

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 Zauberlehrling, 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 starrten, 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.