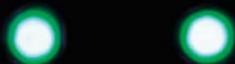


MANUELA LENZEN

KÜNSTLICHE INTELLIGENZ



Was sie
kann & was uns
erwartet

C·H·Beck

C·H·Beck

PAPERBACK

MANUELA LENZEN

KÜNSTLICHE INTELLIGENZ

WAS SIE KANN

&

WAS UNS ERWARTET

C.H.BECK

Zum Buch

Der Klassiker der Wissenschaftsjournalistin in der 4., aktualisierten Auflage

„Manuela Lenzen räumt mit Mythen auf und beschreibt alles, was man derzeit über Künstliche Intelligenz wissen muss.“ – Dana Heide, Handelsblatt

„Ein Wegweiser durch die Zukunft, die längst schon begonnen hat, sollte zur Pflichtlektüre auch in den Schulen werden.“ – Harald Raab, Rhein-Neckar-Zeitung

„Am Ende lässt dieses feine Buch keine Fragen offen.“ – PM

Über die Autorin

Manuela Lenzen hat in Philosophie promoviert und schreibt als freie Wissenschaftsjournalistin über Digitalisierung, Künstliche Intelligenz und Kognitionsforschung, u. a. für die Frankfurter Allgemeine Zeitung. Bei C.H.Beck ist von ihr jüngst erschienen: *Der elektronische Spiegel. Menschliches Denken und künstliche Intelligenz.*

1. Auflage. 2018

2. Auflage. 2018

3. Auflage. 2019

4., aktualisierte Auflage. 2023

© Verlag C.H.Beck oHG, München 2018

Alle urheberrechtlichen Nutzungsrechte bleiben vorbehalten. Der Verlag behält sich auch das Recht vor, Vervielfältigungen dieses Werks zum Zwecke des Text and Data Mining vorzunehmen.

Umschlaggestaltung: Rothfos & Gabler, Hamburg

Umschlagabbildung: © plainpicture/ganguin

Satz: C.H.Beck.Media.Solutions, Nördlingen

ISBN Buch 978 3 406 80663 6

ISBN epub 978 3 406 81282 8

ISBN ePDF 978 3 406 81283 5

Die gedruckte Ausgabe dieses Titels erhalten Sie im Buchhandel
sowie versandkostenfrei auf unserer Website

www.chbeck.de.

Dort finden Sie auch unser gesamtes Programm und viele weitere
Informationen.

INHALT

Einleitung

Verwirrende neue Technikwelt 9

TEIL 1

Was Künstliche Intelligenz kann 23

1. Was ist Künstliche Intelligenz? 23

Optimistische Anfänge, diverse Winter und ein neuer Boom 23

Wer blufft, gewinnt: der Turing-Test 27

Noch einmal: Wann ist eine Maschine intelligent? 29

Die drei Ziele der Künstlichen Intelligenz 33

2. Wie Maschinen denken 37

Denken nach Vorschrift 37

Der Computer und das Gehirn 40

Die gar nicht so geheimnisvolle Welt der Algorithmen 44

Lösungen suchen und bewerten 46

Keine Datenverarbeitung ohne Daten 48

Wie Maschinen lernen 50

Können Programme sich selbst verbessern? 70

Watson oder Cyc? 71

Sisyphos lässt grüßen: Nachvollziehen, was eine Maschine denkt 78

3. Die Roboter kommen 84

Intelligenz braucht einen Körper 86

Auch Roboter müssen lernen 91

Führt der Weg zu Künstlicher Intelligenz durch eine künstliche Kindheit? 95

Die Evolution als Konstrukteur 98
Die Vielfalt der Roboter 101

4. Warum Maschinen Spiele spielen 111

5. Grenzüberschreitungen 114

Viel Mensch, wenig Maschine: der Cyborg 114
Die Experimente der Body-Hacker 116
Gehirn-Computer-Schnittstellen 118
Viel Maschine, ein wenig Tier: biohybride Roboter 121

6. Die letzten Bastionen 124

Malen, Dichten, Komponieren:
die Maschine als Künstler 124
Mythos Autonomie 128
Gefühle und Bewusstsein 133
Stolperstein Verstehen 139
Elektronische Personen? Roboter und das Recht 143
Moralische Maschinen 146

TEIL 2

Was uns erwartet 151

7. Künstliche Intelligenz verändert die Wissenschaft 152

Mehr Durchblick 152
Automatisierte Forschung 153
Hören wir auf, die Welt zu verstehen? 155
Besser als der Arzt? 158

8. Die Algorithmisierung der sozialen Welt 165

Daten sind das neue Öl 165
Algorithmen sortieren die Welt 167
Wenn Maschinen über Menschen urteilen 177
Was also tun? 180

9. Alles vernetzt 186

Das Internet der Dinge 186
Home, smart Home 194

Die kluge Stadt 196
Der gläserne Mensch 198

10. Die Umwälzung der Arbeitswelt 201

Unsichere Prognosen 201
Industrie 4.0 bringt Arbeit 4.0 206
Was zu tun übrig bleibt 208
Was wird aus der Arbeitswelt? 211

11. Künstliche Intelligenz und das Militär 216

Die dritte Revolution der Kriegstechnik 216
Viele offene Fragen 221
Sind Maschinen die besseren Soldaten? 224

12. Pseudogefährten 228

Unsere intelligenten Assistenten 228
Soziale Roboter 230
Das unheimliche Tal 234
Gib dich nicht mit Androiden ab! 235
Pseudogefährten 237

13. Künstliche Intelligenz für eine menschlichere Welt 242

Sind wir zu dumm für unsere komplexe Welt? 242
Künstliche Intelligenz an die Macht? 245
Wie gefährlich ist die Künstliche Intelligenz? 249
(M)eine Utopie 254

Danksagung 257

Anmerkungen 258

Zum Weiterlesen 270

EINLEITUNG

VERWIRRENDE NEUE TECHNIKWELT

Sie ist angekommen. Künstliche Intelligenz ist nicht länger eine ferne Vision. Längst stehen KI-gestützte Übersetzungsprogramme kostenlos im Internet zur Verfügung, Schülerinnen und Schüler lassen Chatbots ihre Hausaufgaben machen, KI-generierte Bilder fluten die sozialen Medien, mit einem Bot zu telefonieren ist fast schon Alltag. Auf den Haushaltsroboter, der für uns putzt und aufräumt, warten wir allerdings bis heute, und auch die selbstfahrenden Autos sind trotz intensiver Forschung und hoher Investitionen noch immer nicht unterwegs. Dennoch reibt sich der Mensch angesichts der rasanten Entwicklungen der letzten Jahre ein wenig verwundert und verwirrt die Augen. Das sei alles nur Mathematik, sagen die einen, diese Systeme seien stochastische Papageien, die kein Wort von dem, was sie von sich geben, verstehen. Das könne nicht sein, sagen andere, die Systeme zeigten ganz klar Zeichen von Verständnis und manchmal vielleicht sogar von Bewusstsein. Das seien nach wie vor Werkzeuge wie andere auch, heißt es auf der einen Seite, das sei ganz klar ein Schritt hin zu einer übermächtigen und sich verselbständigenden Technik, sagen die anderen. Auf jeden Fall ist die Künstliche Intelligenz längst zu einer komplexen Herausforderung geworden: von der basalen Frage, womit wir es hier eigentlich genau zu tun haben, über die gesellschaftliche Herausforderung, zu bestimmen, wie und wo welche Systeme zum Einsatz kommen sollen, bis hin zu Überlegungen, wohin sich die Gesellschaft mit dieser Technik entwickeln wird.

Computersysteme analysieren heute medizinische Daten, sie sagen die komplexe Faltungsstruktur der Proteine voraus, managen den Aktienhandel und steuern Waffensysteme. Sie malen, dichten, übersetzen und komponieren. Sie finden Muster, wo wir nur ein Rauschen sehen. Roboter begrüßen uns im Hotel, im Supermarkt und am Flughafen, sie führen uns durchs Museum und rufen an, wenn sie mit einem Päckchen vor der Tür stehen. Sie pflücken Gurken, beaufsichtigen Kühe, zapfen Bier, mähen Rasen, putzen Fenster, montieren Handys, spielen Dudelsack oder Fußball, braten Burger und schnipseln den Salat dazu.

Die Künstliche Intelligenz boomt. Derzeit vergeht keine Woche, in der nicht ein neuer Roboter, ein neues generatives Programm, ein neues smartes Gadget auf den Markt kommt, kaum ein Tag ohne Schlagzeilen, die eine kommende oder schon in vollem Gange befindliche Revolution beschwören: durch die Künstliche Intelligenz. Immer mehr kluge Maschinen, virtuelle und reale, riesengroße und winzig kleine, fliegende, laufende, rollende und schwimmende, harte und weiche, niedliche und erschreckende, verlassen die Labors und finden Eingang in unser Leben, unser Arbeiten, unsere Kommunikation, in unsere Körper, unser Denken, unser Weltbild. Und obwohl oder vielleicht gerade weil diese Technik intensiver vorgedacht worden ist als jede andere, von den mit Wasser- oder Luftdruck betriebenen Automaten der Antike über den mittelalterlichen Riesen Golem bis zur Science-Fiction, bringt sie uns zurzeit gehörig aus der Fassung.

Eigentlich mögen wir unsere intelligente Technik. Wer wollte schon ohne sein Smartphone auskommen? Erinnert sich noch jemand an (schlecht riechende) Telefonzellen, vor denen sich bisweilen lange Schlangen bildeten? Möchte jemand auf das Internet mit seinen Suchmaschinen verzichten und stattdessen per Postkarte Informationsbroschüren anfordern, wie es vor nicht allzu langer Zeit üblich war? Und wie hat man vor der Erfindung der Navigationsgeräte eigentlich im Dunkeln durch eine fremde Stadt gefunden?

Die intelligente Technik ist praktisch und Roboter sind faszinierend. Jeder Roboterforscher kennt das Phänomen: Kaum öffnet ein auch nur entfernt an ein menschliches Vorbild erinnernder

Roboterkopf mit Papierrohren und Lippen aus roten Gummischläuchen knirschend seine Kameraaugen, konkurrieren die Menschen um seine Aufmerksamkeit. Keine Schulklasse, die nicht in Begeisterungstürme ausbräche, wenn der Lehrer den Dino-Roboter Pleo oder gar den gelenkigen kleinen Humanoiden Nao aus dem Koffer holt. Kein Einkaufszentrum, in dem der fahrende Service-roboter nicht von Neugierigen umringt oder von einem Schwarm Kinder verfolgt würde, kein Roboterfußballmatch, bei dem die Zuschauer nicht mitfieberten, wenn der Stürmer in Zeitlupe den Ball ins Visier nimmt, surrend den Fuß hebt, schießt und vor lauter Schwung gleich hintenüber fällt. Auch Peter und Petra, robotische Museumsführer des Paderborner Heinz-Nixdorf-Computermuseums, sind gut beschäftigt: Sie führen Besucherinnen und Besucher zu verschiedenen Ausstellungsstationen und halten Kinder mit Versteckspielen bei Laune. Kaum hat man sich umgedreht, sind sie surrend zwischen den Exponaten verschwunden.

Intelligente Maschinen, die alle schmutzige, gefährliche, langweilige, gesundheitsschädliche Arbeit erledigen und den Menschen Zeit lassen, sich den interessanten, kreativen und angenehmen Seiten des Lebens zu widmen, sind ein alter Menschheitstraum. Die Visionen der intelligenten Häuser und Städte der Zukunft erinnern nicht zufällig an (modernisierte) Vorstellungen vom Paradies: Mithilfe klug analysierter Datenmengen, Simulationen zukünftiger Entwicklungen und selbstlernender Algorithmen werden wir unsere Umwelt-, Verkehrs-, Energie- und Müllprobleme lösen, Hunger und Krankheiten besiegen und den Klimawandel in den Griff bekommen. Die miteinander kommunizierenden intelligenten Systeme werden den Alltag wie von selbst organisieren, sie werden unsere Wünsche und Befindlichkeiten eher erkennen als wir selbst und uns jederzeit perfekt umsorgen. Durch Künstliche Intelligenz werden wir uns mit Menschen verständigen können, deren Sprache wir nicht sprechen, Wissen und Kommunikation werden ohne Grenzen fließen und uns unsere globalisierte Welt immer besser verstehen lassen. Künstliche Intelligenz wird die weltweite Produktivität steigern und damit Wirtschaftswachstum und Wohlstand für alle ermöglichen. Zumindest wird sie die Produktion verbilligen, Industrien vor dem Abwandern in Niedrig-

lohnländer bewahren und die Produktivität einer alternden Industriegesellschaft aufrechterhalten.¹ Intelligente Implantate und Prothesen werden Menschen mit Behinderungen helfen, alte Menschen werden dank intelligenter Kalender, Waschbecken und Küchen länger selbstbestimmt in ihren eigenen vier Wänden leben können. Das Autofahren wird sicherer und komfortabler, Unterricht und Weiterbildung auf jeden Menschen individuell zugeschnitten. Und auf lange Sicht werden die intelligenten Maschinen nicht einfach Maschinen für dieses oder jenes sein, sondern Universalisten: Für manche Visionäre ist die Künstliche Intelligenz das Problem, das zu lösen alle unsere anderen Probleme lösen wird.²

Doch so richtig freuen können wir uns über die Fortschritte der intelligenten Technik trotzdem nicht. «Computer, erzähl mir alles über die Borgs!»: Während Raumschiff-Enterprise-Fans in Amazons Kommunikationssystem Echo endlich realisiert sahen, wovon sie jahrzehntelang nur träumen konnten, ließen Klagen nicht lange auf sich warten. Bald forderte die Polizei erste Echo-Daten für ihre Ermittlungen an,³ gab es erste Klagen über unerwünscht ins Wohnzimmer gesendete Werbebotschaften des Mitbewerbers GoogleHome.⁴ Facebook machte Schlagzeilen mit einem Algorithmus, der Suizidgefährdete erkennen und ihnen Hilfe anbieten sollte,⁵ während Verbraucherschützer kritisierten, der Konzern verkaufe Daten über die psychische Verfassung seiner Nutzer an Werbekunden. Während das Europäische Parlament noch diskutierte, ob künstliche Intelligenzen als elektronische Personen gelten sollten,⁶ verbot die Bundesnetzagentur die «kluge» Puppe My Friend Cayla, weil sie, kaum gesichert, alles, was im Kinderzimmer gesprochen wurde, aufzeichnete und in der Cloud des Betreibers speicherte und damit jeglichem Schutz der Privatsphäre Hohn sprach.⁷ Im Juni 2023 hat die EU-Kommission im Rahmen ihrer Digitalstrategie den Artificial Intelligence Act, einen Gesetzentwurf über künstliche Intelligenz, beschlossen, um die Verwendung «intelligenter» Systeme zu regulieren und Missbrauch vorzubeugen (<https://artificialintelligenceact.eu/>). Grundlegende Probleme wie der Umgang mit Hassreden im Internet sind noch immer nicht gelöst, nach wie vor wird der Einfluss von

Chatbots auf Wahlkämpfe diskutiert, rätseln die Betreiber sozialer Medien, wie man Fake News identifizieren und die Verbreitung von Bildern entsetzlicher Verbrechen verhindern kann. Seit der Veröffentlichung der großen Sprachmodelle verunsichern immer neue Schätzungen über den Prozentsatz der Arbeitsplätze, die intelligente Maschinen bald übernehmen könnten.

Natürlich, Maschinen waren schon immer dazu da, etwas besser zu können als der Mensch: Lasten heben, Löcher bohren, Korn dreschen, rechnen. Doch inzwischen erinnert das Verhältnis von Mensch und Maschine an ein Rückzugsgefecht: Schach, Jeopardy, Go, sogar Poker, Pac-Man und vieles andere spielen Computerprogramme inzwischen besser als der Mensch. Und irgendwo zwischen dem Taschenrechner und Watson, dem System aus dem Haus IBM, das 2011 das amerikanische Fernsehquiz Jeopardy gewann, eine Zeitlang Mediziner bei der Auswahl von Therapien unterstützte und heute hilft, Geschäftsmodelle zu optimieren, haben wir angefangen, uns Sorgen zu machen: Ist es gut, wenn Maschinen sich in immer mehr Bereichen bewähren, von denen wir die längste Zeit dachten, sie seien dem Menschen vorbehalten? Was kommt da auf uns zu?⁸

Neben die paradiesischen Visionen der schönen neuen intelligenten Technikwelt sind längst Schreckbilder getreten: von Künstlichen Intelligenzen, die jeden Klick, den wir im Internet tun, jedes Wort, das wir in Gegenwart der klugen Lautsprecher wechseln, jeden Schritt, der uns an einer intelligenten Kamera vorbeiführt, analysieren und mit den Ergebnissen Regierungen und ein paar großen Konzernen ermöglichen, uns zu überwachen und zu manipulieren, während ihr eigenes Treiben für uns immer undurchschaubarer wird; Schreckbilder von Maschinen, die unsere Arbeitsplätze übernehmen, den Reichtum einiger weniger mehren und den großen Rest dazu verdammen, sich, bestenfalls versorgt mit einem minimalen Grundeinkommen, in virtuellen Welten zu langweilen; Schreckbilder von alten Menschen, die Maschinen mit Fellüberzug streicheln, und Kindern, denen ein Roboter die Gutenachtgeschichte vorliest, weil niemand mehr Zeit für sie hat; Schreckbilder von Systemen, die sich selbst über menschliches Maß hinaus verbessern, sich unserer Kontrolle entziehen, uns im besten

Fall in den Kaninchenstall sperren, weil wir ihnen zu dumm sind, und im schlimmsten zu Büroklammern verarbeiten, einfach, weil sie darauf programmiert sind, immer mehr davon herzustellen.

Die Künstliche Intelligenz könnte das Beste oder das Schlechteste sein, was der Menschheit je zugestoßen ist, sagte der Physiker Stephen Hawking 2014 in einem BBC-Interview.⁹ Und es scheint, als seien diejenigen, die auf das Beste hoffen, in der Minderheit und ihre Visionen zudem nicht minder beängstigend. Die Entwicklung einer ultra-intelligenten Maschine könne der Menschheit auf ewig das Überleben sichern, vorausgesetzt, sie wäre gutmütig genug, uns wissen zu lassen, wie wir sie unter Kontrolle halten können, schrieb der britische Mathematiker Irving John Good, der zusammen mit Alan Turing im Zweiten Weltkrieg den Code der Verschlüsselungsmaschine Enigma knackte, Mitte der 1960er Jahre.¹⁰ Ray Kurzweil, Informatiker, Erfinder, Futurist und Google-Forschungsdirektor, hofft nicht nur auf das Überleben der Menschheit, sondern auf die Unsterblichkeit des Individuums. Durch das Zusammengehen von Künstliche-Intelligenz-Forschung, Robotik, Genetik und Nanotechnologie werde die Wissenschaft einen gewaltigen qualitativen Sprung tun, auf den dann eine Entwicklung folgt, die wir mit unseren beschränkten intellektuellen Kräften nicht mehr nachvollziehen, geschweige denn steuern können. Diesen Sprung nennt Kurzweil mit einem Begriff des amerikanischen Mathematikers, Informatikers und Science-Fiction-Autors Vernor Vinge Singularität. Danach werde der Mensch mit der Technologie verschmelzen, was ihm ein viel längeres und gesünderes Leben ermöglichen werde, mit kognitiven Fähigkeiten, von denen wir uns gar keinen Begriff machen können. Kurzweil traut der Forschung sogar zu, eines nicht allzu fernen Tages – er (Jahrgang 1948) geht davon aus, es selbst zu erleben – die Sterblichkeit zu überwinden. Aus dem Menschen wird dann allerdings ein Cyborg, ein Mensch-Maschine-Mischwesen, geworden sein.

Die pessimistische Vision formulierte zum Beispiel der erwähnte Vernor Vinge Anfang der 1990er Jahre: «Innerhalb von dreißig Jahren werden wir die technologischen Mittel haben, übermenschliche Intelligenz herzustellen. Kurz darauf wird die Ära der Menschheit zu Ende sein.» Denn diese übermenschlichen In-

telligenzen seien nicht mehr unsere Werkzeuge, ebenso wenig wie wir die Werkzeuge von Hasen, Rotkehlchen oder Schimpansen sind.¹¹ Ähnlich düstere Prognosen, wenn auch mit einem etwas weiteren Zeithorizont, haben in den letzten Jahren unter anderem Elon Musk, der Chef des Elektroauto-Konzerns Tesla, OpenAI-Chef Sam Altman und der renommierte KI-Forscher Geoffrey Hinton formuliert.

Diese düstere Vision findet sich häufig auch in den Schlagzeilen: Begrüßen Sie Ihre neuen Herren! Evolution ohne uns! Wird Künstliche Intelligenz uns töten? Das Bild des Terminators, jenes furchterregenden Cyborgs aus dem gleichnamigen Kultfilm der 1980er Jahre, zierte eine Zeitlang sehr viele Texte zur KI. Dann verschwand er für eine Weile, um mit ChatGPT und seinen Verwandten zurückzukehren.

Die schöne neue Technikwelt präsentiert sich derzeit also vor allem verwirrend. Berechtigte Sorgen, Science-Fiction, ein gehöriger Schuss Lust an der Katastrophe und vor allem ganz verschiedene technische Dinge gehen dabei oft mächtig durcheinander. Sie machen die Künstliche Intelligenz zu einem Heilsbringer oder einem Angst-Gegner.

Beide lähmen, erscheinen aber auch größer, als sie sind. Und zumeist hilft ein genauerer Blick, um sie auf Normalmaß zu reduzieren. Dazu möchte ich mit diesem Buch einladen. Ich möchte den vielen Ausrufezeichen, die die Debatte um die Künstliche Intelligenz begleiten, nur eins hinzufügen, und das gleich zu Beginn: Kommt mal wieder auf den Teppich! Die Künstliche Intelligenz boomt, doch ganz so schnell wie die spektakulären Nachrichten aufeinanderfolgen, ist sie dann doch nicht. Weder Roboter noch Algorithmen, jene digitalen Verfahren, die Maschinen erst klug machen, haben dunkle Machtinstinkte. Und trotz ihrer beeindruckenden Leistungen, sei es beim Pokern oder beim Autofahren, sind die Produkte der KI-Forschung auf absehbare Zeit weit davon entfernt, uns in den Kaninchenstall zu sperren. Noch haben sie ganz andere Probleme, zum Beispiel zu prüfen, ob eine Aussage wahr ist, einen ungewohnten Gegenstand zu erkennen oder eine kompliziertere Situation richtig zu verstehen, eine Spülmaschine einzuräumen, ohne den menschlichen Zuschauer mit

ihrem Schneckentempo in den Wahnsinn zu treiben, oder sich auf den Beinen zu halten, wenn die Tür aufgeht, nachdem sie die Klinke gedrückt haben. Und wenn sie die eine Aufgabe bewältigen, scheitern sie mit einiger Sicherheit an der nächsten.

Der Eindruck, dass wir mit der Künstlichen Intelligenz derzeit von einer Entwicklung überrollt werden, die von nichts und niemandem aufzuhalten ist, ist vermutlich nicht falsch. Doch diese Entwicklung hat ihre Ursache nicht in geheimnisvollen Mächten künstlicher Intelligenzen, sondern darin, dass Menschen sich etwas davon versprechen. Künstliche Intelligenz ist der ganz große Zukunftsmarkt.¹² Staaten, Unternehmen sowie zivile und militärische Forschungsinstitutionen auf der ganzen Welt investieren riesige Summen in die einschlägige Forschung. «AI first», KI zuerst, formulierte etwa Google-Chef Sundar Pichai 2016 die Zielrichtung des Konzerns. Forscher, Unternehmer, Investoren und Militärs hoffen auf neue Märkte, neue Möglichkeiten des Erkenntnisgewinns, neue Waffen und Kontrollmöglichkeiten, auf Effizienzgewinne durch die Einsparung oder Ergänzung menschlicher Arbeitskraft. Kein Unternehmen, keine Armee, keine Volkswirtschaft, keine Regierung, kein Geheimdienst und keine Universität will sich bei diesem Rennen abhängen lassen.

Das macht es nicht unbedingt besser. Technologien müssen nicht klug sein, um gefährlich werden zu können oder die Gesellschaft massiv zu verändern, den sozialen Frieden oder die Demokratie zu bedrohen. Aber es macht deutlicher, worum es geht: Das Stück, das gerade gespielt wird, heißt nicht «Robocalypse» oder «Die Invasion der Superintelligenzen», sondern, mal wieder, immer noch: Wirtschaftswachstum, Gewinnmaximierung, Markt, Machterhalt. Mein Hintergrund ist die Philosophie und ich mag die großen Fragen rund um Singularität, Superintelligenzen und Was-wäre-wenn, aber ich fürchte, ich muss Sie und mich enttäuschen: Unsere Probleme sind die, die schon Aldous Huxley in *Schöne neue Welt* und George Orwell in 1984 beschrieben haben: Macht, Überwachung, die Verteidigung der Freiheit, plus einige neue, etwa die Stabilität der Infrastruktur.

Vielleicht können wir gar nicht anders als bei jedem Roboter und jedem Chatbot all die Science-Fiction mitzudenken, die wir

gelesen oder gesehen haben und in der die gewalttätigen, außer Kontrolle geratenen, manipulativen künstlichen Intelligenzen die Zahl der netten, nützlichen und hilfsbereiten bei Weitem übersteigt. Vielleicht ist auch die Idee einer übermächtigen, aber immerhin von uns selbst geschaffenen Superintelligenz so faszinierend, dass wir immer wieder daran hängen bleiben. Und zu dieser Geschichte gehört, als Strafe für den allzu vorwitzigen Zauberlehrerling, dass sein Geschöpf sich seiner Kontrolle entzieht und Unheil stiftet. Doch auch hier gilt: Kommt mal wieder auf den Teppich! Wesen mit menschlicher Intelligenz in die Welt zu setzen ist uns auf absehbare Zeit nur auf einem Weg gegeben: durch Fortpflanzung.

Wir gehörten tatsächlich in den Kaninchenstall, wenn wir jetzt erstarrt auf die KI-Schlange starren, wie es die bekannte Redensart dem Nagetier andichtet. Denn wir haben es mit einer Technologie zu tun, deren Entwicklung nach ersten Anfängen in den 1940er Jahren gerade erst richtig Fahrt aufnimmt und mit der wir erst noch umzugehen lernen müssen. Zurzeit befinden wir uns in einer Phase, in der technische Ansätze, die bislang von konkurrierenden Schulen gepflegt wurden, zusammenwachsen, in einer Phase, in der die nun verfügbaren großen Datenmengen neue Lernverfahren und mit diesen unerwartet schnelle Fortschritte ermöglichen. Seit Produkte der KI-Forschung marktauglich geworden sind, tritt zudem die Forschung privater Unternehmen massiv neben die der Universitäten und des Militärs und beschleunigt die Entwicklung zusätzlich. Noch passen die am weitesten entwickelten Roboterkörper und die klügsten Programme nicht recht zusammen. Die beweglichsten unter den Robotern haben wenig im Kopf, die klugen Antwortmaschinen hingegen residieren in Superrechnern und haben bestenfalls eine Sprachausgabe. Doch das Zusammengehen von Robotik und Künstlicher Intelligenz gilt als einer der nächsten großen Schritte in der Entwicklung intelligenter Maschinen. Ein weiterer ist das Unsichtbarwerden der KI: Statt digitale Assistenten auf dem Handy herumzutragen oder auf dem Tablet aufzurufen, soll sich die intelligente Technik so unauffällig wie allgegenwärtig in unseren Alltag integrieren. Dann könnten wir den Backofen per Zuruf steuern und

Informationen vom Badezimmerspiegel erhalten. Zugleich rückt die lange als utopisch angesehene künstliche allgemeine Intelligenz – anstelle von bzw. zusätzlich zu Systemen, die auf einen sehr kleinen Einsatzbereich spezialisiert sind – wieder auf die Agenda der Forscher.

So dürfen wir täglich von neuen erstaunlichen, teils faszinierenden, teils absonderlichen Dingen lesen: Wie wäre es mit einer App, die uns sagt, wie wir mit dem Partner umgehen sollten, wenn er oder wir selbst schlecht gelaunt oder müde sind? Oder einer App, die den Chef mit leichter Vibration seines intelligenten Armbands dazu ermahnt, statt den Untergebenen anzumaulen, erst einmal an die Luft zu gehen und durchzuatmen? Oder einer intelligenten Puppe, die uns die Gefühle unserer Kinder entschlüsselt? Emotionen haben es einigen Forschern derzeit angetan. Mir wäre ein Roboter, der im Garten herumkriecht und Unkraut ausreißt, lieber – auch ein solcher ist (natürlich) in der Entwicklung.

Wir befinden uns in einer Phase, in der die Digitalisierung sich in der Gesellschaft bemerkbar zu machen beginnt und die Frage nach ihren Auswirkungen immer dringender wird; einer Phase, in der die Technikfolgenabschätzung sich vorsichtshalber auch mit sehr unwahrscheinlichen Szenarien befasst, die dann oft gerade die meiste Aufmerksamkeit bekommen: Wenn das Schicksal der Menschheit auf dem Spiel steht, rechtfertigt doch auch eine sehr kleine Wahrscheinlichkeit, dass man sich mit einem Szenario befasst, oder etwa nicht?

Wir befinden uns in einer Phase, in der die Politik sich um die Künstliche Intelligenz und ihre Auswirkungen zu kümmern beginnt und feststellen muss, dass bereits monopolartige Strukturen bestehen, gegen die schwer und von einzelnen Nationen allein vermutlich gar nicht anzukommen ist; in einer Phase, in der die Big Player der Digitalisierung massiv um Vertrauen in die neuen Technologien werben, während zugleich mehr und mehr Institutionen und Nichtregierungsorganisationen es in die Hand nehmen, kritisch über Hintergründe und die möglichen Gefahren der KI aufzuklären. Wir sind hin- und hergerissen zwischen den Möglichkeiten einer neuen Technologie und ihren vermuteten und realen Nebenwirkungen.

Manche Forscher geben die Schuld daran dem allzu griffigen und werbewirksamen Begriff «Künstliche Intelligenz». Ohne ihn hätte die Disziplin nie die Aufmerksamkeit erfahren und nie die Befürchtungen und Hoffnungen geweckt, die wir heute diskutieren. «Künstliche Intelligenz» verspreche zu viel und rücke die Technik zu nah an den Menschen heran. Hätte das ganze Unterfangen einen weniger charismatischen Namen bekommen, etwa «anthropic computing», kämen wir heute nicht in die Versuchung, Roboter als eine Art Vormenschen zu beschreiben, meint der Computerwissenschaftler Jerry Kaplan.¹³ Da ist etwas dran, doch Maschinen, die sprechen, die herumgehen, die lächeln oder mit den Augen rollen, sind etwas Besonderes, egal wie man sie bezeichnet. Wir sind nun einmal so eingerichtet, dass wir gar nicht anders können, als sie als Wesen mit Gedanken, Plänen, Absichten und Wünschen zu betrachten. Anthropomorphismus nennen das die Forscher. Und dem erliegen nicht nur Laien: Die einschlägige wissenschaftliche Literatur ist voller Systeme, die denken, lernen, planen, voraussagen, analysieren, abwägen, entscheiden, wissen und handeln. Da gibt es autonome Systeme und sogar selbstbewusste Systeme. Diese Begriffe meinen teilweise dasselbe wie in der Alltagssprache, teilweise aber auch etwas ganz anderes. Auf den folgenden Seiten wird es auch darum gehen, das auseinanderzusortieren.

Nicht umsonst hat der böse Roboter, den es zwecks Rettung der Welt in der Science-Fiction regelmäßig zu vernichten gilt, zumeist die Gestalt einer schönen Frau: da fällt dem Helden das Abdrücken besonders schwer – und uns die Einschätzung, womit wir es eigentlich zu tun haben: Was kann ein Roboter und womit steht er in Verbindung? Ist ChatGPT mehr als ein beeindruckendes, aber letztlich dummes Statistikprogramm? Kann es denken? Versteht es die Texte, die es generiert? Hat ein System, das mit Belohnungsfunktionen arbeitet, nicht doch so etwas wie Gefühle? Und wie sollen wir mit den intelligenten Maschinen umgehen? Sollen Roboter aussehen wie Menschen? Darf man sie schlagen? Könnten sie irgendwann zu Bewusstsein kommen? Würden wir das merken? Können sie schuld sein, wenn sie Fehler machen?

Vielleicht werden wir schon in wenigen Jahren ganz selbstver-

ständig mit intelligenten Maschinen umgehen und solche Fragen als naiv belächeln. Aber um dorthin zu kommen, müssen wir erst einmal versuchen, sie zu beantworten. Wir müssen unsere Intuitionen und unsere Begrifflichkeiten klären. Wir müssen ausprobieren, wie wir auf künstliche Systeme reagieren, welche gut für uns sind und welche nicht. Wir müssen darüber nachdenken, wie wir die klugen Maschinen nutzen können, um die Gesellschaft menschlicher zu machen, nicht nur effizienter. Wir müssen zusehen, dass uns Profit- und Machtstreben die Freude an einer Technologie, von der die Menschheit so lange geträumt hat, nicht gleich wieder verderben.

Denn intelligente Maschinen sind noch in einer anderen Hinsicht etwas Besonderes: Sie waren nie nur dazu da, uns schmutzige, gefährliche oder langweilige Arbeit abzunehmen. Von Beginn der KI-Forschung an dienten sie auch dazu, den Menschen besser zu verstehen. Man habe nur wirklich verstanden, was man auch bauen könne, formulierte der Physiker und Nobelpreisträger Richard Feynman. Das bewegt auch heute viele KI-Forscher. Der Mensch mit seiner vielseitigen Intelligenz stellt das Vorbild. Und viele Programme und Roboter sind Hypothesen darüber, wie der Mensch funktioniert. Sie bringen Forscher dazu, Menschen (und Tiere) noch einmal ganz genau zu betrachten, um Hinweise zu bekommen, wie sie intelligentes Verhalten bewerkstelligen. Und sie zwingen sie, ihre Theorien so präzise zu formulieren, dass man sie in einer Maschine realisieren kann. Wenn wir die Leistungen der Roboter mit unseren eigenen vergleichen, zwingen uns die Maschinen dazu, auch über uns selbst noch einmal neu nachzudenken: Was genau ist eigentlich Intelligenz? Was ist Autonomie? Was ist Kreativität? Was macht den Menschen aus? Wenn wir heute ein viel differenzierteres Bild der menschlichen Intelligenz haben als in den 1950er Jahren, liegt das auch daran, dass sich Computermodelle immer wieder als zu einfach erwiesen haben. Die Verwirrungsmaschinen können hier zu Präzisierungsmaschinen werden, die uns zeigen, worin wir wirklich gut sind.

Dieses Buch hat zwei Teile: Im ersten Teil geht es um die Grundlagen, um Algorithmen und maschinelles Lernen, um Transformer, Roboter und Cyborgs und um das, was dem Menschen aller

Voraussicht nach vorbehalten bleiben wird. Im zweiten Teil geht es um die Folgen der Digitalisierung und diejenigen Fragen, über die die Gesellschaft sich in den nächsten Jahren verständigen muss: um den gläsernen Menschen und die Verletzlichkeit der Infrastruktur, die Veränderungen in der Arbeitswelt, um Roboter als Kriegsmaschinen und als Gefährten des Menschen. Den Schluss bildet ein Blick in die Zukunft: Wie könnte unser Leben mit der Künstlichen Intelligenz aussehen? Gibt es vielleicht doch ein wenig Grund, sich auf unsere intelligenten neuen Werkzeuge zu freuen?

TEIL 1

WAS KÜNSTLICHE INTELLIGENZ KANN

1. WAS IST KÜNSTLICHE INTELLIGENZ?

Künstliche Intelligenz ist in. Es ist ein Etikett, das Maschinen interessant macht und gerne und inflationär auf vielerlei geklebt wird, ob Autos, Suchmaschinen oder Onlineshops. Noch inflationärer wird das Wörtchen «smart» gebraucht: Handys sind schon lange smart, und glaubt man der Werbung, gilt das auch für Kameras und Kühlschränke, Feuermelder, Staubsauger und Fernseher, Zahnbürsten, und Bettlaken. Künstliche Intelligenz, so scheint es, ist überall. Aber was bedeuten die beiden Wörter überhaupt?

Optimistische Anfänge, diverse Winter und ein neuer Boom

Anfang September 1955 reichte der 28-jährige John McCarthy, damals Assistenzprofessor für Mathematik am Dartmouth College in Hanover, New Hampshire, zusammen mit drei Kollegen bei der Rockefeller-Stiftung einen Antrag zur Förderung eines ambitionierten Projekts ein:¹ Zehn ausgewählte Forscher sollten im Sommer 1956 zwei Monate lang herauszufinden versuchen, «wie Maschinen dazu gebracht werden können, Sprache zu benutzen, Abstraktionen und Begriffe zu bilden, Probleme zu lösen, die zu lösen bislang dem Menschen vorbehalten sind, und sich selbst zu verbessern». Dieses Projekt nannte er «Artificial Intelligence», Künstliche Intelligenz.

Die «Dartmouth-Konferenz» gilt heute als Startschuss der KI-Forschung. Tatsächlich steckten die Forscher damals schon mit-tendrin, nur ein griffiger Name hatte dem Projekt bislang gefehlt. Viele, die heute zu den Gründervätern der KI zählen, waren da-mals dabei: Warren McCulloch, der schon 1943 zusammen mit Walter Pitts in Anlehnung an die Architektur des menschlichen Gehirns erste künstliche neuronale Netze entworfen hatte; Allen Newell und Herbert Simon, die ihr Programm «Logical Theorist» präsentierten, das logische Theoreme beweisen konnte. Zwei Mo-nate später stellte der Linguist Noam Chomsky auf einer Tagung am Massachusetts Institute of Technology seine Theorie der Spra-che vor, der zufolge ein unbewusst bleibendes Regelsystem es Sprechern ermöglicht, immer neue Sätze zu bilden. Sollte man diese Regeln nicht auch einer Maschine beibringen können?

Die Computertechnik war in den 1950er Jahren gerade da-bei, von Röhren auf Transistoren umzusteigen, erste Firmen wie die US-amerikanischen Bell Labs und die deutsche Firma Sie-mens begannen, kleiderschrankgroße Rechner in Serie zu bauen. Trotz ihrer aus heutiger Sicht bescheidenen Rechenkapazität – UNIVAC I, der erste in den USA für den Markt produzierte Com-puter, schaffte knapp 2000 Rechenoperationen in der Sekunde – kannte der Optimismus der jungen Forscher kaum Grenzen: Das Problem sei nicht die Leistungsfähigkeit der Maschinen, sondern «unsere Unfähigkeit, Programme zu schreiben, die das Beste aus dem machen, was wir haben», hieß es in McCarthys Forschungs-antrag. Eine Maschine zu programmieren, die aus Versuch und Irrtum lernen könne, sei nicht schwierig. Er selbst sei zuversicht-lich, bis zum Sommer 1956 eine Maschine, die ein Modell ihrer Umwelt entwickeln und Probleme durch Experimente mit diesem Modell lösen könne, so weit zu haben, dass man sie programmie-ren könne.²

Die großen Fortschritte kamen, wenn auch deutlich später als erwartet. Erst in den letzten Jahren zeichnet sich ab, dass und wie sich einige der kühnen Ideen des legendären Dartmouth-Som-mers verwirklichen lassen. Eine Maschine, mit der man sprechen kann, trägt heute jeder in der Tasche herum, lernende Programme können selbstständig Muster in großen Datenmengen erkennen,

schon das Navigationsgerät im Auto löst ein alles andere als triviales Problem – wie komme ich am besten von A nach B? –, und in Ansätzen können Programme sich auch selbst verbessern. John McCarthy, der 2011 starb, hat unter anderem mit der Entwicklung der Programmiersprache Lisp viel dazu beigetragen.

Künstliche Intelligenz: Diese beiden Wörter stehen heute für ein weites, interdisziplinäres und wenig übersichtliches Gebiet, auf dem Forscher mit ganz unterschiedlichen Fragestellungen und Methoden daran arbeiten, Systeme zu entwickeln, die Dinge können, wie McCarthy sie aufzählte: Sprache verwenden, Begriffe bilden, Probleme lösen. Künstliche Intelligenz, so heißt es bisweilen flapsig, das ist, Maschinen zu bauen, die können, was die Maschinen in der Science-Fiction können.

Zu den Teilgebieten der KI gehört die Sprachverarbeitung ebenso wie die Darstellung von Wissen in Datenverarbeitungssystemen, automatisches Schlussfolgern, Wahrnehmung und Bildanalyse, die Robotik und diejenige Disziplin, die aktuell am meisten von sich reden macht, die Geschichte der KI jedoch von Beginn an begleitet hat: das maschinelle Lernen.

Die KI-Forschung baut auf Arbeiten vieler Disziplinen auf, vor allem natürlich auf der Informatik: Künstliche Intelligenz ist die Suche nach Computer- und/oder Roboterintelligenz. Andere Möglichkeiten, intelligente Leistungen künstlich zu realisieren, sind bislang nicht bekannt. Die Psychologie steuert Erkenntnisse zu zwischenmenschlicher Interaktion, Lernen und Wahrnehmen bei. Aus der Logik kommen Einsichten in die Struktur von Aussagen, aus der Philosophie Überlegungen, wie Begriffe Bedeutung bekommen oder wie Geist und Materie zusammenhängen. Hinzu kommen mathematische Arbeiten, die elementare Konzepte wie Berechenbarkeit, Algorithmus und Wahrscheinlichkeit definieren, entscheidungs- und spieltheoretische Überlegungen aus der Ökonomie, Erkenntnisse der Neurowissenschaft über das Funktionieren des Gehirns, Einsichten aus der Sprachwissenschaft, der Pädagogik und der Ethnologie sowie, für den Bau der Roboter, der unerschöpfliche Pool biologischer Vorbilder für Körperbau, Sinnesorgane und Gliedmaßen bei Mensch und Tier.

Zeitgleich mit der KI entstand die ebenfalls interdisziplinär

ausgerichtete und größtenteils auf dieselben Disziplinen zurückgreifende Kognitionswissenschaft. Im Vergleich zur KI kehrt sie die Perspektive um: Statt sich am Menschen zu inspirieren, um Computer besser zu machen, bedient sie sich – unter anderem – der Computer, um besser zu verstehen, wie die menschliche Kognition funktioniert. Beide Gebiete, KI wie Kognitionswissenschaft, führen das jeweils andere als Teilgebiet des eigenen Unternehmens. Die parallele Entstehung beider Disziplinen zeigt: Der Computer weckte von Anfang an nicht nur die Hoffnung auf intelligente Maschinen, sondern auch auf ein neues, besseres Verständnis von Intelligenz überhaupt.

Bis heute prägt ein Auf und Ab zwischen Phasen der Euphorie und Phasen der Ernüchterung die Geschichte der KI. Mitte der 1950er Jahre schien alles möglich und zum Greifen nahe. Doch die Ernüchterung ließ nicht lange auf sich warten: Erste Programme zum automatischen Übersetzen zwischen Englisch und Russisch, die sich die US-Armee im Kalten Krieg gewünscht hatte, scheiterten an der Komplexität der Aufgabe, autonome Panzer ließen sich nicht realisieren, erste lernende Systeme auf der Basis künstlicher neuronaler Netze zeigten gravierende Mängel. Nachdem Kommissionen von staatlicher wie von militärischer Seite zu dem Schluss gekommen waren, die vollmundigen Ankündigungen der Forscher würden sich nicht realisieren lassen, fuhr die DARPA, die Forschungsabteilung des US-Militärs, die die KI-Forschung in ihren Anfängen massiv und ohne große Vorgaben gefördert hatte, ihre Zahlungen mehrmals zurück. Ende der 1970er und Ende der 1980er Jahre kam es so zu Einbrüchen in der KI-Forschung, die heute als «KI-Winter» bezeichnet werden. Auch in diesen frostigen Zeiten wurde freilich weitergeforscht und -entwickelt. Im Schatten des Scheiterns zu groß angekündigter Projekte wurden die Grundlagen des aktuellen Hypes gelegt. Denn zurzeit, man kann es kaum anders sagen, herrscht in der KI ein Sommer mit nie da gewesenen Höchsttemperaturen.

Wer blufft, gewinnt: der Turing-Test

Nicht jedes Gerät mit einem Prozessor oder einem Internetanschluss ist intelligent. Digitalisierung und Datenverarbeitung, die Erfassung von Informationen in digitaler, also computerlesbarer Form und ihre Verarbeitung sind noch keine Künstliche Intelligenz. Aber wann ist eine Maschine intelligent? Die berühmteste Antwort auf diese Frage gab 1950 der britische Mathematiker Alan Turing. Er erklärte sie zu einer Frage des Sprachgebrauchs und für zu belanglos, um sie überhaupt zu diskutieren. Bis zum Ende des 20. Jahrhunderts, so seine Prognose, werde man sich schlicht daran gewöhnt haben, von denkenden Maschinen zu sprechen. Deshalb schlug er vor, die Frage «Kann eine Maschine denken?» durch die Frage zu ersetzen, ob eine Maschine sich in einem Spiel bewähren könne, das er «Imitationsspiel» nannte. Ein Mensch kommuniziert dabei so, dass er seine Gesprächspartner nicht hören oder sehen kann, mit einem Menschen und einer Maschine. Seine Aufgabe besteht darin, durch Fragen herauszufinden, wer von beiden der Mensch, wer die Maschine ist. Dieses Imitationsspiel ist heute als «Turing-Test» bekannt. Halten 30 Prozent der durchschnittlichen Anwender den Computer nach einer Fragezeit von fünf Minuten für einen Menschen, sollte der Test als bestanden gelten.

Die Vorteile dieses Verfahrens liegen auf der Hand: Eine klare Aufgabe tritt an die Stelle einer Definition, die mit dem notorisch unklaren Begriff des Denkens arbeiten müsste oder mit einer ebenso notorisch umstrittenen Liste, welche Fähigkeiten denn nötig seien, damit ein System als intelligent gelten kann. Das Aussehen des Gegenübers spielt keine Rolle. Zudem ist der Test anspruchsvoll: Ein System, das ihn bestehen kann, muss Sprache verarbeiten können und Weltwissen besitzen, es muss speichern können, was es schon gesagt hat, und sich am besten auch auf den Fragenden einstellen können.

Doch ist dieser Test überzeugend? Zuerst einmal funktioniert er nur bei Systemen, mit denen man einen Dialog führen kann. Das Programm, das Lee Sedol im Go-Spiel besiegte, käme nicht infrage, weil man sich mit ihm nicht unterhalten kann. Auch auto-