

Zwischen Phantasie und Wirklichkeit

—
Mythos Künstliche Intelligenz

CHRISTOPH MEINEL

Geboren 1954 in Meißen, Institutsdirektor und Geschäftsführer des Hasso-Plattner-Instituts in Potsdam.

Kaum ein Thema beflügelt die Phantasie der Menschheit so wie die künstliche Nachbildung menschlicher Intelligenz. In unterschiedlichen Formen lässt sich der Wunsch der Imitation menschlichen Denkens bis in die Zeit der Renaissance zurückverfolgen. So beschreibt der Arzt und Alchemist Paracelsus in seinem 1578 erschienenen Werk *De natura rerum*, wie ein „Homunculus“ („kleiner Mensch“) mit allen kognitiven Fähigkeiten eines Menschen synthetisch hergestellt werden könnte. Es ist kein Wunder, dass literarische Belege zur Synthetisierung menschlicher Intelligenz mit der Entwicklung des mechanistischen Welt- und Menschenbilds in der Neuzeit zunehmen. Die Naturforscher begannen, die Welt zu vermessen und in mechanischen Bildern zu beschreiben. So ist die Welt für Johannes Kepler ein präzises Uhrwerk,¹ und René Descartes stellt sich in den *Meditationen zur Ersten Philosophie* den Menschen als „eine Art

Maschine [vor], die aus Knochen, Nerven, Muskeln, Adern, Blut und Haut [...] zusammengesetzt ist“.² Besonders Ärzte wie Julien Offray de La Mettrie prägen das Bild des ausschließlich mechanisch verstandenen menschlichen Körpers mit Schriften wie *L'homme machine* und inspirieren künftige Generationen von Wissenschaftlern und Künstlern.

Seither fand das Sujet der Imitation menschlicher vegetativer und kognitiver Fähigkeiten großen Anklang in der Kunst und Literatur. In Mary Shelleys *Frankenstein* und Goethes *Faust* werden Menschen von Forschern künstlich erschaffen. In der gruseligen Erzählung von E. T. A. Hoffmann *Der Sandmann* konstruiert Spalanzani den mechanischen Automaten Olimpia, den er als seine Tochter ausgibt. In der Cineastik erfreut sich das Thema künstliche Intelligenz seit Anbeginn großer Beliebtheit. In Fritz Langs Klassiker *Metropolis* wird eine Darstellerin als „Maschinen-Maria“ nachgebildet. Seither fehlen intelligente Maschinen in praktisch keinem Science-Fiction-Film, ob intergalaktisch (*Star Wars*, *Star Trek*, *2001 Odyssee im Weltraum*) oder terrestrisch (*Terminator*, *Ghost in the Shell*, *Matrix*, *I, Robot*). Oft setzt sich dabei die künstlerische Darstellung sehr kritisch mit diesem Thema auseinander, begründet sicherlich in seinem blasphemischen Charakter und einer apokalyptischen Sorge um den Fortbestand der Menschheit.

Die meisten dieser Vorstellungen haben mit der Wirklichkeit wenig zu tun. Wir wissen, dass sich mit der grobschlächtigen Nachbildung von Venen und Adern durch mechanische oder biologische Systeme keine menschenähnliche Intelligenz nachbilden lässt – weder zum Nutzen noch zum Schaden der Menschheit. Bis vor Kurzem hat es außerdem keine technischen Voraussetzungen und Konzepte gegeben, um relevante Anwendungen der künstlichen Intelligenz möglich zu machen. Selbst wenn die Ergebnisse weit hinter der Phantasie früher Naturforscher und Künstler zurückliegen, gab es gerade in den letzten Jahren signifikante Fortschritte in der Forschung zu künstlicher Intelligenz.

Die wichtigsten Anwendungen zeigen sich in der automatischen Sprach- und Bilderkennung sowie deren korrekter Einordnung in die jeweiligen sachlichen Kontexte. Allein diese Anwendungen werden in der Lage sein, die Zukunft der Arbeit, der Mobilität, der medizinischen Versorgung, der Produktion und der Freizeitgestaltung zu revolutionieren.

NEURONALE NETZE

Das grundlegende Problem früher Naturforscher bei der Entwicklung menschenähnlicher Intelligenz lag in einer zu mechanistischen Vorstellung vom menschlichen Denken. In diesem Kontext konnte niemand erklären, wie der „Geist“ in die Maschine kommt. Das hat sich allerdings mit dem Aufkommen der Informatik in der Mitte des 20. Jahrhunderts radikal verändert. So

interessierte sich der deutsche Informatikpionier Karl Steinbuch für selbststeuernde Systeme. In den 1960er-Jahren veröffentlichte er ein Konzept der Lernmatrizen zur Modellierung künstlicher neuronaler Netze.³ Eine Lernmatrix repräsentiert ähnlich den menschlichen Neuronen und Synapsen die Verknüpfung aus „Eigenschafts- und Bedeutungsleitungen“. Dem Kybernetiker schwebte dabei ein künstliches neuronales Netz vor, in dem mehrere Lernmatrizen nacheinander geschaltet werden.

Das der Vorstellung vom menschlichen Gehirn nachempfundene Konzept der neuronalen Netze bestand darin, ein komplexes System von vernetzten, sehr einfachen Rechenelementen über Eingaben (Sinnesreize) verarbeiten zu lassen. In jedem Rechenelement werden Teile des Inputs mit den Berechnungsergebnissen der unmittelbar benachbarten Rechenelemente kalkuliert und das Ergebnis wird im nächsten Schritt seinerseits den benachbarten Zellen als Input zugeleitet. Ein neuronales Netz bietet so ein dynamisches System, das man auf bestimmte Muster „trainieren“ kann. Annotierte Informationen werden als Bild(punkt) und Bedeutung durch die Lernmatrizen verarbeitet, nachdem in einer Lernphase entsprechende „Durchfluss-Kanäle“ generiert worden sind. Nach Abschluss der Trainingsphase – so die Erwartung – würde es möglich sein, dass die künstliche Intelligenz automatisch Bildinformationen (Buchstaben, Farben, Wörter) erkennt.

ALGORITHMEN-BASIERTE KI-VERFAHREN

In der Frühzeit der Informatik gab es noch keine ausreichend rechenstarken Computer, sodass erste Umsetzungen des Konzepts neuronaler Netze keine auch nur annähernd brauchbaren Ergebnisse liefern konnten. Daher verlegte man sich in den Forschungen zur künstlichen Intelligenz (KI) auf algorithmische Lernmethodiken. Hier werden explizite Ausführungsregeln programmiert und katalogisierte Beispiele hinterlegt (zum Beispiel Metadaten). Jeder Berechnungsschritt ist hier ex post nachvollziehbar. Mit solchen algorithmenbasierten KI-Techniken mit ihren regulativen und katalogisierten „Whitelisting“- und Referenzierungsverfahren ließ sich allerdings kein „neues Wissen“ generieren, sondern lediglich Bekanntes schnell erfassen und zuordnen. Beispiele für diese Art der künstlichen Intelligenz sind das Zählen von Pixeln und Wörtern, die Erstellung komplizierter mathematischer Modelle zur Klima- und Wirtschaftsforschung sowie die noch sehr holprigen Google-Übersetzungen vor der Umstellung auf *Deep-Learning Neural Networks* im vergangenen Jahr.

Die Grenzen solcher Nachbildung menschlicher Intelligenz sind offensichtlich. Die Maschine weiß nie mehr als ein Mensch, sie ist nur schneller in der Ausführung, ohne bei neuen Ereignissen adaptionsfähig zu sein. Außerdem können nicht sämtliche Regeln und Beispiele, die für die Erkennung

eines Phänomens notwendig sind, „vorprogrammiert“ werden. Daher bleibt die „Intelligenz“ immer grob und mangelhaft differenziert. Algorithmenbasierte Verfahren können lediglich relativ einfache funktionale Aufgaben lösen beziehungsweise automatisieren. Das Erkennen von Sprache, Bildern, Aktionen lässt sich damit nur schwer realisieren. Warum es für solche Verfahren schwierig ist, Schrift und Bilder zu erkennen, wird deutlich, wenn man sich vor Augen führt, dass eine Handschrift für einen Computer nichts weiter als ein Bild mit unterschiedlich farbigen Pixeln ist. Wie kann ein Programm Text von Bildern unterscheiden? Ein Programmierer müsste alle denkbaren Handschriftpixel programmieren, damit eine Maschine lernen könnte, den Text vom Bild zu unterscheiden und ihn dann auch noch lesen zu können. Analog dazu stellt sich die Schwierigkeit in der Spracherkennung mit undeutlicher, dialektgefärbter Sprache dar.

NOTWENDIGER PARADIGMENWECHSEL

Die Nachbildung menschlicher Intelligenz auf der Grundlage vorgegebener Algorithmen war nicht die Vision früher Naturforscher, sondern sie hatten einen autonom agierenden menschähnlichen Automaten vor Augen, der auch mit unbekannten Situationen umgehen, neue Erkenntnisse generieren und sich selbstständig weiterentwickeln kann. Um diesem Ziel näherzukommen, brauchte es einen Paradigmenwechsel vom rein naturwissenschaftlich-mechanistischen Welt- und Menschenbild zu einem korrelativ-dynamischen. Die erkenntnistheoretischen Voraussetzungen für diese neue Weltsicht wurden in der Philosophie durch die Phänomenologie seit Edmund Husserl und in den Naturwissenschaften um und in der Nachfolge von Albert Einstein geschaffen. Darauf aufbauend entwickelten Neurowissenschaftler eine Vorstellung vom menschlichen Gehirn als komplexer Sammlung von Neuronen und Synapsen, die mittels komplexer elektronischer Signale die Wahrnehmung und Kontextualisierung in Form von Netzen steuern und speichern. Das Gehirn ist demnach ein riesiges neuronales Netz, das Informationen in mehrstufigen konsekutiven Verfahren verarbeitet. Für den Beobachter bleiben das Verfahren und das Ergebnis zufällig. Wieso sich eine bestimmte Intuition einstellt, können Hirnforscher nicht sagen.

Als Ende des letzten Jahrhunderts die algorithmen-basierten Verfahren technisch ausgereizt wurden und sich keine signifikanten Verbesserungen in dem stark geförderten Bereich der künstlichen Intelligenz abzeichneten, glaubte man, dass Fortschritte auf dem Feld der Imitation menschlicher kognitiver Fähigkeiten nicht möglich seien. Erst vor wenigen Jahren, als die Digitalisierung in weite Bereiche der modernen Lebenswelt vordrang und die Rechenleistung moderner Computer bis dahin ungeahnte Kapazitäten erreichte, erhielt die KI-Forschung neuen Aufschwung.

Anstelle der mechanistisch-kausalen Verfahren, wie sie von algorithmen-basierter KI genutzt werden, greifen moderne KI-Entwickler auf die Idee der neuronalen Netze zurück und entwickeln mithilfe von Big Data, leistungsstarken Rechnern und komplexen neuronalen Netzen selbstlernende Systeme, die sich darüber hinaus selbstständig weiterentwickeln, also lernen können. In der Fachwelt wird dieser Ansatz als „Deep Learning“ bezeichnet. Ein trainiertes *Deep-Learning Neural Network* beinhaltet eine Input-Schicht, in der die Rohdaten aufgenommen werden, eine Output-Schicht, die das Ergebnis der Bild-, Text- oder Spracherkennung ausgibt, und sogenannte *Hidden Layers*, also versteckte Schichten. Letztere sind für die eigentlichen Analysen der Eingaben verantwortlich.

Ein „tiefes“ neuronales Netzwerk zeichnet sich dadurch aus, dass es ähnlich dem menschlichen Gehirn über mehrere (zurzeit schon über 100) hintereinander geschaltete und aufeinander aufbauende Schichten verfügt, in denen Informationen anhand verschiedener Parameter verarbeitet werden. In jeder Schicht können mehrere Tausend Parameter (als mathematische Funktion) eingestellt werden, die unterschiedliche Abstraktionsebenen eines Bildes oder einer sprachlichen Äußerung erfassen. Ein solches tiefes neuronales Netzwerk muss vor seinem Einsatz je nach Einsatzzweck mit riesigen Mengen von Bild-, Text- oder Sprachdaten „trainiert“ werden. Auch hier erfolgt das Training mithilfe annotierter Datensätze. Dieser Prozess, der noch zu Zeiten der Kybernetik-Pioniere aufgrund mangelnder Datenlage und langsamer Rechner scheiterte, ist heute gut machbar.

Deep-Learning-Netze können mit Milliarden annotierten großen Datenmengen gespeist werden. Das System trainiert sich nach dem Trial-and-Error-Prinzip selbst, indem die riesige Zahl von Schichtenparametern automatisch anhand der annotierten Daten optimiert wird. Am Ende des Trainingsprozesses ist es dem System möglich, nicht-annotierte Daten (Bilder, Texte, Sprache) mit sehr hoher Wahrscheinlichkeit richtig zu erkennen. Ähnlich wie bei Abläufen im menschlichen Gehirn ist es in tiefen neuronalen Netzen von außen nicht möglich, nachzuvollziehen, wie das System zu dem Ergebnis kommt, dass auf einem Bild eine Katze abgebildet und in einem Video ein tanzendes Paar gezeigt wird. Elektrische Signale durchfluten die *Hidden Layers* von Schicht zu Schicht, gesteuert von den im Training erlernten Parametern. That's it.

REVOLUTIONEN DER KI-ENTWICKLUNG

Künstliche Intelligenz wird in den kommenden Jahren unsere Lebenswelt nachhaltig verändern. In wenigen Jahren werden neuronale Netzwerke in der Lage sein, Bilder, Gesichter, Texte in Bildern, Handlungen, Videos und so weiter sicher und schnell erkennen zu können sowie deren Bedeutung zu

erfassen. Das heißt, dass in zahlreichen Bereichen der Produktions- und Arbeitswelt umfangreich automatisiert werden kann. Komplexe selbstlernende KI-Technologien werden die Voraussetzung dafür schaffen, schon in absehbarer Zeit das vielbeschworene autonome Fahren möglich zu machen.

Auch die Nutzung von Medien wird sich in Zukunft vermutlich radikal ändern. In jeder Minute werden etwa 400 Stunden Videomaterial allein auf YouTube hochgeladen. Um die riesigen Informationsmengen zu verarbeiten, wird man nicht auf automatisierte Verfahren verzichten können. Kein Mensch ist in der Lage, solch riesige Datenmengen überhaupt nur zu erfassen, geschweige denn inhaltlich zu klassifizieren und zu bewerten. Mithilfe von KI-Technologien werden wir in die Lage versetzt, das Wissen, das in den Milliarden von Stunden von Videomaterial zu finden ist, sinnvoll zu nutzen.

In der Produktion und Wartung werden ebenfalls verstärkt KI-Agenten eingesetzt werden. Je stärker sich Maschinen, Fabriken und ganze Produktions- und Lieferketten vernetzen, desto wichtiger wird es sein, Muster zu erkennen und Abläufe automatisch anzupassen. Wenn aufgrund von Verkehrsbehinderungen eine Lieferung zu spät kommt, soll die Fabrik nicht stillstehen. Außerdem können *Deep-Learning*-Systeme genutzt werden, um Wartungsarbeiten zu erleichtern und präventiv einzugreifen, wenn Materialversagen droht. Mit *Predictive Maintenance* („vorausschauende Wartung“) muss man nicht mehr darauf warten, bis ein Schaltkreis durchbrennt, bevor er ersetzt werden kann.

Der Medizinsektor wird von KI-Technologien vollständig revolutioniert werden. Nach derzeitigem Stand ist unser Gesundheitssystem unangemessen teuer und in den Verfahren und Diagnosen viel zu wenig personalisiert. Mithilfe von intelligenten Systemen werden individuelle Therapien möglich und Medikamente zielgenauer einsetzbar. Dadurch kann, um nur ein Beispiel zu nennen, die exzessive Nutzung von Antibiotika in Deutschland eingeschränkt werden, die Medikamentendosis auf den individuellen Stoffwechsel angepasst und vorausschauend diagnostiziert werden. Schon heute ist es neuronalen Netzwerken möglich, die Hautkrebsvorsorge zu verbessern. Anhand von Hautpigmentbildern können Bildanalyseverfahren für den Menschen unsichtbare Indikatoren für voraussichtlich bösartige Tumore erkennen. Ausreichend trainierte Systeme werden es ermöglichen, die Belastung durch bildgebende Diagnoseverfahren (Bestrahlung) zu reduzieren, indem aus MRT-Aufnahmen automatisch präzise Röntgenaufnahmen generiert werden. Darüber hinaus werden KI-Verfahren auf Röntgen- und MRT-Aufnahmen selbstständig Tumore erkennen und klassifizieren können. Auch Herz-Kreislauf-Krankheiten werden mithilfe von Wearables, also tragbaren Datenverarbeitungsgeräten, die beständig die Herzfrequenz messen, zuverlässiger behandelt werden können. Das Vorhofflimmern, das als ein Indikator für einen nahenden Herzinfarkt oder Schlaganfall gilt, kann so sicher erfasst und rechtzeitig behandelt werden.

Diese Beispiele werfen nur wenige Schlaglichter darauf, was mit moderner KI-Technologie möglich sein und in welchem Verhältnis der Mensch zur KI stehen wird. Anders als in den utopischen und dystopischen Phantasien von Neuzeitforschern und Künstlern wird KI den Menschen wohl nicht abschaffen (weder im Positiven noch im Negativen), sondern eine unterstützende Funktion einnehmen. Selbstverständlich wird sich die Lebenswelt des Menschen durch die Nutzung der Technologie verändern, aber das ist angesichts der Entwicklung der Weltbevölkerung, des demografischen Wandels, der Herausforderungen steigender Mobilität, Urbanität und Energienutzung auch dringend erforderlich. Die Furcht aber vor der Abschaffung des Menschen ist unbegründet, da selbst mit den modernen Methoden, zumindest vom heutigen Wissensstand aus beurteilt, kein umfassendes Maschinenbewusstsein entwickelt werden kann.

WORAUF SICH UNSERE GESELLSCHAFT EINSTELLEN MUSS

Trotzdem ist klar, dass immer mehr Arbeiten und zunehmend auch intellektuelle Tätigkeiten von Maschinen übernommen werden können, wie das Beispiel der Radiologie zeigt. Daneben werden Maschinen auch immer mehr lästige, langweilige, unangenehme und gefährliche Arbeiten übernehmen können. Das hat selbstverständlich die Entstehung neuer und bisher unbekannter Arbeitsfelder zur Folge. Menschen müssen außerdem frühzeitig Kompetenzen entwickeln, die sie befähigen, sich in der schnell wandelnden digitalen Welt neue Beschäftigungsfelder zu erschließen. Das größte Problem besteht darin, dass unsere Gesellschaft regulatorisch und strukturell mangelhaft auf die Lebenswelt von morgen angepasst ist. Die digitale Transformation wird Deutschland zu einem Verlierer machen, wenn wir nicht frühzeitig die Infrastruktur dafür schaffen, dass jeder von überall Zugang zur digitalen Welt erlangen, also überall und jederzeit auf einen Breitband-Internet-Zugang zurückgreifen kann. Auch die Schule darf kein digitalfreier Raum bleiben. Wenn dort nicht bald digitale Lernangebote und -techniken Einzug halten, wird die soziale Spaltung der Gesellschaft zunehmen. Ein aus dem letzten Jahrhundert stammender, unangemessen prohibitiver Datenschutz darf nicht dazu führen, dass Deutschland den Anschluss an wichtige datengetriebene digitale Innovationen verpasst und im globalen Wettbewerb an Bedeutung verliert. Nur wenn sich unsere Gesellschaft neugierig und verantwortungsbewusst den unterstützenden Möglichkeiten der KI-Technologien öffnet, werden wir die Herausforderungen der digitalen Transformation meistern.

¹ Johannes Kepler: Brief an Herwart von Hohenburg, 10.02.1605.

² René Descartes: *Meditationes de prima philosophia*, 1642.

³ Karl Steinbuch: *Automat und Mensch*, 1961.