

Einleitung

- » Seit 20 Jahren liest man regelmäßig, dass der Durchbruch der Künstlichen Intelligenz unmittelbar bevorsteht. Aber jetzt stimmt es wirklich. (Sascha Lobo)
- » Wenn man ihnen gestattet, in einer komplexen Welt eigenständig zu handeln, bergen geistlose Maschinen enorme Risiken gemeinsam mit ihrer enormen Leistungsfähigkeit, gleichgültig, ob sie als Roboter verkörpert sind oder einfach algorithmisch abgeleitete Urteile ausgeben. (Nicolas G. Carr)

Mythos und Wahrheit, Hype und Realität lagen kaum jemals so nah beieinander wie bei der künstlichen Intelligenz. Ist KI wirklich schon zu dicht am biologischen Vorbild – unserer eigenen Intelligenz? Oder geht es hier eher um eine prinzipiellere, grundlegendere Frage? Schließlich haben wir Menschen ja bereits Myriaden an ausgeklügelten Werkzeugen entwickelt. Aber noch nie waren wir in der Situation, die Potenziale, Chancen und Risiken eines unserer Werkzeuge nicht eindeutig abschätzen zu können. Wird uns KI befreien oder versklaven? Erlösen oder ersetzen? »Empowern« oder entmachten?

Mit dieser binären Logik wird der digitalen Transformation rund um KI und Robotik eine geradezu paranoide Entscheidungsnot eingeprägt. Doch KI besitzt, wie jede andere Technologie, vom ersten Faustkeil bis zur Nuklearkraft, *dual use*-Charakter: Sie kann nutzen oder schaden. Die überdeutlich aufscheinende Ambivalenz von Technik hat hier zwar eine völlig ungekannte Dimension erreicht, ist aber im Kern alles andere als neu oder unerhört. Bereits der Faust-

keil des frühen Menschen konnte zum Zerkleinern von Nahrung oder zum Zerschmettern der Schädelkapsel eines unliebsamen Zeitgenossen verwendet werden. Und selbst Waffentechnologie kann zu zivilisatorisch wünschenswerten Zwecken wie der Selbstverteidigung genutzt werden. Dieser zweischneidige *dual use*-Charakter von Technik ist seit unseren Pioniertagen als Werkzeugmacher präsent, sollte jedoch der Weiterentwicklung der Tools nicht entgegenstehen. Denn die spezifische Nutzung des Werkzeugs war schon immer entscheidender als das Werkzeug selbst.

Und so paradox es klingen mag: Gerade, weil eine abschließende Beurteilbarkeit so schwer fällt, forciert dies die Entscheidung für eine der beiden möglichen Optionen. Mit anderen Worten: Eben weil es nur zwei Möglichkeiten zu geben scheint, muss eine eindeutige Entscheidung her, die aus den beiden Möglichkeiten eine Wahrheit zieht. Doch folgt künstliche Intelligenz wirklich mit algorithmischer Unausweichlichkeit dieser binär-oppositionellen Logik? Oder liegt darin vielmehr eine Zuschreibung durch uns Menschen, den zunehmend verzweifelten und verzweifelnden Erfinder dieser so potenten wie rätselhaften Technologien?

Wie so oft liegt die Wahrheit auch hier auf einem Spektrum – genau genommen auf einem Wahrscheinlichkeitsspektrum, auf das wir Menschen allerdings entscheidenden Einfluss haben. Denn wie bei jeder anderen Technologie auch kommt es ganz entscheidend auf die Art und Weise des menschlichen Einsatzes an. Lernen wir als kreative Menschen, KI-Systeme zu unserem Vorteil zu nutzen oder verweigern wir uns der gesamten Entwicklung – und werden genau durch diese Haltung abgehängt? Überlassen wir der Technologie das Steuer oder bleiben wir Kapitän? Nutzen wir KI für zivile Zwecke oder lassen wir zu, damit neue Formen der Kriegsführung zu etablieren? Das sind nur drei von zahlreichen möglichen Grundsatzen.

Festzuhalten bleibt: Je fortgeschrittener und weniger verbreitet die Technologie ist, desto undurchsichtiger erscheinen ihre Kernfunktionen. Zudem ist mit dem Hype rund um KI eine Fülle an Experten aus dem Kraut geschossen, die sich teils nicht nur untereinander uneins sind, sondern sich sogar regelmäßig selbst widersprechen. Es ist also gar nicht so leicht, aus der Laiensphäre eine fundierte digitale Perspektive aufzubauen. Aber nichtsdestotrotz ist es möglich! Und dazu soll dieses Buch zugleich ein Ansporn und eine praktische Hilfestellung sein. Fakt ist jedenfalls: KI ist gekommen, um zu bleiben.

Totgesagte leben offenbar wirklich länger. Denn künstliche Intelligenz wurde schon mehrmals abgeschrieben, für ultimativ unerreichbar gehalten oder auf andere Weise in Zweifel gezogen. Zwei kalte »KI-Winter« (der erste Mitte der 1970er bis Anfang der 1980er Jahre, der zweite Ende der 1980er bis Anfang der 1990er Jahre) wurden überstanden, in denen die Erwartungen der Investoren nicht erreicht worden waren, wodurch das Investitionsvolumen naturgemäß zurückgefahren wurde. Doch die Realität verhält sich nicht streng binär wie die Logik der maschinellen Intelligenz. Zwischen Mythos und Wahrheit kann demnach ein weites Spektrum an Zwischenstufen liegen. Dennoch lassen sich einige dieser Vorstellungen eindeutig in das Reich des Mythologischen verweisen, während andere fundiert sind. Ihre schiere Existenzmöglichkeit wurde angezweifelt, ja sogar vehement abgestritten. Aber jetzt ist künstliche Intelligenz – vor allem dank der »ChatGPT Revolution« – endlich im Mainstream angekommen. Ähnlich wie einst beim Internet oder der allgemeinen Digitalisierung etabliert sich KI gerade als neuer Standard, an dem man einfach nicht mehr vorbeikommt. Und das gilt nicht nur für den Einzelnen – welcher Profession auch immer angehört – sondern in einer anderen Schärfe auch für Unternehmen und Konzerne. Wer jetzt noch zögert, spezifische KI-Anwendungsfälle (sogenannte *use cases*) zu identifizieren, versäumt die Chance, seine Prozesse effizienter und sein Unternehmen produktiver zu machen. Rund zwei Jahre nach der Markteinführung des Epoche

machenden Tools »Chat GPT3« durch die Firma »OpenAI« kann sich kein Unternehmen mehr zu den frühen Anwendern (*early adopters*) zählen, wenn jetzt erst damit begonnen wird, KI-basierte Technologien zu nutzen. Erste fundierte Statistiken wie die 2024 veröffentlichte Adesso-Studie zum Einsatz von generativer KI zeigen: Auch in Deutschland stehen wir bereits eher irgendwo zwischen früher und später Mehrheit (*early/late majority*). Insofern hat der eingangs zitierte Autor und Blogger Sascha Lobo mit seiner Einschätzung natürlich einen validen Punkt getroffen. Diesmal sind die Erwartungen an KI-Systeme offenbar mehr als gerechtfertigt. Denn sie besitzen das Potenzial, fast alles zu verändern.

Kernthesen

Die Kernthesen dieses Buches auf einen Blick:

1. Der Mensch wird immer mehr sein als KI. Das menschliche Gehirn kann durch KI weder zur Gänze nachgebildet noch ersetzt werden.
2. KI wird immer nur ein Werkzeug bleiben. Durch die Simulation bzw. die geschickte Nachahmung menschlicher Fähigkeiten wird die Unterscheidung jedoch immer anspruchsvoller. Es besteht zunehmend die Gefahr der Anthropomorphisierung (Vermenschlichung) von KI.

In diesen Kernthesen spiegelt sich weniger Technikskepsis als vielmehr die feste Überzeugung, dass der Mensch immer vielschichtiger und komplexer sein wird als sein vielseitigstes Werkzeug. Wir dürfen nicht vergessen, warum KI überhaupt erfunden wurde – nämlich, um den Menschen in all seinen Tätigkeiten zu unterstützen. Nur, weil unsere Werkzeuge langsam, aber sicher immer besser darin werden, diese vielfältigen Aufgaben zu erfüllen, heißt das nicht, wir

würden dadurch entmachtet. Im Gegenteil: Wer versteht, die mannigfachen Potenziale geschickt für sich zu nutzen, wird ermächtigt!

Mensch > KI

Der Mensch, so lautet die erste Kernthese, wird sich durch künstliche Intelligenz weder vollumfänglich erfassen noch abbilden oder gar ersetzen lassen. Somit wird der Mensch als Schöpfer von KI auch niemals von seinem bislang potentesten Werkzeug evolutionär auf die hintersten Ränge verwiesen werden können. Dieser Umstand hat wenig bis nichts mit der anscheinenden Substratunabhängigkeit kognitiver Prozesse zu tun (also ob sie biologisch oder künstlich sind), sondern erscheint vielmehr als logisches Resultat aus dem Verhältnis von Original zu Simulation. Die durch künstliche Intelligenz herbeigeführten Simulationen menschlicher Geisteskraft können in Teilbereichen dem Original sogar weit überlegen sein (siehe Schach-Computer oder »AlphaGo Zero« etc.). Der menschliche Geist hat all diese Expertensysteme allerdings ersonnen, wodurch sich paradoxerweise sein uneinholbarer Vorsprung zeigt. Zudem hat er dies durch seine intrinsische Motivation und seine Ambitionen geschafft, die Welt, seine Gegenwart und Zukunft zu gestalten.

Untrennbar mit dem »menschlich sein« verbunden ist zudem die Subjektivität, mit der eine große Unvorhersehbarkeit einhergeht.

Der Mensch als ultimativer Werkzeugmacher

KI ist unser Werkzeug, der Mensch sein Schöpfer, so die zweite Kernthese. Binnen rund einer Million Jahren etablierte sich der Homo sapiens als ultimativer Werkzeugmacher, der stets geschickter und einfallsreicher hinsichtlich Konzeption und Herstellung seiner Gerätschaften wurde. Bis zum heutigen Tag hat diese Evolution keineswegs nachgelassen – im Gegenteil. Als einziger Werkzeugmacher des Tierreichs gestaltete der Mensch seine Umwelt und sich selbst

wie kein anderes Wesen; eine Entwicklung, die schließlich vorerst in der Erfindung des wohl potentesten unserer Werkzeuge gipfelte. Der künstlichen Intelligenz. Die ohne ihr biologisches Vorbild allerdings vollkommen undenkbar wäre. Die Entstehung und vor allem die weitere Entwicklung künstlicher Intelligenz kann nicht isoliert vom Menschen gedacht werden. Nicht nur, weil der Mensch als ihr Schöpfer und Urheber für die Entwicklung von KI verantwortlich ist – sondern auch, weil unsere eigenen kognitiven Werkzeuge von ihren künstlichen Mit- und Gegenspielern immer wieder herausgefordert werden. Welchen ethischen Normen soll KI gehorchen? Welche menschlichen Schwächen dürfen sich auf keinen Fall in der Maschinenintelligenz fortsetzen? Wie können wir gewährleisten, die Potenziale für Nutzen zu maximieren und die Schadenspotenziale zu minimieren?

Als dominantester Werkzeugmacher des Tierreichs hat der Mensch mit KI sozusagen sein Meisterstück abgeliefert. Glücklicherweise verfügte der Mensch auch über das passende *embodiment* (allen voran die Nutzung des opponierenden Daumens), um seine kognitiven »PS« auch auf die Straße bringen zu können. Doch KI hat mit der Intelligenz ihrer Schöpfer tatsächlich nur Schnittmengen gemeinsam.

Betrachten wir zur Annäherung an das faszinierende kognitive Phänomen der Intelligenz einmal zwei diametral einander gegenüberstehende Definitionen:

- Definition 1: Intelligenz ist die Fähigkeit, aus einem Minimum an Informationen ein Maximum an relevanten Schlüssen zu ziehen.
- Definition 2: Intelligenz ist die Fähigkeit, aus einem Maximum an Informationen ein Minimum an relevanten Schlüssen zu ziehen (Lierfeld 2018).

Was auf den ersten Blick wie eine formale Spielerei wirken mag, beschreibt auf den zweiten Blick recht exakt die Unterschiede von

biologischer und maschineller Intelligenz. Die Menschheit hat von je her konzeptionelle Durchbrüche auf dünner Informationslage »ge-wagt« und ist damit regelmäßig auf die nächsthöhere Entwicklungsstufe gelangt. Beispielsweise liegt das Minimum an Informationen bei der Entstehung von Feuer in zwei Beobachtungen: Reibung erzeugt Hitze und unterschiedliche Materialien können diese Hitze unterschiedlich gut speichern und abgeben. Diese Informationen enggeführt ergeben den konzeptionellen Durchbruch, der zum Beherrschen des Feuers erforderlich war. Ähnlich verhält es sich mit der Erfindung des Rades und vielen weiteren Beispielen mehr, wobei freilich mit der jeweiligen Entwicklungsstufe auch der Komplexitätsgrad kontinuierlich stieg (Lierfeld 2018).

Betrachten wir die zweite Intelligenzdefinition, so haben wir hiermit eine Schablone für den Umgang mit Big Data – überwältigend großen Datenmengen, auf die Maschinenintelligenz viel besser vorbereitet ist als menschliche Intelligenz. Lediglich Autisten verfügen ebenfalls über die Fähigkeit, aus einem überbordenden Maximum an Informationen ein Minimum an relevanten Schlüssen zu ziehen. Maschinen können dies jedoch auf reproduzierbar zuverlässige Weise und in Datenmengen, die den leistungsfähigsten Savant auf die hintersten Plätze verweisen würde. In einer Utopie würde unsere biologische Intelligenz Wege finden, mit der künstlichen Intelligenz Synthesen einzugehen, ohne die eigene Identität völlig preis- oder aufgeben zu müssen. In einer dystopischen Alternativzukunft würden wir allerdings weder konzeptionelle Durchbrüche noch die Beherrschung der Big-Data-Welt unserer Spezies zuschreiben können. Künstliche Intelligenz würde damit tatsächlich zu unserer vom amerikanischen Autor und Dokumentarfilmer James Barrat ausgerufenen »letzten Erfindung« (Barrat 2014). Im positiven Fall postuliert Barrat KI als die letzte Erfindung, die wir je machen müssen, da die Maschinenintelligenz von nun an jegliche Weiterentwicklung für uns übernehmen würde, aber im negativen Fall spricht er von der

letzten Erfindung, die wir je machen können werden, da wir durch eben jene potenten Maschinen entmachtet werden.

KI hat das Potenzial, mehr zu verändern als die Einführung der Elektrizität und die Erfindung des Internets zusammen. Daher ist diese Forschungsrichtung vollkommen zurecht gerade in aller Munde. Kein Trendthema ist dabei zugleich mit einem derart gewaltigen Hype, aber auch mit einer solchen Fülle an Vorurteilen, Halbwahrheiten, Falschvorstellungen und gar Mythologisierungen verbunden wie KI. Relativ harmlos, aber zugleich auf fatale Weise irreführend sind Kategorienfehler wie die Gleichsetzung von generativen KI-Tools wie »Chat GPT« mit künstlicher Intelligenz als solcher. Hierbei handelt es sich um eine Art von pars-pro-toto-Fehler, in dem ein Teil mit der Gesamtheit gleichgesetzt wird. Ein Autoreifen gehört zweifellos zu einem Auto, konstituiert dieses aber nicht. Genauso verhält es sich mit generativer KI. Text- und Bildgeneratoren werden zwar durch künstliche Intelligenz ermöglicht, repräsentieren aber nur einen verhältnismäßig kleinen Teil des riesigen Spektrums, das durch KI abgedeckt wird.

Der menschliche Verstand ist das Territorium, KI die Karte

Von jeher entsprach das Verhältnis zwischen menschlichem Verstand und künstlicher Intelligenz dem zwischen Territorium und Karte; ersteres entspricht dabei dem Original, zweiteres der Nachbildung. Diese klare Grenzziehung gilt nach wie vor. Die größten Herausforderungen ergeben sich aktuell aus einem Hang zur Anthropomorphisierung von Technologie sowie den vermeintlich unvermeidlichen Debatten um künstliche Formen von Bewusstsein.

Auch die Idee, Kognition, Intelligenz und gar Bewusstsein könnten substratunabhängig sein, hat im Kern gar nichts mit dieser Frage zu tun. Sprich, nur weil aktuell nichts dagegen zu sprechen scheint, Denkprozesse nicht nur im Protein unseres biologischen Gehirns, sondern auch im Silizium eines Computers realisieren zu können, bedeutet dies keineswegs, Bewusstsein könnte in Chatbots oder anderen KI-Systemen entstehen. Es scheint sich hierbei vielmehr um ein Problem auf der Nutzer-Seite zu handeln. Zunehmend überwältigt von der Performance intelligenter Systeme, erscheint der menschliche User zugleich oftmals unfähig, zu erkennen, dass die Simulation menschlicher Fähigkeiten ja gerade ein Hauptziel bei der Entwicklung jener Systeme war – und schreibt diesen Systemen dann in einem sentimental Reflex fälschlicherweise pseudo-menschliche Qualitäten zu.

Dabei kommt die Revolution künstlicher Intelligenz tatsächlich Dekaden (genau genommen sogar ca. ein halbes Jahrhundert) später als ursprünglich erwartet. John von Neumann, einer der KI-Urväter, mutmaßte bereits in den 1960er Jahren, starke künstliche Intelligenz sei praktisch »um die Ecke«, maximal zehn Jahre entfernt. Dieses Level wurde aber selbst 2024 noch nicht erreicht. Doch nicht nur das; es erscheint nach wie vor als durchaus fraglich, ob *Artificial General Intelligence* (AGI), also KI, die Parität zu allen kognitiven Fähigkeiten der Menschheit erlangt hat, überhaupt im Bereich des Möglichen liegt. Bei intensiverer Betrachtung erscheint der menschliche Geist als zu komplex, vielschichtig und letztlich unvorhersehbar, um ihn auf allen Ebenen künstlich replizieren zu können.

Der Mensch, seit jeher aufgespannt zwischen existenziellen Bedürfnissen wie dem schieren Überleben und der sinnhaften Suche nach Bedeutung, hat seine höchst diversen Motivationen aus evolutionären Rahmenbedingungen und dem damit verbundenen Anpassungsdruck heraus entwickeln können. Erst als die frühen Menschen ihre Nahrungsversorgung hinreichend sicherstellen konnten, blieb

beispielsweise genug Zeit für die Entwicklung rudimentärer Kultur-techniken. Mit anderen Worten: Auch ein besonders tiefsinniger Homo sapiens würde zuerst seinen Verpflichtungen nachkommen und zur Mammutjagd aufgebrochen sein, bevor er in der Höhle den ersten Malereiversuchen frönen konnte. Erst die Pflicht, dann die Kür. Unterm Strich bleibt festzuhalten: Der Mensch ist seit Anbeginn von der Gewissheit seiner Endlichkeit herausgefordert, etwas für die Nachwelt zu hinterlassen. Besonders kreative Menschen stehen zudem unter dem inneren Druck, etwas zu erschaffen. Beides fehlt Maschinen völlig.

Solche existenziellen Probleme stellen sich KI-Systemen nämlich überhaupt nicht. Um mit den intellektuellen Potenzialen der Menschheit gleichzuziehen, müsste KI aber unter anderem in der Lage sein, neue Weltreligionen aufzubauen oder Charles-Manson-artige Kulte zu begründen. Sie müsste aus eigenem Antrieb neue Formen von klassischer und moderner Musik entwickeln, neue literarische und filmische Genres erfinden etc. KI müsste aus einer (ir-)rationalen Überzeugung heraus Kriege anzetteln oder Frieden stiften, paradoxe oder surreale Kunstwerke herstellen oder andere Systeme aus betrügerischer Absicht hereinlegen. All dies erscheint zum Glück derzeit nicht nur als nicht absehbar, sondern nachgerade als höchst unwahrscheinlich. Wie das bei Werkzeugen nun mal so üblich ist, werden KI-Systeme bislang nämlich nur dann aktiv, wenn ihnen der Mensch dazu einen Auftrag erteilt. Was im Zweifel schon schlimm genug sein kann.

Im Jahr 2022 kreierte das Team des US-amerikanischen Pharmaziekonzerns »Collaborations Pharmaceuticals« mit Hilfe der KI-basierten Software »MegaSyn« 40.000 neue Giftstoffe – binnen sechs Stunden! Anstatt die gesammelten toxikologischen Daten zu nutzen, um gefährliche Nebenwirkungen von Medikamenten zu vermeiden, beschritt man hier einfach den umgekehrten Weg. Dazu war nicht viel mehr notwendig, als sprichwörtlich den Schalter um-