf ihren Stühlen hin und her. Einige zuckten regelrecht zusammen und hüstelten nervös. Dann begannen das Gejammer und die Ausflüchte. Niemand wollte glauben, dass dies möglich war. Sicherlich sei das nicht der Fall, mit Sicherheit gebe es wirksame Kontrollmechanismen, sicherlich seien die Krankheiten schwer zu erzeugen, mit Sicherheit könnten die Datenbanken abgeschaltet werden, zweifellos lasse sich die Hardware sichern. Und so weiter.

Die kollektive Reaktion im Seminar war mehr als nur ablehnend. Die Leute weigerten sich schlicht, die Vision des Referenten zu akzeptieren. Niemand wollte sich mit den Auswirkungen der harten Fakten und nackten Möglichkeiten auseinandersetzen, denen sie gelauscht hatten. Ich sagte nichts und war, ehrlich gesagt, erschüttert. Kurz darauf war das Seminar zu Ende. An diesem Abend gingen wir alle gemeinsam zum Essen und unter-

hielten uns wie immer. Wir hatten gerade einen Tag lang über das Ende der Welt gesprochen, aber es gab immer noch Pizza zu essen, Witze zu erzählen, ein Büro, in das wir zurückkehren mussten, und außerdem würde schon irgendetwas auftauchen, oder die Argumentation war eben teilweise falsch. Auch ich machte mit.

Aber die Präsentation nagte noch Monate danach an mir. Warum habe ich, warum haben wir alle, das nicht ernster genommen? Warum entziehen wir uns peinlich berührt einer weiteren Diskussion? Warum werden manche boshaft und werfen Menschen, die diese Fragen aufwerfen, vor, sie würden eine Katastrophe heraufbeschwören oder «das unglaublich Gute» der Technologie übersehen? Diese weitverbreitete emotionale Reaktion, die ich beobachtet habe, ist etwas, das ich als die Falle der Pessimismus-Aversion bezeichne: die fehlgeleitete Analyse, die entsteht, wenn man von der Angst vor der Konfrontation mit potenziell düsteren Realitäten überwältigt wird, und die daraus resultierende Neigung, lieber wegzuschauen.

So gut wie jeder kennt diese Reaktion in irgendeiner Form, und sie hat zur Folge, dass wir eine Reihe kritischer Trends übersehen, die sich direkt vor unseren Augen entfalten. Es ist fast eine natürliche physiologische Reaktion. Unsere Spezies ist nicht dafür gemacht, sich wirklich mit einem Wandel dieses Ausmaßes auseinanderzusetzen, ganz zu schweigen von der Möglichkeit, dass die Technologie uns auf diese Weise enttäuschen und im Stich lassen könnte. Ich habe dieses Gefühl während meiner gesamten beruflichen Laufbahn erlebt, und ich habe gesehen, dass viele, viele andere dieselbe instinktive Reaktion zeigten. Die Auseinandersetzung mit diesem Gefühl ist eines der Leitthemen dieses Buches. Es geht darum, einen kalten, nüchternen Blick auf die Fakten zu werfen, auch wenn es unbequem ist.

Mit dieser Welle richtig umzugehen, die Technologie einzuhegen und sicherzustellen, dass sie stets der Menschheit dient, bedeutet, die Pessimismus-Aversion zu überwinden. Es bedeutet, sich der Realität dessen, was auf uns zukommt, ganz offen zu stellen.

Das vorliegende Buch ist mein Versuch, dies zu tun. Ich möchte die Konturen der kommenden Welle erkennen und beleuchten. Danach fragen, ob eine Eindämmung möglich ist. Die Dinge in einen historischen Kon-

text stellen und das größere Bild in den Blick bekommen, indem ich mich vom täglichen alarmistischen Gerede über Technik zurückziehe. Mein Ziel ist es, mich mit dem Dilemma auseinanderzusetzen und die zugrunde liegenden Prozesse zu verstehen, die die Entstehung von Wissenschaft und Technologie vorantreiben. Ich möchte diese Ideen so klar wie möglich für ein möglichst breites Publikum darstellen. Ich habe dieses Buch im Geiste der Offenheit und des Forscherdrangs verfasst: Ich mache Beobachtungen, spüre ihren Implikationen nach, bin aber auch offen für Widerlegungen und bessere Interpretationen. Ich wünsche mir nichts sehnlicher, als in diesem Punkt eines Besseren belehrt zu werden, dass also eine Einhegung ohne weiteres möglich ist.

Manche Leute erwarten von jemandem wie mir, einem Gründer von zwei KI-Unternehmen, verständlicherweise ein eher techno-utopisches Buch. Als Technologe und Unternehmer bin ich zwangsläufig Optimist. Ich erinnere mich, wie ich als Teenager völlig fasziniert war, als ich Netscape zum ersten Mal auf meinem Packard Bell 486er PC installierte. Ich war verzaubert von den surrenden Lüftern und dem verzerrten Pfeifen meines 56-Kbps-Einwahlmodems, das seine Hand nach dem World Wide Web ausstreckte und mich mit Foren und Chatrooms verband, die mir Freiheit verschafften und mir so viel beibrachten. Ich liebe die Technik. Sie ist *der* Motor des Fortschritts und ein Grund, stolz auf die Errungenschaften der Menschheit zu sein.

Aber ich glaube auch, dass diejenigen von uns, die die technologische Entwicklung vorantreiben, den Mut haben müssen vorherzusagen – und die Verantwortung dafür zu übernehmen –, wohin diese Entwicklung uns in den kommenden Jahrzehnten führen wird. Wir müssen endlich Vorschläge machen, was zu tun ist, wenn es so aussieht, als bestünde die Gefahr, dass die Technologie uns im Stich lässt. Es bedarf einer gesellschaftlichen und politischen Reaktion, nicht nur individueller Maßnahmen, aber diese Reaktion muss bei mir und meinen Kollegen beginnen.

Einige werden sagen, dass das alles übertrieben ist. Dass der Wandel viel schleichender vonstattengeht. Dass es sich nur um eine weitere Drehung des Hype-Zyklus handelt. Dass die Systeme zur Bewältigung von Krisen und Veränderungen eigentlich recht robust sind. Dass meine Sicht der menschlichen Natur viel zu düster ist. Dass die Bilanz der Menschheit, nun ja, zumindest bislang ganz gut ist. Die Geschichte ist voller falscher

Propheten und Unkenrufer, die sich geirrt haben. Warum sollte es dieses Mal anders sein?

Die Abneigung gegen Pessimismus ist eine emotionale Reaktion, eine tief verwurzelte instinktive Weigerung, die Möglichkeit ernsthaft destabilisierender Ergebnisse zu akzeptieren. Sie findet sich in der Regel bei Menschen in sicheren und mächtigen Positionen mit eingefahrenen Weltanschauungen, Menschen, die, oberflächlich betrachtet, mit Veränderungen umgehen können, sich aber schwertun, eine echte Infragestellung ihrer Weltordnung zu akzeptieren. Viele von denen, denen ich vorwerfe, dass sie in der Falle der Pessimismus-Aversion feststecken, teilen die wachsende Kritik an der Technologie voll und ganz. Aber sie nicken nur zustimmend, ohne tatsächlich etwas zu unternehmen. Wir schaffen das, sagen sie, so wie wir es schon immer geschafft haben.

Wer in technischen oder politischen Kreisen verkehrt, merkt schnell, dass die Standardideologie die des «Kopf-in-den-Sand-Steckens» ist. Wer etwas anderes glaubt und anders handelt, läuft Gefahr, vor lauter Angst vor und Empörung über gewaltige, unaufhaltsame Kräfte so gelähmt zu werden, dass sich alles sinnlos anfühlt. Die seltsame intellektuelle Halbwelt der Pessimismus-Aversion rumpelt also weiter. Ich muss es wissen, denn ich war zu lange darin gefangen.

In den Jahren seit der Gründung von DeepMind und seit diesen Vorträgen hat sich der Diskurs verändert – zumindest bis zu einem gewissen Grad. Die Debatte über die Automatisierung von Arbeitsplätzen wurde schon unzählige Male geführt. Eine globale Pandemie hat sowohl die Risiken als auch die Möglichkeiten der synthetischen Biologie aufgezeigt. In den Regulierungshauptstädten Washington, Brüssel und Peking kam es zu einer Art «Techlash», in dessen Zuge Kritiker in Kommentaren und Büchern gegen Technologie und Technologieunternehmen wetterten. Vormals randständige Ängste im Zusammenhang mit Technologie wurden zum Mainstream, die öffentliche Skepsis gegenüber Technologie nahm zu, und die Kritik aus Wissenschaft, Zivilgesellschaft und Politik wurde schärfer.

Doch angesichts der kommenden Welle und des großen Dilemmas und angesichts einer pessimismus-aversen Techno-Elite reicht das bei weitem nicht aus.

Die These

Wellen gibt es überall im menschlichen Leben. Diese hier ist nur die jüngste. Oft scheinen die Leute zu glauben, sie sei noch weit entfernt, sie klinge so futuristisch und absurd, dass sie nur die Domäne einiger weniger Nerds und Randdenker sei, eher Hype, Technogeschwätz, Überschwang. Das ist ein Irrtum. Die Welle ist real, so real wie ein Tsunami, der aus dem offenen blauen Ozean kommt.

Das Ganze ist nicht bloß Fantasie oder eine knifflige intellektuelle Übung. Selbst wenn Sie mit meinem Deutungsrahmen nicht einverstanden sind und nichts davon für wahrscheinlich halten, möchte ich Sie bitten weiterzulesen. Ja, ich komme aus der KI-Branche und bin dazu prädestiniert, die Welt durch eine technologische Brille zu betrachten. Ich bin voreingenommen, wenn es um die Frage geht, ob dies von Bedeutung ist. Dennoch bin ich, nachdem ich diese Revolution in den letzten anderthalb Jahrzehnten hautnah miterlebt habe, davon überzeugt, dass wir an der Schwelle zur wichtigsten Transformation unseres Lebens stehen.

Als jemand, der diese Technologien entwickelt hat, bin ich der festen Überzeugung, dass sie außerordentlich viel Gutes bewirken, unzählige Leben zum Besseren verändern und grundlegende Herausforderungen bewältigen können – von der Erschließung der nächsten Generation sauberer Energie bis zur Entwicklung preiswerter und wirksamer Therapien für unsere am schwersten heilbaren Krankheiten. Technologien können und sollten unser Leben bereichern; in der Vergangenheit haben die Erfinder und Unternehmer, die hinter ihnen stehen, den Fortschritt vorangetrieben und den Lebensstandard von Milliarden von Menschen verbessert.

Aber ohne Einhegung ist jeder andere Aspekt der Technologie, jede Diskussion über ihre ethischen Unzulänglichkeiten oder die Vorteile, die sie bringen könnte, irrelevant. Wir brauchen dringend hieb- und stichfeste Antworten darauf, wie die kommende Welle kontrolliert und eingedämmt werden kann, wie die Schutzmechanismen und die Möglichkeiten des demokratischen Nationalstaats aufrechterhalten werden können, aber im Moment hat niemand einen solchen Plan. Es ist eine Zukunft, die niemand von uns will, aber ich fürchte, sie wird immer wahrscheinlicher, und ich werde in den folgenden Kapiteln erklären, warum.

In Teil I werfen wir einen Blick auf die lange Geschichte der Technologie und ihrer Ausbreitung – Wellen, die sich über Jahrtausende hinweg aufbauen. Was treibt sie an? Was macht sie wirklich allgemein? Wir fragen zudem, ob es Beispiele für Gesellschaften gibt, die bewusst Nein zu einer neuen Technologie sagen. Statt von einer Abwendung von Technologien ist die Vergangenheit durch ein ausgeprägtes Muster der Proliferation gekennzeichnet, das zu ausufernden Ketten beabsichtigter und unbeabsichtigter Folgen geführt hat.

Ich nenne dies das «Containment-Problem». Wie können wir die wertvollsten Technologien, die je erfunden wurden, in den Griff bekommen, wenn sie billiger werden und sich schneller verbreiten als alle anderen in der Geschichte?

Teil II befasst sich mit den Einzelheiten der kommenden Welle. Im Mittelpunkt stehen zwei enorm vielversprechende, leistungsstarke und gefährliche Allzwecktechnologien: künstliche Intelligenz und synthetische Biologie. Beide kündigen sich schon seit langem an, aber dennoch glaube ich, dass das Ausmaß ihrer Auswirkungen oft noch unterschätzt wird. Um sie herum entwickelt sich eine Vielzahl von verwandten Technologien wie Robotik und Quantencomputing, deren Entwicklung sich auf komplexe und turbulente Weise überschneiden wird.

In diesem Abschnitt gehen wir nicht nur der Frage nach, wie diese Technologien alle entstanden sind und was sie leisten können, sondern auch, warum sie so schwer zu bändigen sind. Die verschiedenen Technologien, von denen ich hier spreche, haben vier Hauptmerkmale gemeinsam, die erklären, warum wir es in ihrem Fall nicht mit «business as usual» zu tun haben: Sie sind von Natur aus allgemein und daher vielseitig einsetzbar, sie entwickeln sich rasant weiter, sie haben asymmetrische Auswirkungen, und sie sind in mancherlei Hinsicht zunehmend autonom.

Ihre Entwicklung wird durch starke Anreize vorangetrieben: geopolitischen Wettbewerb, massive finanzielle Belohnungen und eine offene, dezentrale Forschungskultur. Eine Vielzahl staatlicher und nichtstaatlicher Akteure wird sie ungeachtet der Bemühungen um Regulierung und Kontrolle vorantreiben und dabei Risiken eingehen, die jeden betreffen, ob es uns gefällt oder nicht.

Teil III erkundet die politischen Auswirkungen einer kolossalen Umverteilung der Macht, die durch eine ungebremste Welle hervorgerufen

wird. Der Grundpfeiler unserer heutigen politischen Ordnung – und der wichtigste Akteur bei der Eindämmung von Technologien – ist der Nationalstaat. Er ist bereits durch Krisen erschüttert und wird durch eine Reihe von Schocks, die durch die Welle noch verstärkt werden, weiter geschwächt werden: das Potenzial für neue Formen von Gewalt, eine Flut von Falschinformationen, verschwindende Arbeitsplätze und die Aussicht auf katastrophale Unfälle.

Im weiteren Verlauf wird die Welle eine Reihe von tektonischen Machtverschiebungen erzwingen, die gleichzeitig zentralisierend und dezentralisierend wirken. Dadurch werden riesige neue Unternehmen entstehen, der Autoritarismus wird gestärkt werden, und gleichzeitig werden Gruppen und Bewegungen ermächtigt, außerhalb der traditionellen Gesellschaftsstrukturen zu leben. Das heikle Geschäft des Nationalstaats wird gerade dann unter immensen Druck geraten, wenn wir Institutionen wie die seinen am meisten brauchen. Auf diese Weise geraten wir in ein Dilemma.

In Teil IV geht es um die Frage, was wir dagegen tun können. Gibt es überhaupt eine kleine Chance der Eindämmung, eine Möglichkeit, uns aus dem Dilemma herauszuwinden? Und wenn ja, wie soll das gehen? In diesem Schlussabschnitt werden zehn Schritte skizziert, die sich von der Ebene des Codes und der DNA bis zur Ebene internationaler Verträge vorarbeiten und ein festes, verschachteltes Gefüge von Beschränkungen, eine Art Rahmenplan für die Eindämmung bilden.

Dieses Buch ist nicht zuletzt eine Auseinandersetzung mit dem Versagen. Technologien können im ganz banalen Sinne versagen: Der Motor springt nicht an, die Brücke stürzt ein. Aber sie können auch in einem weiteren Sinne versagen. Wenn Technologie Menschenleben beschädigt oder Gesellschaften voller Leid hervorbringt oder sie unregierbar macht, weil wir einen chaotischen Rattenschwanz von üblen (oder unbeabsichtigt gefährlichen) Akteuren ermächtigen – wenn, kurz gesagt, Technologie insgesamt schädlich ist –, dann kann man sagen, dass sie in einem anderen, tieferen Sinne versagt hat, dass sie ihr Versprechen nicht erfüllt hat. Das Scheitern in diesem Sinne ist nicht der Technologie als solcher zuzuschreiben, sondern hängt mit dem Kontext zusammen, in dem sie funktioniert, mit den Regierungsstrukturen, denen sie unterworfen ist, mit den Netzwerken der Macht und der Nutzung, denen sie unterliegt.

Der beeindruckende Erfindungsreichtum, der heute so viele Dinge hervorbringt, bedeutet, dass wir besser in der Lage sind, die erste Art von Versagen zu vermeiden. Es stürzen weniger Flugzeuge ab, Autos sind sauberer und sicherer, Computer sind leistungsfähiger und dennoch sicherer. Die große Herausforderung besteht darin, dass wir immer noch nicht mit der zweiten Art des Versagens rechnen.

Im Laufe der Jahrhunderte hat die Technologie das Wohlergehen von Milliarden von Menschen dramatisch verbessert. Dank der modernen Medizin sind wir unermesslich gesünder, die Mehrheit der Welt lebt in einem Überfluss an Nahrungsmitteln, die Menschen waren noch nie so gebildet, so friedlich und materiell so gut gestellt. Dies sind entscheidende Errungenschaften, die zum Teil auf den großen Antriebsmotor der Menschheit zurückzuführen sind: die Wissenschaft und die Entwicklung der Technik. Aus diesem Grund habe ich mein Leben der sicheren Entwicklung dieser Instrumente gewidmet.

Aber jeder Optimismus, den wir aus dieser außergewöhnlichen Geschichte schöpfen, muss in der nackten Realität gründen. Sich gegen das Scheitern zu wappnen heißt, zu verstehen, was schiefgehen kann, und sich schließlich damit auseinanderzusetzen. Wir müssen die Gedankenkette bis zu ihrem logischen Endpunkt verfolgen, ohne Angst davor zu haben, wohin das führen könnte, und, wenn wir dort ankommen, etwas dagegen tun. Die kommende Technologiewelle droht schneller und in größerem Umfang zu versagen als alles, was die Menschheit bisher erlebt hat. Angesichts dieser Situation bedarf es weltweiter, breiter Aufmerksamkeit. Es braucht Antworten – Antworten, die noch niemand hat.

Eine Eindämmung ist auf den ersten Blick nicht möglich. Und doch *muss* eine Eindämmung um unser aller willen möglich sein.