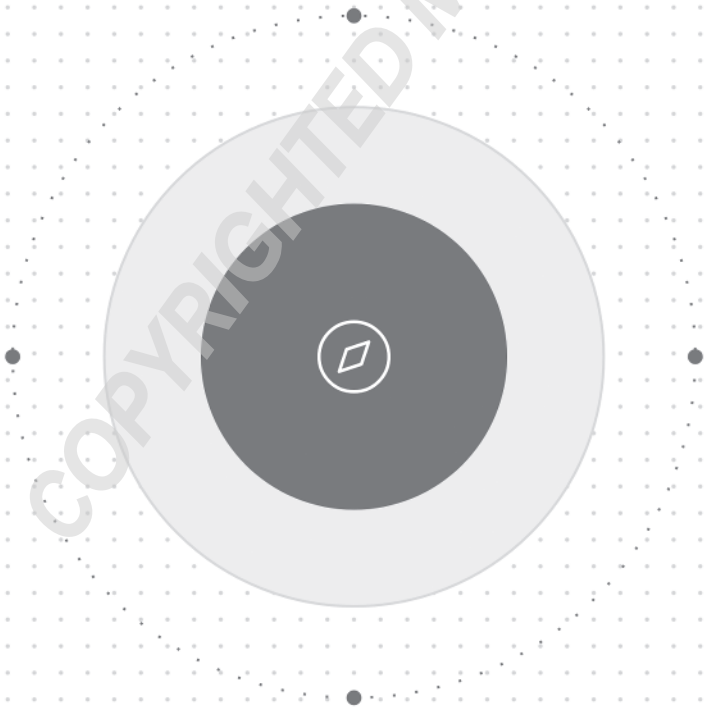
KAPITEL 1

Orientierung im KI-Dschungel

Den Überblick gewinnen – KI einordnen, statt sich im Dschungel zu verlieren.

COPYRIGHTED MATERIAL



In den letzten vier Jahren bis zum Erscheinen dieses Buches hat das Thema Künstliche Intelligenz eine solch starke Präsenz in unserem privaten und beruflichen Leben erfahren, dass kein Tag vergeht, ohne dass wir in irgendeiner Form von ihr sprechen. Dazu nimmt KI in immer mehr Anwendungen unseres täglichen Lebens Einzug. WhatsApp, unsere Bank und das Restaurant um die Ecke wollen nun alle, dass wir mit einer KI sprechen.

Trotz des allgegenwärtigen Erscheinens von KI-Lösungen fehlt uns oft ein Verständnis, was unter der Oberfläche passiert. Führungskräfte stellt dies vor die Herausforderung, in einem meistens fachfremden Thema Wissen und Fähigkeiten aufzubauen. Gleichzeitig müssen sie mit der Geschwindigkeit der Entwicklung mithalten. Eine solide Wissensgrundlage ist notwendig, um Entscheidungen treffen zu können. Sollen im ganzen Unternehmen Copilot-Lizenzen gekauft werden? Sollte man mit einem deutschen Start-up kooperieren? Können wir unsere Daten dort unbesorgt ablegen? Wie bilden wir unsere Teams weiter? In diesem Kapitel stellen wir daher sowohl eine Übersicht über die wichtigsten technologischen Entwicklungen dar, eine Einordnung der aktuell kursierenden Begriffe, als auch einen Referenzrahmen, wie neue Entwicklungen schnell verstanden werden können. Technologie ist kein Selbstzweck. Sondern sie soll immer einen Wert für uns Menschen herbeiführen. Es geht also darum, genug Verständnis aufzubauen, um bewerten zu können, ob die Technologie einen Wert für das eigene Unternehmen beitragen kann oder nicht. Wir werden auch einsteigen in die Limitationen der KI. Dabei betrachten wir, wieso künstliche Intelligenz manchmal wie Magie wirkt und im nächsten Moment an sehr banalen Aufgabenstellungen scheitert. Mit dem Kontext aus diesem Kapitel wird es einfacher zu bewerten, welche Anwendungsfälle ein Kostensumpf sind und wo das wirkliche Potenzial steckt.

1.1 Warum 2022 nicht der Anfang war

Dann begeben wir uns also einmal auf eine Zeitreise und auf eine Reise hinein in die Technologie. Wir werden dies jedoch immer mit dem Blick tun: Was sind die Informationen, die wirklich helfen, Entscheidungen hier und jetzt zu treffen, und nicht nur bei »Wer wird Millionär?« zu gewinnen.

Stellen wir uns vor, wir haben gerade eine Zeitmaschine bestiegen. Wir werden nicht zu lange unterwegs sein, aber ein kurzer Blick in die Vergangenheit lohnt sich. Er erklärt, warum sich der aktuelle Moment so anders anfühlt als alle KI-Wellen davor, und er hilft uns einzuordnen, ob das, was wir gerade erleben, wirklich neu ist oder ob wir schon einmal hier waren.

Der Begriff Künstliche Intelligenz ist 1956 auf der Dartmouth-Konferenz das erste Mal offiziell in die Welt gesetzt worden. Eine Gruppe von Wissenschaftlern beschäftigte sich mit der Suche, wie Menschliche Intelligenz auf Maschinen übertragen werden könne. Der Begriff »Artificial Intelligence« wurde für diesen Forschungszweig gewählt.[1] Der Begriff KI wird auch oft mit Alan Turing in Verbindung gebracht. 1950 veröffentlichte er im Fachjournal *Mind* ein Gedankenexperiment, das Wissenschaftsgeschichte schreiben sollte. Er fragte nicht direkt, ob Maschinen denken können, sondern schlug eine konkrete Prüfung vor.[2] Kann eine Maschine in einem schriftlichen Gespräch so überzeugend auftreten, dass ein Mensch sie nicht von einem Menschen unterscheiden kann? Dieser sogenannte Turing-Test hat die Debatte über maschinelle Intelligenz bis heute nicht losgelassen.

Was die meisten Menschen nicht wissen, ist, dass es in den Jahrzehnten danach nicht eine, sondern gleich zwei große Wellen der KI-Begeisterung gab, und beide endeten mit dem, was Forscher nüchtern als »KI-Winter« bezeichnen. In den 1960er und 1970er

Jahren glaubten viele Wissenschaftler, echte maschinelle Intelligenz sei nur noch eine Frage von zwanzig Jahren. Jedoch stellte sich heraus, die Rechenleistung, die es dafür gebraucht hätte, existierte schlicht nicht. In den 1980er Jahren erlebte KI dann ihr nächstes Comeback, diesmal in Form von sogenannten Expert Systems, also regelbasierten Programmen, die das Wissen von Fachexperten nachahmen sollten. Auch diese Welle verlor sich im Sand, als die Systeme zu unflexibel für die Komplexität der realen Welt wurden.

Diese Geschichte ist deshalb so wichtig, weil sie erklärt, warum viele erfahrene Führungskräfte beim Thema KI zu Recht eine gesunde Skepsis mitbringen. Wer schon mehrfach erlebt hat, dass eine Technologie mehr versprochen hat als sie geliefert hat, der fragt sich zu Recht, ob es diesmal wirklich anders ist. Aber ja, diese KI-Welle ist anders. Um zu verstehen, warum, müssen wir ein paar Jahrzehnte vorspulen.

1997 verlor Schachweltmeister Garri Kasparov gegen Deep Blue, den Schachcomputer von IBM. Das war ein prägender Moment. Der beste Schachspieler der Welt verliert gegen eine Maschine, live übertragen, weltweit beobachtet. Das war ein globales Medienereignis. Es war ein symbolischer Bruch, der die Öffentlichkeit faszinierte und gleichzeitig verwirrte. Doch wer genauer hinsah, erkannte die Grenzen des Triumphs. Deep Blue besiegte den Schachweltmeister nicht, weil die Maschine »denken« oder menschliche Fähigkeiten erlernen konnte, sondern weil Schach ein klar definiertes, mathematisches System ist, das sich mit ausreichend Rechenleistung perfektionieren lässt, ganz ohne Verständnis von Sprache, Bedeutung oder Weltwissen. Es berechnete Millionen von Zugmöglichkeiten pro Sekunde und wählte den statistisch besten. Es war Rechenleistung in reinster Form. Viele Experten sagten damals, das sei kein echtes Denken, nur das Durchrechnen aller möglichen Züge. Das stimmte. Aber es zeigte, was passiert, wenn Rechenleistung auf die richtige Architektur trifft.

Der wirklich entscheidende technische Durchbruch kam etwas leiser. 2006 entwickelten der kanadische Forscher Geoffrey Hinton und sein Team eine Methode, Röntgen- und Scandaten, die das stabile Training tiefer neuronaler Netze möglich machte und zeigte, dass Modellgröße und Datenmenge direkt zu besserer Leistung führen. Maschinen hörten auf, nach fixen Regeln zu arbeiten, und begannen stattdessen, aus Daten selbst zu lernen.[3] Also maschinelles Lernen inspiriert von der Funktionsweise des menschlichen Gehirns. Diese Idee war nicht neu, das Konzept der Neuralen Netze und des Maschinellen Lernens existierte in der Forschung schon einige Jahrzehnte. Nun hatte jedoch die Rechenleistung aufgeholt und man verfügte nun über einen viel größeren Schatz an Daten. Dazu kamen kleine Anpassungen in der Architektur der Kalkulationen und Modell-Optimierung. Deep Learning, wie man diese Methode bald nannte, sollte alles verändern.

2011 setzte IBMs Watson noch einen drauf und gewann in der US-Quizshow *Jeopardy!* gegen die besten menschlichen Champions. Anders als Deep Blue musste Watson nicht rechnen, sondern Sprache verstehen. Mehrdeutige Fragen voller Anspielungen und Wortspiele. Zum ersten Mal zeigte sich öffentlich, dass Maschinen mit Sprache umgehen können.

Neben dem Sprachenverständnis machten neuronale Netze und Deep Learning es erstmals möglich, visuelle Informationen auszuwerten, Maschinen konnten erkennen, was auf einem Bild zu sehen ist, etwa Gesichter oder andere Objekte. Dies wird in Zukunft ermöglichen, dass KI Schäden auf Bauteilen findet oder Abweichungen in Röntgen- und Scandaten. Bei einem internationalen Wettbewerb zur Bilderkennung halbierte ein Deep-Learning-System auf einen Schlag die Fehlerquote aller anderen menschlichen Teilnehmer. Maschinen konnten plötzlich Katzen von Hunden unterscheiden, und das sogar besser als Menschen. Das mag banal klingen, war es aber nicht. Es war der Beginn einer neuen Ära.

2016 folgte der nächste Moment, der viele aufgerüttelt hat. AlphaGo, eine Erfindung der Firma DeepMind, schlug den weltbesten

Go-Spieler. Ein Spiel, das lange als zu komplex für Maschinen galt. Bemerkenswert war das Wie: AlphaGo spielte Millionen Partien gegen sich selbst und lernte aus selbst erzeugten Daten. Diese Methode nennt man Reinforcement Learning. Manche seiner Züge hielten Experten zunächst für Fehler, bis sie sich als genial erwiesen.

Was dann 2017 passierte, ging in der breiten Öffentlichkeit weitgehend unter, ist aber der technische Ursprung von allem, was wir heute erleben. Ein Team bei Google veröffentlichte eine Arbeit mit dem Titel »Attention is all you need« und beschrieb darin eine neue Architektur für KI-Systeme, den sogenannten Transformer.[4] Ohne zu tief in die Technik einzusteigen, was ein eigenes Buch füllen würde, lässt sich sagen, dass diese Erfindung es möglich machte, Sprache in einem völlig neuen Maßstab technisch zu verarbeiten. Alle modernen KI-Sprachmodelle, die heute in aller Munde sind, basieren auf dieser Grundlage.

Der ChatGPT-Moment

Im November 2022 öffnete OpenAI die Tür in eine neue Zeitrechnung. ChatGPT erschien und erreichte in zwei Monaten 100 Millionen Nutzerinnen und Nutzer.[5] Kein Produkt in der Geschichte des Internets war bis zu diesem Zeitpunkt schneller gewachsen. Nicht Facebook, nicht Instagram, nicht TikTok. ChatGPT war nicht das erste Sprachmodell und auch nicht das technisch ausgereifteste, aber es war das erste, das sich so einfach bedienen ließ wie eine Suchmaschine. Plötzlich konnten alle mitmachen: Schüler, Manager, Ärzte, Anwälte, Handwerker. KI verließ die Kellerbüros der Forschung und zog in unseren Alltag ein.

Was ChatGPT so besonders macht, war nicht allein die Technologie. Das Trainieren eines Modells dieser Größenordnung war erst möglich geworden, weil drei Dinge gleichzeitig stimmten. Die Transformer-Architektur, die wie eben beschrieben 2017 ihren

Durchbruch feierte, lieferte die richtige Methode. Zusätzlich hatte das Internet über Jahrzehnte einen nie dagewesenen Schatz an Textdaten angehäuft, aus dem das Sprachmodell lernen konnte. Leistungsstarke Grafikprozessoren, ursprünglich für Computerspiele entwickelt, lieferten die nötige Rechenleistung. Das Trainieren kostete Hunderte Millionen Dollar. Das war kein Selbstexperiment in einem Labor, sondern ein industrieller Kraftakt, den sich kaum eine Organisation der Welt allein hätte leisten können. Ein wichtiger Faktor für den Erfolg von OpenAI war die frühere Kollaboration mit Microsoft. Zugang zu Geld und Rechenleistung machten Microsoft zum entscheidenden Partner. So war OpenAI in der Lage, als Erstes ein bevölkerungstaugliches großes Sprachmodell zu entwickeln. Mit ChatGPT hat das Unternehmen das Ergebnis danach für alle Menschen zugänglich gemacht.

Die Interaktion mit einem Algorithmus, also einer Künstlichen Intelligenz, war vor ChatGPT den meisten Menschen, die nicht programmieren können, verschlossen. Einen ähnlichen Moment gab es schon einmal: Auch das Internet existierte längst in Militär und Forschung, bis Mitte der 1990er die ersten Browser, allen voran Netscape, das Internet für alle zugänglich machten. Jetzt lohnt es sich, an die Zeit vor dem Internet zu denken und was sich in den letzten dreißig Jahren alles getan hat. Wir sind überzeugt: Rückblickend wird KI schneller und disruptiver unser Leben und unsere Unternehmen verändert haben.

Nun sind vier Jahre seit der Veröffentlichung von ChatGPT vergangen. Was danach kam, entwickelt sich in einem Tempo, das selbst Experten überrascht. Microsoft hatte bereits seit 2019 in OpenAI investiert. Als der Durchbruch kam, waren sie vorbereitet, diese neue Technologie auch gewinnbringend zu nutzen. Schon im April 2023 ermöglichte Microsoft die Nutzung des Modells GPT-3.5 über die Azure Cloud. Das führte dazu, dass deutsche Unternehmen ihr eigenes sicheres Unternehmens-ChatGPT bauen konnten. In den Medien war schnell von dmGPT, BoschGPT, AllianzGPT, ogGPT (Otto Gruppe) und E.ON GPT zu lesen.[6]

Auch die Unternehmen ohne eigenes ChatGPT mussten nicht lange warten. Innerhalb weniger Monate wurde GPT 3.5 in die gesamte Office-Welt integriert. Word, Excel, Teams, PowerPoint, Outlook – überall tauchte plötzlich der Copilot auf. Das ist das Ergebnis dieser engen Partnerschaft zwischen OpenAI und Microsoft.

Gleichzeitig entstanden neue Wettbewerber. Anthropic wurde von ehemaligen OpenAI-Mitarbeitern gegründet, die eigene Überzeugungen zur sicheren Entwicklung von KI mitbrachten, und entwickelte mit Claude eine eigene Modellfamilie. Google, das mit dem Transformer-Paper überhaupt erst die technische Grundlage für alle modernen Sprachmodelle gelegt hatte, brachte Gemini auf den Markt. Meta veröffentlichte seine Llama-Modelle als Open Source und veränderte damit die Spielregeln grundlegend. Plötzlich konnte jedes Unternehmen ein leistungsstarkes Grundmodell herunterladen, auf den eigenen Servern betreiben und an die eigenen Bedürfnisse anpassen, meist ohne Lizenzgebühren und ohne Datenweitergabe an Dritte. Selbst Elon Musk brachte mit seiner Firma xAI und dem Modell Grok einen eigenen Wettbewerber auf den Markt.

Aus Europa meldeten sich Mistral und Aleph Alpha an. Das französische Mistral zeigte, dass man mit effizienten Modellen mithalten kann, ohne die US-Riesen zu kopieren. Das deutsche Aleph Alpha, einst als europäischer Hoffnungsträger mit 500 Millionen Dollar gefeiert, konnte auf dem Massenmarkt nicht mithalten und richtet sich heute auf KI-Infrastruktur für Unternehmen und Behörden aus. Vereinfacht gesagt fehlt Europa das Investitionskapital, um vorn mitzuspielen. Deshalb lohnt es sich, die geopolitische Dimension noch kurz zu betrachten.

Geopolitische Unterschiede bei der Entwicklung von Künstlicher Intelligenz

Die dominierenden KI-Modelle kommen bislang aus zwei Regionen. Den USA und China. Auf amerikanischer Seite setzen OpenAI, Google, Anthropic und Meta die Standards. Sie nutzen dafür

ihr momentan endloses Kapital. In China ist Künstliche Intelligenz eine staatlich vorgeschriebene Initiative. China geht das Thema KI grundlegend anders an als der Westen. Seit 2017 verfolgt der chinesische Staat einen nationalen KI-Entwicklungsplan mit einem klaren Ziel. Sie wollen bis 2030 die weltweite Führungsposition in Künstlicher Intelligenz einnehmen. Was diesen Plan von westlichen Digitalstrategien unterscheidet, ist die Konsequenz der Umsetzung. Der Staat weist den größten Technologieunternehmen des Landes konkrete Aufgaben zu. Baidu forscht an autonomem Fahren, Alibaba entwickelt intelligente Städte, Tencent arbeitet an KI-gestützter Medizindiagnose, iFlyTek übernimmt Spracherkennung.[7] Im Gegenzug erhalten die Unternehmen etwas, das im Westen kaum vorstellbar wäre: staatlich organisierten Zugang zu Daten. Schon vor 15 Jahren hat man gesagt, Daten sind das neue Gold. Damit werden wir uns später noch intensiv beschäftigen.

Chinas KI-Politik erhält wenig Aufmerksamkeit in unseren Medien. 2025 sorgte dann das chinesische Unternehmen DeepSeek für Aufsehen: Sein Sprachmodell hielt in vielen Benchmarks mit GPT und Claude mit, bei berichteten Trainingskosten von rund sechs Millionen Dollar, während US-Unternehmen Hunderte Millionen investieren. Die Nvidia-Aktie verlor daraufhin an einem einzigen Tag fast 600 Milliarden Dollar an Börsenwert. Der Markt hatte angenommen, dass nur große Budgets und große Rechenleistung zu starker KI führen. DeepSeek bewies das Gegenteil.[8]

Wir sind fast am Ende unserer Zeitreise angekommen. Es gibt noch ein letztes Kapitel dieser Geschichte und es spielt sich gerade jetzt ab. 2025 tauchte eine Weiterentwicklung auf, die die KI-Welt erneut in Bewegung versetzte: KI-Agenten. Eine Weile war unklar, ob es sich um einen weiteren Trend handelt oder um eine echte technologische Trendwende. Inzwischen zeigt die Beweislage eindeutig auf eine revolutionäre Veränderung. KI-Agenten kombinieren ein Sprachmodell mit weiteren Werkzeugen, zum Beispiel einem Browser, einer E-Mail, einem Kalender oder einer Datenbank. Sie

können mit anderen Systemen interagieren und eigenständig handeln. Man gibt ihnen ein Ziel, und sie setzen es um. Das ist der entscheidende Unterschied zu allem, was vorher kam. Früher hat man eine Antwort in Text erhalten. Jetzt bekommt man eine umgesetzte Lösung. Der Schritt von einer KI, die tippt, zu einer KI, die handelt, ist größer, als er klingt. Auch hier war die Technologie schon länger da und dann kam der öffentliche Durchbruch und das ausgerechnet aus Österreich. Besser gesagt von Peter Steinberger, einem österreichischen Softwareentwickler.

Im November 2025 veröffentlichte er ClawdBot.[9] Bei ClawdBot handelt es sich um einen persönlichen KI-Assistenten. Die Anleitung dafür hat Steinberger öffentlich gemacht. Wieder überzeugte die einfache Anwendbarkeit. Man schreibt eine Nachricht in WhatsApp, gibt eine Aufgabe, und der Agent erledigt sie. Er bucht Flüge, beantwortet E-Mails, füllt Formulare aus und programmiert Software selbstständig. Nach einem Markenrechtsstreit mit Anthropic, die lamentieren, der Name sei zu nah an ihrem Modellnamen »Claude«, wurde das Produkt in OpenClaw umbenannt. Der Streit hat sicherlich auch dazu beigetragen, die Bekanntheit von KI-Agenten zu steigern.

An dieser Stelle machen wir nun den Absprung aus der Zeitreise. Als Nächstes tauchen wir in die Technologie ein. Einige Namen und Begriffe sind schon gefallen. Nun wollen wir uns anschauen, was dahinter steckt.

1.2 Was ist KI eigentlich und warum kann das niemand einfach erklären?

Stellen wir uns eine Szene vor, die sich gerade in unzähligen Unternehmen täglich abspielt. Eine Führungskraft sagt im Meeting: »Wir müssen jetzt KI einsetzen.« Alle nicken. Dann fragt jemand: »Was genau meinst du damit?« Und plötzlich reden fünf Menschen über fünf völlig verschiedene Dinge. Eine Person denkt an

ChatGPT. Eine andere an automatisierte Entscheidungen in der Kreditvergabe. Eine dritte an Roboter. Und die vierte hat gerade einen Artikel über Quantencomputer gelesen und ist sich nicht mehr sicher, ob das jetzt auch KI ist.

Das ist kein Zeichen mangelnder Intelligenz. Es ist das direkte Ergebnis davon, dass es bis heute keine einheitliche, allgemein anerkannte Definition von Künstlicher Intelligenz gibt. Selbst Wissenschaftler streiten sich darüber, was den Begriff eigentlich ausfüllt. Aus technischer Sicht beschreibt Künstliche Intelligenz Systeme, die Aufgaben erledigen, für die man früher menschliche Intelligenz gebraucht hätte. Aus wirtschaftlicher Sicht interessiert weniger, was im Inneren passiert, und mehr, welchen Nutzen ein System erzeugt. Beide Perspektiven sind legitim. Und beide führen zu unterschiedlichen Erwartungen, Prioritäten und Missverständnissen.

Was hilft, ist eine pragmatische Haltung. Künstliche Intelligenz ist keine einzelne Technologie. Es ist ein Sammelbegriff für eine ganze Familie von Technologien und Methoden, die alle dasselbe Ziel verfolgen: Maschinen beizubringen, Aufgaben zu übernehmen, die bislang menschliches Denken, Urteilen oder Lernen voraussetzten. Was genau darunter fällt, hängt vom Kontext, vom Zeitpunkt und manchmal auch vom Interesse des Erzählenden ab. Die Grenze, was als Künstliche Intelligenz bezeichnet wird, verschiebt sich konstant. Wir gewöhnen uns an gewisse Algorithmen und auch Applikationen, und gleichzeitig kommen ständig neue Technologien auf den Markt. Vor zehn Jahren hätten wir gesagt, Netflix nutzt KI, um Filme vorzuschlagen. Inzwischen haben wir uns an den Algorithmus gewöhnt, sodass wir es gar nicht mehr als KI wahrnehmen. Da es keine einheitliche Definition von Künstlicher Intelligenz gibt, lohnt es sich, im Arbeitskontext im Unternehmen eine gemeinsame Definition zu finden. Wir wollen nun einen Rahmen aufbauen, der uns hilft, den Begriff zu ordnen, ohne dabei in technische Tiefenbohrungen abzurutschen.

Ein einfaches Ordnungsmodell als Kompass im KI-Dschungel

Wir versuchen nun, den Rahmen zu bauen, den viele vermissen: eine klare Ordnung der wichtigsten Begriffe. Keine vollständige Taxonomie, sondern ein Kompass. Etwas, das hilft, neue Entwicklungen schnell einzuordnen und im nächsten Meeting nicht mehr nur zu nicken, sondern auch zu verstehen.

Den äußersten Kreis bildet Künstliche Intelligenz als Überbegriff. Alles, was wir in diesem Buch besprechen, fällt darunter. KI bezeichnet alle Systeme, die intelligentes Verhalten simulieren oder ermöglichen.[10]

Innerhalb dieses Kreises liegt Machine Learning. Im Gegensatz zu klassischen, regelbasierten Systemen wird ein System des maschinellen Lernens nicht mit festen Wenn-dann-Logiken programmiert, sondern es lernt aus Daten und passt sein Verhalten statistisch an. Das ist der Unterschied zur regelbasierten Prozessautomatisierung. Solche Systeme führen vorab definierte Regeln zuverlässig und deterministisch aus. Ob man diese Form der Automatisierung allerdings zur Künstlichen Intelligenz zählt, ist umstritten. Aber für uns ist weniger die Begriffsfrage als das Wirkprinzip entscheidend.

Während regelbasierte Automatisierung starr bleibt, arbeitet maschinelles Lernen mit Wahrscheinlichkeiten. Es ist nicht immer »richtig«, sondern wird mit mehr Daten besser. Dieser Lernmechanismus bildet den Kern nahezu aller modernen KI-Anwendungen, von neuronalen Netzen bis hin zu großen Sprachmodellen. Diese statistischen Algorithmen des maschinellen Lernens erweitern die Möglichkeiten von KI erheblich: Statt dass Menschen jede Eventualität in starre Regeln übersetzen müssen, lernen die Systeme selbst aus Daten, welche Muster relevant sind, und entwickeln ihre Leistungsfähigkeit aus diesen Mustern.

Eine besondere Klasse des maschinellen Lernens sind neuronale Netze. Sie sind an die Funktionsweise des menschlichen Gehirns

angelehnt und bestehen aus vielen miteinander verbundenen Recheneinheiten, die Informationen schrittweise verarbeiten. Entscheidend ist dabei nicht die einzelne Verbindung, sondern die Gesamtheit des Netzes: Je größer und tiefer ein solches Netzwerk wird, desto komplexere Muster kann es erkennen.[11]

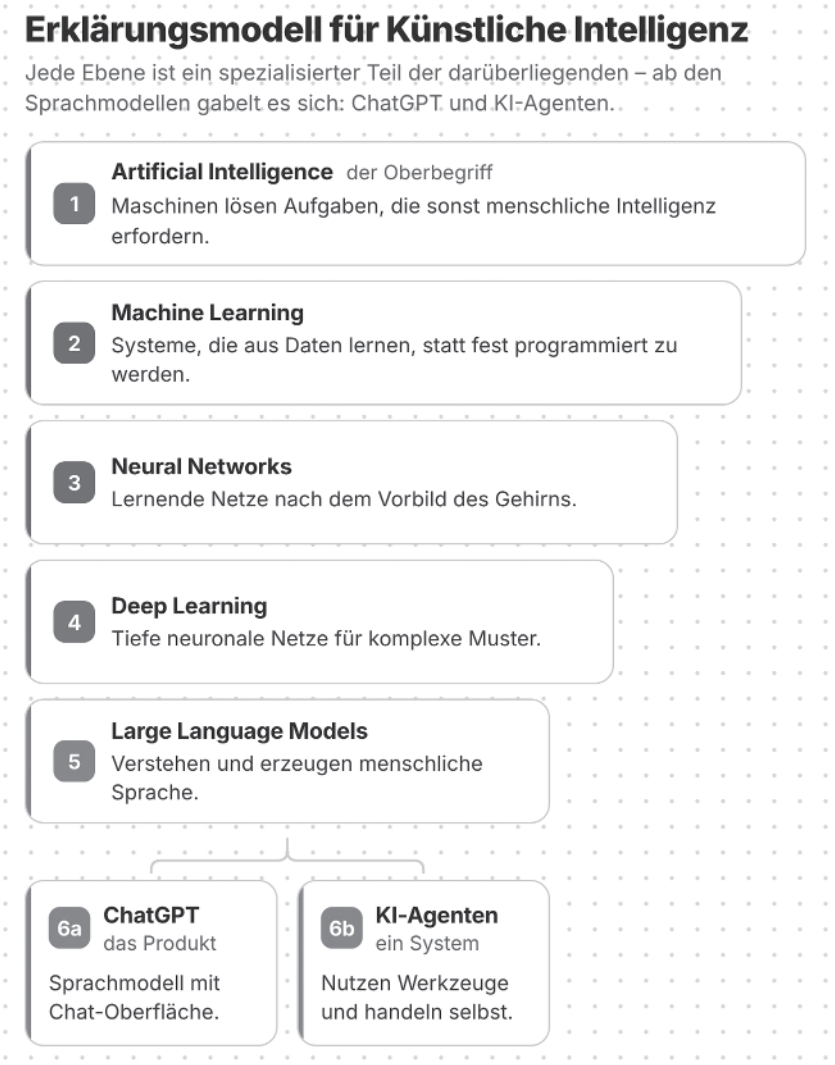


Abbildung 1.1: Erklärungsmodell für Künstliche Intelligenz

Genau hier setzt Deep Learning, das wir in Abschnitt 1.1 erwähnt haben, an. Der Begriff beschreibt keine neue Methode, sondern den Punkt, an dem neuronale Netze so groß, so tief und mit so vielen Daten trainiert werden, dass ihre Leistungsfähigkeit sprunghaft zunimmt. Sprache, Bilder oder komplexe Zusammenhänge lassen sich so nicht mehr nur grob erfassen, sondern in einer Qualität, die zuvor technisch nicht erreichbar war.

Noch eine Ebene tiefer landen wir bei der Generativen KI. Das ist der Teil, der die letzten Jahre dominiert hat. Sie ist keine neue Form von Künstlicher Intelligenz, sondern eine Weiterentwicklung derselben Prinzipien, die wir bereits kennengelernt haben. Auch sie basiert auf maschinellem Lernen und neuronalen Netzen. Der Unterschied liegt nicht im »Ob«, sondern im »Was«. Generative KI erzeugt neue Inhalte: Texte, Bilder, Code, Audio, Video. Sie lernt nicht nur, Bestehendes zu klassifizieren, sondern etwas Neues zu schaffen, das dem ähnelt, womit sie trainiert wurde.

Innerhalb der Generativen KI sind Large Language Models, kurz LLMs, die bekannteste Unterkategorie. GPT-4, Claude oder Gemini sind solche Modelle: sehr große neuronale Netze, trainiert auf enormen Textmengen, die Sprache analysieren und erzeugen können, die sich für Menschen natürlich anfühlt.

Noch einen Schritt weiter landen wir bei KI-Agenten. Auch sie sind keine neue Form von Künstlicher Intelligenz, sondern eine Weiterentwicklung generativer Modelle. Im Kern bleibt das Sprachmodell gleich. Der Unterschied liegt darin, wie es eingesetzt wird. Ein KI-Agent kombiniert ein großes Sprachmodell mit weiteren Werkzeugen, etwa einem Browser, einem E-Mail-Postfach, einem Kalender oder einer Datenbank und kann diese selbstständig nutzen. Während ein klassisches Sprachmodell auf Fragen reagiert, arbeitet ein KI-Agent zielorientiert. Man gibt ihm ein Ziel, nicht jeden einzelnen Schritt. Der Agent plant den Weg dorthin, führt Aufgaben aus und passt sein Vorgehen an, bis das Ziel erreicht ist.

Der entscheidende Fortschritt liegt damit nicht im »Denken«, sondern im Handeln. KI-Agenten verschieben KI von der reinen Antwortmaschine hin zu Systemen, die Arbeit tatsächlich erledigen.

Wer KI im Unternehmen ernsthaft einsetzen will, wird um dieses Thema in den nächsten Jahren nicht herumkommen. In unserem Modell gehören die KI-Agenten in die Ebene der LLMs. So wie ChatGPT ein Produkt ist, das ein LLM (nämlich das LLM GPT) und einen Chat nutzt, so ist ein KI-Agent ein Chat, ein LLM und verschiedene Zugänge zu anderen Werkzeugen. So sind KI-Agenten im Kern eine Erweiterung von Sprachmodellen. Sie können beispielsweise dieselben Sprachmodelle wie ChatGPT nutzen, werden jedoch so eingebettet, dass sie zusätzliche Werkzeuge eigenständig verwenden können, etwa um Informationen abzurufen, Systeme zu bedienen oder Aufgaben auszuführen.

Mit der Abbildung 1.2 wollen wir auch aufzeigen, dass KI so viel mehr kann als Generative KI. Die Ringe sind auch proportional zu verstehen zu der Wichtigkeit des Bereiches. Das Schöne an diesem Ordnungsmodell ist, dass es nicht starr ist. Wenn morgen ein neuer Begriff auftaucht, lässt er sich meistens in diesem Rahmen einordnen. Und das ist eigentlich das Wichtigste: nicht jedes Detail zu kennen, sondern zu wissen, wo man neue Information verankern kann.

Was dieses Modell jedoch nicht zeigt, ist die Breite dessen, was KI konkret bereits heute kann. Das Ordnungsmodell erklärt die Struktur. Die praktischere Frage für Führungskräfte ist oft eine andere: Wofür kann ich KI eigentlich einsetzen? Wenn man den Blick auf das gesamte Spielfeld richtet, wird schnell klar, wie weit das Feld reicht. KI wird oft mit ChatGPT und Chatbots gleichgesetzt, jedoch geht es viel weiter.



Abbildung 1.2: KI-Fähigkeiten auf einen Blick

Diese Breite hat praktische Konsequenzen. Wer KI nur mit dem letzten Punkt der Liste gleichsetzt, übersieht elf weitere Felder, auf denen KI bereits längst arbeitet, oft unsichtbar, oft tief in bestehende Geschäftsprozesse eingebettet. Das Ordnungsmodell hilft einzuordnen, welche wissenschaftlichen Ebenen hinter der KI-Technologie stecken. Diese Übersicht hilft zu verstehen, was KI für ein Unternehmen tun kann.

Um überhaupt eine der KI-Fähigkeiten nutzen zu können, braucht es Daten. Daten sind das neue Gold. Diesen Satz hören wir seit vielen Jahren und er ist so wichtig und richtig wie nie zuvor. Daten sind das Rohmaterial der Künstlichen Intelligenz. Ohne Daten keine Muster. Ohne Muster kein Lernen. Ohne Lernen keine KI. Es lohnt sich deshalb kurz innezuhalten und zu verstehen, womit KI-Systeme eigentlich arbeiten. Jetzt mag man denken: Stimmt, da war doch etwas mit 0 und 1. Damit rechnen Computer doch, oder? Ja. Im Grundsatz ist das weiterhin richtig. Der entscheidende Unterschied liegt jedoch darin, wie KI diese Daten nutzt. An dieser Stelle werden wir oft gefragt: Wann ist KI eigentlich zu 100 Prozent richtig?

Die ehrliche Antwort lautet: bei lernenden Systemen fast nie. KI arbeitet, sofern man rein regelbasierte Systeme ausklammert, nicht mit klaren Ja-Nein-Entscheidungen, sondern mit Wahrscheinlichkeiten. Sie bewertet, wie gut etwas zu den Daten passt, mit denen sie gelernt hat, nicht, ob etwas eindeutig richtig oder falsch ist. Deshalb wird KI bei lernenden Systemen nie zu 100 Prozent korrekt sein. Zugleich liegt darin ihre Stärke: Sie kann mit komplexen, unvollständigen oder widersprüchlichen Informationen umgehen, an denen regelbasierte Systeme scheitern. Wie zuverlässig diese Einschätzungen sind, hängt maßgeblich von der Qualität und Relevanz der zugrunde liegenden Daten ab.

Daten kommen in sehr unterschiedlichen Formen. Zahlen sind die offensichtlichste Kategorie, wie zum Beispiel Verkaufszahlen, Temperaturen, Laufzeiten und Klickraten. Textdaten umfassen alles, was in geschriebener Sprache existiert, von E-Mails über Verträge bis hin zu Kundenbewertungen. Bilddaten ermöglichen es Systemen, Fotos, Scans oder Videos zu analysieren, also Produktionsfehler auf Förderbändern zu erkennen oder Gesichter in Zugangssystemen zu identifizieren. Audiodaten liegen jedem Sprachassistenten zugrunde und erlauben es, gesprochene Sprache zu verstehen oder zu erzeugen. In mancher Fachliteratur werden Audiodaten als Textdaten klassifiziert, da sie in Audiodateien

transkribiert werden. Das gleiche Argument wird für Videodaten verwendet, auch hier wird es als Mischung aus Bild- und Textdaten gesehen. Dann unterscheiden wir noch in strukturierte Daten und unstrukturierte Daten. Die meisten Daten sind unstrukturiert. Das war lange ein Hindernis für KI. Strukturierte Daten sind das, was in Tabellenform lebt, sauber formatiert in Datenbanken, gut sortiert und direkt verarbeitbar. Egal ob numerische Daten, Text- oder Bild-Daten – alle Arten werden am Ende in Zahlen transferiert, damit die KI damit weiter rechnen kann.

Das alles klingt sehr technisch, ist es aber eigentlich nicht. Die entscheidende Frage für Führungskräfte ist deshalb eine ganz andere: Welche Daten haben wir und sind sie »gut« genug? Denn ein KI-System ist immer nur so gut wie das, womit es gelernt hat.

Dabei geht es nicht nur um technische Qualität, sondern auch um Klarheit in den Prozessen. Wenn schon Menschen sich nicht einig sind, wie etwas zu bewerten ist, spiegelt sich diese Uneinheitlichkeit zwangsläufig in den Daten wider. Fragt man zehn Sachbearbeiter, was als Rechnung oder als Beschwerde gilt, und erhält fünf unterschiedliche Antworten, dann lernt auch die KI diese Widersprüche.

Lernende Systeme können solche Unschärfen bis zu einem gewissen Grad ausgleichen. Sie können aber keinen schlechten oder inkonsistenten Prozess in einen guten verwandeln.

Schlechte Daten erzeugen schlechte Ergebnisse, egal wie ausgereift die Methode dahinter ist. Das werden wir in späteren Kapiteln noch sehr konkret spüren.

Es ist alles Mathematik

Jetzt kommt der Teil, bei dem manche reflexartig wegscrollen würden: Mathematik. Aber keine Sorge, wir werden keine Gleichungen lösen. Uns geht es darum zu verstehen, was unter der Motorhaube steckt. So viel wie nötig und so wenig wie möglich.

Künstliche Intelligenz ist im Kern angewandte Statistik. Sehr viel davon, sehr schnell und mit sehr vielen Daten. Während klassische Computersysteme mit festen Wenn-dann-Logiken arbeiten, lernen KI-Systeme anders. Sie erkennen Muster in großen Datenmengen und leiten daraus Vorhersagen ab.

Nehmen wir einen E-Mail-Filter, der Spam erkennen soll. Wir zeigen ihm tausende E-Mails und markieren, welche Spam sind und welche nicht. Das System analysiert Muster, etwa typische Wörter, Absender oder Formulierungen, und baut daraus ein Modell, das neue E-Mails nicht eindeutig einordnet, sondern mit einer bestimmten Wahrscheinlichkeit bewertet.

Genau daran knüpft der Gedanke aus dem vorherigen Abschnitt an: KI liefert keine Wahrheiten, sondern Wahrscheinlichkeiten. Wenn ein Sprachmodell antwortet, berechnet es im Kern, welches Wort als nächstes am besten passt, basierend auf allem, womit es gelernt hat. Das klingt zunächst ernüchternd, ist es aber nicht. Denn auch Menschen urteilen, entscheiden und kommunizieren auf Basis von Erfahrungen und Mustern, nicht auf Grundlage lückenloser Logik.

Mustererkennung und Optimierung sind deshalb die beiden Grundprinzipien hinter fast allem, was wir heute als KI bezeichnen. Ein System, das Röntgenbilder auf Auffälligkeiten prüft, erkennt Muster. Ein System, das Lieferketten optimiert, minimiert eine Zielfunktion. Ein System, das Texte zusammenfasst, gewichtet und priorisiert Informationen. Es ist immer dieselbe Grundidee, angewendet auf unterschiedliche Problemstellungen.

Wenn es im Grunde immer der gleichen Logik folgt, wieso gibt es dann trotzdem so viele verschiedene KI-Begriffe? Zum Anfang des Kapitels haben wir das Ordnungsmodell eingeführt. Damit lassen sich die Begriffe in Relation zueinander erklären. Die nächste Verständnisebene bezieht sich auf die verschiedenen Arten der Daten. Unterschiedliche Datenarten verlangen unterschiedliche Werkzeuge. Daher hat jede dieser Disziplinen ihren

eigenen Namen bekommen. Ein Large Language Model ist schlicht die Antwort auf die Frage, wie eine Maschine Sprache und Text verarbeitet. Computer Vision ist die Antwort auf dieselbe Frage, gestellt für Bilder. Predictive Analytics beantwortet sie für Zahlen und Messwerte.

Wer das einmal verstanden hat, kann neue Begriffe sofort einordnen. Taucht ein unbekanntes KI-Schlagwort auf, lohnt sich eine einzige Frage: Mit welcher Art von Daten arbeitet dieses System? Die Antwort zeigt meistens sehr schnell, womit man es zu tun hat. Das erklärt auch, wieso es so viele verschiedene Namen für KI-Produkte im Markt gibt.

Dazu noch eine wichtige Ergänzung. Moderne KI-Systeme arbeiten zunehmend mit mehreren Datenarten gleichzeitig. Aktuelle Sprachmodelle können Texte lesen, Bilder beschreiben und Tabellen auswerten, alles in einem Gespräch. Das nennt man multimodale KI.

Übersicht der verschiedenen Datenarten

Texte, Bilder und Zahlen sind typische Datenarten – jede hat ihre eigene KI-Disziplin und Modellgruppe.

1 · DATENART	2 · KI-DISZIPLIN (FACHBEGRIFF)	3 · SUBKATEGORIE / MODELLGRUPPE
Texte Unstrukturiert	**Natural Language Processing (NLP)**	z. B. Large Language Models (LLMs)
Bilder Unstrukturiert	**Computer Vision**	z. B. Generative / Convolutional Models
Zahlen Strukturiert	**Predictive Analytics**	z. B. Zeitreihen-, Regressions- und Klassifikationsmodelle

Abbildung 1.3: Übersicht der verschiedenen Datenarten

Was Führungskräfte nicht brauchen, sind die technischen Unterkategorien dahinter. Spracherkennung, Edge-KI oder industrielle KI sind Spezialistenterritorium – das müssen Teams und Anbieter beherrschen.

Was zählt, ist die Fähigkeit, eine KI-Initiative einzuordnen: Welche Daten liegen zugrunde? Welche Modelle kommen von welchem Anbieter zum Einsatz? Welches Werkzeug wird genutzt? Und was kann das gebaute System realistisch leisten?

Erst wenn diese Fragen beantwortet sind, wird klar, an welchen Stellschrauben gedreht werden kann, damit KI für das Unternehmen relevant wird. Ob es um bessere Datenverfügbarkeit geht, um höhere Datenqualität, um klarere Prozesse oder um realistischere Erwartungen an das System, all das sind Führungsentscheidungen, nicht technische Details.

Prompting: Die Fähigkeit der Zukunft?

Mit dem Verständnis dafür, wie Generative KI und große Sprachmodelle funktionieren, wird auch klar, warum sich ab 2023 der Begriff »Prompting« etabliert hat.

Ein Prompt ist die Texteingabe in ein Sprachmodell. Es ist also die Frage, die Anweisung oder Aufgabe, die ein Mensch der Maschine gibt. Mit der Fähigkeit, natürliche Sprache zu verstehen, brach eine sehr große technologische Barriere ein. Jedoch öffnete sich gleichzeitig eine neue Herausforderung. Die Kommunikation mit dem Sprachmodell bedarf etwas Übung. Wie in einem guten Bewerbungsinterview liegt der entscheidende Unterschied in den richtigen Fragen.

Vielen Menschen gelang es schnell, ihre Texteingaben so zu formulieren, dass brauchbare Ergebnisse entstanden. Andere wiederum taten sich schwer. Da es doch nicht so intuitiv ist, mit einer Maschine zu reden wie mit einem Menschen, sind sehr viele Trainings zum Thema Prompting entstanden. Dazu haben viele Unternehmen

sogenannte »Prompt Libraries« erstellt. Also Sammlungen an Eingabebefehlen, die dann gute Ergebnisse produzieren. In den ersten zwei Jahren nach dem Start von ChatGPT wurde die Fähigkeit zu prompten beinahe verheiligt. Es gab sogar Diskussionen, ob es neue Berufe geben wird wie den »Prompt Engineer«. Heute, nochmal zwei Jahre später, ist diese Fähigkeit schon wieder aus den Gesprächen verschwunden. Unsere Vorhersage ist, dass dies auch mehr und mehr passieren wird. Prompting wird eine Fähigkeit sein wie das Schreiben mit zehn Fingern. Ohne sie hätte man sich auf kaum eine Stelle beworben – und im Lebenslauf hätte sie eher ein Schmunzeln als Anerkennung ausgelöst.

Bedeutet das im Umkehrschluss, dass wir empfehlen, als Führungskraft das bereits gebuchte Prompt-Training abzusagen? Ein klares Nein. Wir werden uns später noch mit dem Technologie-Einführungs-Modell beschäftigen. Die Annahme neuer KI-Lösungen hängt sehr davon ab, wie nützlich wir Menschen die Lösung finden. Ein Training, das dabei hilft, mehr Nutzen aus der Technologie zu ziehen, erhöht die Annahme und sichert damit auch das Investment. Ein gutes Prompt-Training vermittelt das Handwerkszeug. Es erklärt aber auch die Logik, wie KI-Systeme funktionieren. Es ist essenziell zu verstehen, was ein Sprachmodell weiß – und was eben nicht. Wie beschrieben hängt das maßgeblich davon ab, auf welchen Daten das Modell trainiert wurde. In jedem Fall ist es unwahrscheinlich, dass ein solches Modell fundierte Kenntnisse über das eigene Unternehmen, die Kunden oder den konkreten Arbeitsalltag besitzt. Diese Informationen müssen mitgegeben werden, um gute Antworten zu erhalten.

Gleichzeitig ist wichtig zu verstehen, dass ein Sprachmodell im Kern genau das ist: ein Modell für Sprache. Seine eigentliche Stärke liegt darin, Texte zu erzeugen, die sprachlich plausibel sind, nicht darin, Wissen zu prüfen oder Fakten zu verifizieren. Es handelt sich um ein statistisches Modell, das Wahrscheinlichkeiten berechnet.

Dieses Defizit lässt sich jedoch gezielt ausgleichen. Sprachmodelle können mit zusätzlichem Wissen versorgt werden, etwa über Dokumente, Datenbanken oder andere angebundene Quellen. Erfolgt diese Anbindung sauber, können Modelle sehr wohl unternehmensspezifisches Wissen berücksichtigen. Fehlt dieser Kontext, greifen sie auf das zurück, was sie am besten können: sprachlich überzeugende Antworten zu formulieren. Aus diesem Sachverhalt heraus erklären sich auch sogenannte Halluzinationen, also plausibel klingende Aussagen, denen die faktische Grundlage fehlt.

Deshalb ist der Umgang mit Sprachmodellen keine Frage von Tricks, sondern von Verständnis. Wer weiß, wie ein Modell arbeitet, kann gezielt beeinflussen, welchen Kontext es bekommt, welche Aufgabe es lösen soll und wo seine Grenzen liegen.

Prompting ist damit kein Selbstzweck, sondern ein Hilfsmittel, um Sprachmodelle sinnvoll zu steuern und ihre Stärken nutzbar zu machen.

Als Wiederholung oder für eine schnelle Einführung haben wir die Acht Grundsätze eines guten Prompts in Abbildung 1.4 zusammengefasst.

Es gibt kein perfektes Prompting-Regelwerk. Erfahrung und Ausprobieren sind durch nichts zu ersetzen. Es gilt, sich nicht entmutigen zu lassen von der ersten schlechten Antwort, sondern weiter zu feilen an der Anweisung.

Unter allen genannten Punkten ist der Kontext die wichtigste Information für die Künstliche Intelligenz. Um diesen Schritt einfacher zu machen, haben viele KI-Lösungen noch einen sogenannten »System Prompt« eingeführt. Dabei handelt es sich um eine Arbeitsanweisung, die die KI jedes Mal liest, bevor sie die Aufgabe des Nutzers ausführt. System-Prompts erhöhen die Qualität, da mehr Kontext vermittelt wird. Sie sparen zudem Zeit, da die Nutzerinnen nicht immer alle Informationen mitgeben müssen. Ein LLM-getriebener Chatbot wie ChatGPT kann eine gewisse Menge an Informationen über einen Unterhaltungsverlauf im Gedächtnis



Abbildung 1.4: Acht Grundsätze eines guten Prompts

behalten. Diese Informationen werden dann mitverwendet. Jedoch sind diese sogenannten Kontextfenster klein. So kann es vorkommen, dass man eine wirklich gute Unterhaltung führt und auf einmal die Antworten immer generischer werden.

Neben der Nutzer-Eingabe, dem Prompt und dem System-Prompt, der potenziell dahinter liegt, bekommt das Thema Context Engineering mehr und mehr Aufmerksamkeit. Hinter dem Begriff versteckt sich das Bestreben, den KI-Modellen möglichst viele Informationen zu geben, damit sie besser arbeiten können. Das geht weit über Texteingaben hinaus. Dabei stellt man sich auch die Frage, auf welche Daten, Wissensdatenbänke und andere Informationen die KI-Modelle zugreifen dürfen. Besonders im Kontext von KI-Agenten wird es zu einer kritischen Fähigkeit, dem Agenten die relevanten Informationen zu geben.

Von ChatBots zu KI-Agenten

Kommen wir nun von ChatBots zu dem Thema KI-Agenten. Es ist Montagmorgen. Eine Führungskraft bittet einen Chatbot um einen Überblick über offene Kundenanfragen. Der Chatbot fasst zusammen, listet Kategorien auf und formuliert eine saubere Antwort. Hilfreich, aber danach endet seine Arbeit. Ein KI-Agent würde anders vorgehen. Er würde sich selbstständig im Ticketsystem anmelden, neue Anfragen prüfen, relevante Informationen aus früheren Fällen heranziehen, Entwürfe für Antworten vorbereiten und diese zur Freigabe vorlegen. Tauchen Rückfragen auf, stellt er sie gezielt. Ändert sich die Situation, passt er sein Vorgehen an.

Der Unterschied ist entscheidend: Ein Chatbot antwortet. Ein KI-Agent handelt.

Kommen wir zurück zu den KI-Agenten, auch als agentische Künstliche Intelligenz bezeichnet. Vereinfacht gesagt lassen sie sich mit einem cleveren Werkstudenten vergleichen. Ein Sprachmodell, also ein LLM, wird mit weiteren Werkzeugen kombiniert, etwa einem Internetbrowser, einem E-Mail-Postfach oder einer Datenbank.

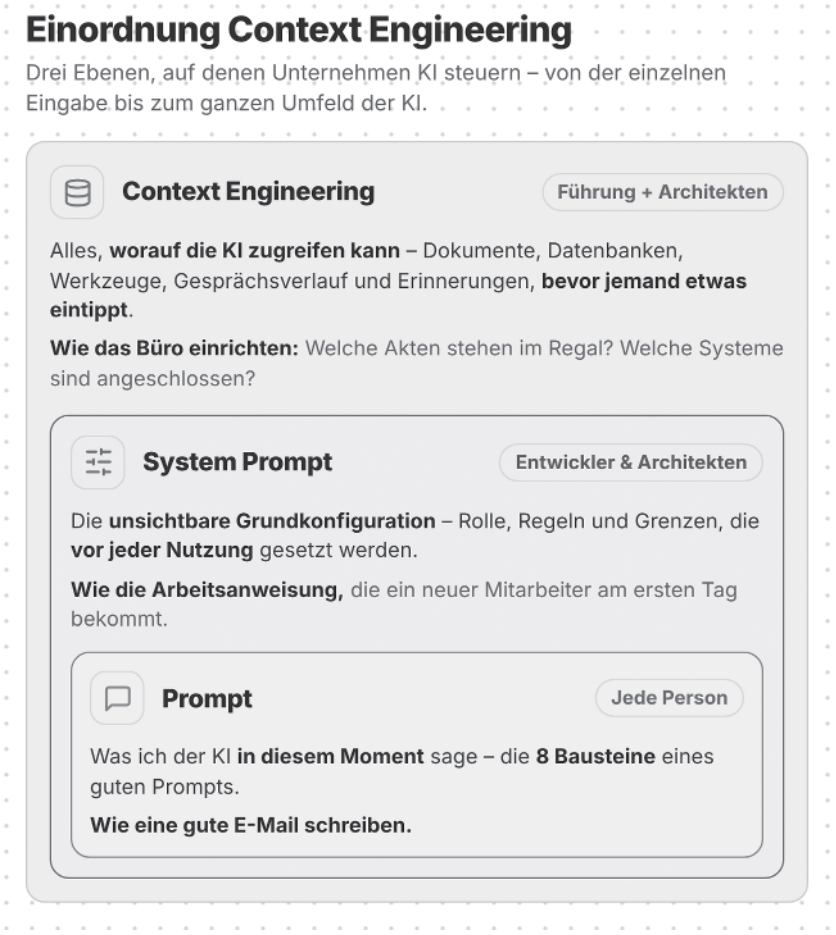


Abbildung 1.5: Einordnung Context Engineering

Das Sprachmodell ist dabei das »Gehirn«. Es plant, entscheidet und steuert die Werkzeuge. Während ein reines Sprachmodell vor allem auf Fragen reagiert, gibt man einem KI-Agenten ein Ziel. Den Weg dorthin bestimmt er selbst. So entsteht der große Schritt von reinen Antworten hin zur eigenständigen Umsetzung.

Wie ein Werkstudent müssen auch KI-Agenten eingewiesen und angelernt werden. Ist diese Grundlage geschaffen, können sie die

persönliche Produktivität deutlich steigern. Gleichzeitig lassen sich KI-Agenten in Prozessen einsetzen: Sie erledigen einzelne Schritte autonom oder führen, in klar abgegrenzten Fällen, ganze Abläufe selbstständig aus.

Mit wachsender Verantwortung stellt sich jedoch dieselbe Frage wie bei menschlichen Mitarbeitenden: Wem geben wir welchen Zugriff? Einem Werkstudenten würden wir nicht einfach alle Passwörter anvertrauen. Diese Logik gilt auch für KI-Agenten.

Regeln, Arbeitsmittel und Grenzen müssen klar definiert sein. Eine bewusste Einschränkung dessen, was ein KI-Agent darf und was nicht, ist kein Misstrauen, sondern Voraussetzung für einen sicheren und sinnvollen Einsatz. Sie schützt vor Fehlern und schafft Vertrauen.

OpenClaw ist aktuell einer der bekanntesten KI-Agenten. Da es sich jedoch um ein öffentlich entwickeltes Produkt handelt, ist fraglich, wie relevant es langfristig im Unternehmenskontext sein wird.

Genau genommen ist auch Copilot ein KI-Agent. Im Hintergrund arbeitet ein Sprachmodell, das mit Werkzeugen wie einem Internetbrowser oder weiteren Systemen verknüpft ist. Microsoft und viele andere Anbieter ermöglichen inzwischen zudem die Entwicklung eigener KI-Agenten. Neu ist dabei vor allem eines: Der Bau solcher Agenten wird zunehmend auch für Nicht-Entwickler möglich. Über Low-Code- und No-Code-Ansätze lassen sich KI-Agenten konfigurieren, an Datenquellen anbinden und mit klaren Aufgaben versehen, ohne selbst Software schreiben zu müssen. Damit rückt die Gestaltung von KI-Agenten näher an Fachbereiche und Führungskräfte heran.

Von LLM-getriebenen Chatbots wie ChatGPT wissen wir, wie entscheidend der Kontext für gute Ergebnisse ist. Bei KI-Agenten wird dieser Kontext zunehmend erweitert. Informationen werden nicht nur genutzt, sondern auch strukturiert abgelegt. Das reduziert eines der zentralen Probleme klassischer Chatbots: begrenzte Kontextfenster.

Neben Prompt und System-Prompt spricht man bei agentischer KI deshalb auch von »Context Engineering«, also der gezielten Modellierung von Kontext. KI-Agenten lassen sich mit relevanten Informationen versorgen, die dauerhaft verfügbar bleiben. Abhängig vom Ziel greifen sie dann selbstständig auf passende Informationsquellen zu.

So können Aufgaben nicht nur schneller, sondern auch in gleichbleibend höherer Qualität erledigt werden.

Als Gedankenstütze zum Merken: Ein Chatbot folgt einem definierten Pfad. Endet dieser, stößt der Bot an seine Grenzen. Ein KI-Agent hingegen kennt ein Ziel und arbeitet so lange darauf hin, bis dieses erreicht ist.

KI-Agenten können dabei zunehmend komplexe Aufgaben bewältigen. In ersten Messungen zeigen sich Szenarien, in denen ein Agent über mehrere Stunden hinweg selbstständig an einer Aufgabe arbeitet. Das ist weniger als Versprechen zu verstehen, sondern als Hinweis auf die Richtung, in die sich diese Systeme entwickeln.

Als nächster Schritt rücken Agentensysteme in den Fokus. Dabei arbeiten mehrere spezialisierte Agenten zusammen, tauschen Ergebnisse aus und überprüfen sich gegenseitig. So lassen sich Prinzipien wie Arbeitsteilung oder sogar ein Vier-Augen-Prinzip technisch abbilden.

Wohin sich diese Entwicklung genau bewegt und wie KI-Agenten künftig mit Menschen und Teams interagieren werden, ist offen. Zunächst werden sie in bestehende Strukturen integriert. Im nächsten Schritt passen wir unsere Arbeitsweisen an, damit Mensch und KI besser zusammenarbeiten können. Und perspektivisch ist auch denkbar, dass Organisationen entstehen, die von Grund auf für den Einsatz von KI-Agenten konzipiert sind. Hier liegen die eigentlichen Potenziale von KI-Agenten. Sie sind nicht nur ein weiteres Werkzeug, sondern eröffnen mehrere neue

Möglichkeiten für Automatisierung. Erstmals lassen sich komplexe Aufgaben nicht nur unterstützen, sondern weitgehend selbstständig ausführen, über Systemgrenzen hinweg, über längere Zeiträume und mit klarer Zielorientierung.

Damit werden Formen von KI-Automatisierung möglich, die mit klassischen Ansätzen kaum erreichbar waren: flexibler, anpassungsfähiger und in großem Maßstab einsetzbar. KI-Agenten schaffen so die Grundlage für leistungsfähige Automatisierungs-Anwendungen, die weit über einzelne Use Cases hinausgehen und KI als produktive Kraft im Unternehmen verankern.

1.3 Richtig oder wahrscheinlich? Warum KI kein Taschenrechner ist

Hat die KI gelogen?

2023 reichte der amerikanische Anwalt Steven Schwartz beim Bundesgericht in New York eine Klageschrift ein. Er hatte ChatGPT genutzt, um relevante Gerichtsurteile zu recherchieren, und die vorgeschlagenen Fälle ungeprüft in seine Klageschrift übernommen. Sechs Fälle mit überzeugenden Namen, Aktenzeichen, Zitaten und richterlichen Begründungen. Was er nicht wusste: Keiner dieser Fälle hatte je existiert. Kein einziger. Der zuständige Richter P. Kevin Castel bezeichnete die Dokumente als »legal gibberish«, als juristischen Unsinn. Schwartz wurde zu einer Geldstrafe von 5 000 Dollar verurteilt.[12]

Hat die KI also gelogen? Nein. Die KI hat genau das getan, wofür sie gebaut wurde. Das Problem war das Missverständnis, wie diese KI eigentlich funktioniert. Um das zu verstehen, steigen wir tiefer in die Funktionsweise ein. Wir schauen uns an, welche Systeme deterministisch agieren und welche probabilistisch. Und wir sprechen über einen Begriff, der seit diesem Fall in keiner KI-Diskussion mehr fehlt: Halluzination.

Der Taschenrechner – ein deterministisches System

Ein Taschenrechner lügt nicht. Tippe 7 mal 8 ein, und das Ergebnis ist 56. Heute. Morgen. In zehn Jahren. Auf jedem Gerät der Welt. Das nennt sich Determinismus: Gleicher Input ergibt immer denselben Output. Jedes Mal. Ohne Ausnahme.

Diese Logik kennen wir aus dem Unternehmensalltag. Das ERP-System, das Lagerbestände berechnet. Das GPS, das eine Route plant. Excel, wenn es eine Formel ausführt. All diese Systeme folgen festen Regeln, die jemand explizit einprogrammiert hat. Das Ergebnis ist vorhersehbar, auditierbar und im Zweifelsfall erklärbar. Wenn etwas falsch herauskommt, liegt es an den Daten oder den Regeln, aber das System hat nie »geraten«.

Das ist die technologische Welt, in der die meisten Führungskräfte Technologie gelernt haben: verlässlich. Entweder richtig oder falsch. Diese Erwartung macht den Übergang zu KI so schwierig. Denn KI folgt einer anderen Logik. Wer sich noch an Abschnitt 1.1 erinnert, kennt dieses Muster aus der Geschichte. Die Expert Systems der 1980er Jahre waren ebenfalls deterministisch. Experten codierten ihr Wissen in Wenn-dann-Regeln, und das System folgte ihnen. Das war verlässlich und starr. Die Realität ließ sich aber nie vollständig in Regeln fassen, und deshalb war die Starrheit, die aus den genauen Regeln folgte, ihr Verhängnis.

Und diese deterministischen Systeme haben Enormes geleistet. Sie sind der Grund, warum wir Milliarden von Transaktionen täglich digital und automatisch verarbeiten können, warum Lieferketten funktionieren, warum ein Gehalt pünktlich auf dem Konto landet. Verlässlichkeit ist kein kleines Feature. In regulierten Branchen wie Finanzen, Gesundheit oder Energie ist sie Grundbedingung. Ein System, dessen Entscheidungen sich lückenlos nachvollziehen und erklären lassen, ist in vielen Kontexten genau das Richtige.

Beispiele gibt es überall: Buchungssysteme, die Verfügbarkeiten prüfen. Betrugserkennung, die nach fest definierten Schwellenwerten flaggt. Ampelschaltungen. Payroll-Systeme. Einfache Entscheidungsbäume in alten Chatbots, die bei »Frage 3« immer zur selben Antwort kommen. All diese Systeme folgen derselben Grundlogik wie der Taschenrechner.

Und jetzt kommt die Frage, die es sich lohnt zu stellen: Ist das eigentlich KI? Die Antwort ist nicht eindeutig, hier streiten sich die Geister. Viele würden sagen: nein. Deterministische Systeme sind klassische Software. Jemand hat die Regeln aufgeschrieben, und das System führt sie aus.

In diesem Buch sprechen wir von KI dort, wo ein System nicht mehr nur Regeln befolgt, sondern aus Daten lernt und daraus selbst Schlüsse zieht. Im Hintergrund werden Wahrscheinlichkeiten berechnet, deshalb bezeichnen wir solche Systeme als probabilistisch.

Gleichzeitig bedienen sich moderne KI-Systeme sehr wohl auch deterministischer Regeln. Dort, wo Ergebnisse eindeutig sind oder sicherheitskritisch entschieden werden muss, werden feste Logiken bewusst eingesetzt, um Präzision und Verlässlichkeit zu erhöhen. In der Praxis entstehen so hybride Systeme, die probabilistisches Lernen mit deterministischer Steuerung kombinieren. Also ist die Antwort aus unserer Sicht ein klassisches Jein.

Probabilistisch heißt, plausibel ist nicht gleich korrekt

Im Hintergrund werden für jede einzelne Entscheidung Wahrscheinlichkeiten kalkuliert. Daraus ergeben sich zum Beispiel bei einem Sprachmodell die erzeugten Antworten: Jedes Wort wird auf Basis der Wahrscheinlichkeiten ausgewählt und eingefügt. Das führt dazu, dass trotz gleicher Eingabe ein anderes Ergebnis entstehen kann. Und es führt auch dazu, dass Antworten plausibel klingen können, ohne korrekt zu sein.

Basierend auf dem bisherigen Kapitel könnte man nun meinen, KI sei immer probabilistisch. Das wäre zu kurz gesprungen. KI-Systeme und KI-Lösungen kombinieren oft beide Arten. Die Grenze verschwimmt: Wann handelt es sich um regelbasierte Programmierung und wann ist es wirklich Künstliche Intelligenz? Technologien wie RPA (Robotic Process Automation) werden seit Jahrzehnten in Unternehmen verwendet. Sie pauschal als KI zu klassifizieren, führt oft zu Verwirrung. Gleichzeitig hilft es zu verstehen, dass die Grundlogik der Prozessautomatisierung auch im KI-Zeitalter weiterhin gilt: Auch KI-Lösungen führen letztlich Abfolgen von Schritten und Entscheidungen aus. Der Unterschied liegt darin, wie deterministisch oder wie flexibel diese Entscheidungen getroffen werden. Je komplexer ein Problem wird, desto stärker bewegt sich die Lösung in ein probabilistisches, also nicht mehr vollständig vorherbestimmbares Verhalten.

In der einfachsten Form handelt es sich um klassische Workflow-Automatisierung. Hier werden klar definierte, regelbasierte Abläufe umgesetzt: Wenn ein bestimmtes Ereignis eintritt, folgt eine festgelegte Aktion. Diese Logik ist vollständig deterministisch und eignet sich besonders für standardisierte, wiederkehrende Prozesse.

Darauf aufbauend entstehen Workflows, die gezielt durch KI-Elemente ergänzt werden. Der Ablauf bleibt grundsätzlich strukturiert, einzelne Schritte werden jedoch durch KI unterstützt – etwa bei der Klassifikation von Anfragen, der Analyse von Texten oder der Generierung von Vorschlägen. Die KI erweitert den Prozess, ersetzt ihn aber nicht vollständig.

Die nächste Stufe ist agentische KI. Hier werden nicht mehr nur einzelne Schritte unterstützt, sondern ganze Aufgabenbereiche eigenständig bearbeitet. Die KI trifft innerhalb eines definierten Rahmens eigene Entscheidungen, passt ihr Vorgehen an und führt mehrere Schritte flexibel aus. Der Prozess ist nicht mehr vollständig vorgegeben, sondern entwickelt sich dynamisch entlang des Ziels. Der entscheidende Unterschied zwischen diesen

drei Formen liegt somit weniger in der Technologie selbst, sondern im Grad der Autonomie und Flexibilität der Lösung.

Die ersten Sprachmodelle standen in der Kritik. Ihr revolutionärer und unternehmerischer Nutzen wurde angezweifelt. Das Halluzinieren war noch zu groß. Ein LLM konnte ein Gedicht zum 85. Geburtstag dichten, aber unternehmerische Fragestellungen schienen weit weg. Halluzinationen hängen mit der probabilistischen Vorhersage zusammen: Modelle wurden darauf trainiert, das nächste wahrscheinlichste Wort vorherzusagen, nicht darauf, Fakten zu prüfen.[13] Plausibilität und Korrektheit sind zwei unterschiedliche Ziele. Anfangs kannten die Modelle nur das erste.

Dazu kommt die Limitierung durch Trainingsdaten. Anfangs wurden nur historische Daten verwendet, es entsteht ein Zeitversatz. Bei ChatGPT lag dieser anfangs über zwei Jahre. Alle Informationen, die in dieser Zeit entstanden, waren dem Modell unbekannt. Sprachmodelle können nur »wissen«, was in ihren Trainingsdaten war. Für alles, was fehlt, zum Beispiel persönliche oder Unternehmensinformationen, Nischenwissen, sehr aktuelle Ereignisse, lokale Details hat das Modell zwei Optionen: »Ich weiß es nicht« sagen oder improvisieren. Und strukturell neigt es eher zum Improvisieren.

So entstehen die bekannten Fälle, in denen LLMs Gerichtsurteile erfinden, die nie existiert haben: überzeugend formuliert, mit korrekter Syntax, plausiblen Namen, realistischen Aktenzeichen. Wichtig ist: Es lügt nicht. Es tut, wofür es optimiert wurde. Halluzinationen zu begrenzen war eines der zentralen Entwicklungsziele der letzten Jahre, und große Schritte wurden gemacht.

Interessanterweise betrachten manche das Probabilistische, also auch Improvisation, nicht nur als Schwäche, sondern als Stärke. Bei kreativen Arbeiten oder kniffligen Problemen improvisieren Menschen ebenfalls. Richtig eingesetzt, wird KIs Hang zur Improvisation zur Stärke. Je nach Aufgabenstellung kann der Freiheitsgrad, also der Grad der geforderten Richtigkeit, bewusst gesteuert

werden. Weniger Strenge führt oft zu ungewöhnlicheren, kreativeren Lösungen. So gilt es zu entscheiden, wie viele Freiheitsgrade eine KI erhalten soll: Beim Austausch eines Ersatzteils in einem Umspannwerk möchten wir keine Kreativität. In der Kundenkommunikation schon eher. Für eine erste Ideensammlung zu neuen Geschäftsideen ist viel Kreativität sogar hilfreich.

Als Führungskraft wird von uns verlangt, Entscheidungen zu treffen. Entscheidungen auf Basis falscher oder unvollständiger Informationen erzeugen Schaden, manchmal klein, etwa für die eigene Reputation, manchmal groß, zum Beispiel durch falsche Abbuchungen. Die Führungsaufgabe dahinter ist simpel und zugleich anspruchsvoll: Wir müssen nicht nur entscheiden**,** wo wir der KI in unserer eigenen Arbeit vertrauen**,** sondern auch, wo wir probabilistischen Systemen in unseren Geschäftsprozessen vertrauen können. Wir müssen ein Gefühl dafür entwickeln, wie sich Halluzinationen dort begrenzen lassen, wo es notwendig ist**,** durch klare Quellen, Prüfmechanismen und Eskalation. Was zunächst wie eine klassische IT-Aufgabe erscheint, ist in Wirklichkeit keine reine IT-Aufgabe, sondern vor allem eine Führungsaufgabe. Zwar ist KI eine Technologie, ihr erfolgreicher Einsatz hängt jedoch vor allem von Strategie, Prozessen, Daten und Menschen ab. Führen bedeutet, das Potenzial größerer Automatisierung durch KI zu erkennen und gleichzeitig die Risiken zu steuern, die entstehen, wenn ein System überzeugend klingt, aber irrt.

Nicht jedes Team braucht perfekte KI. Jedes Team braucht klare Regeln, wie mit Unsicherheit umzugehen ist.

1.4 Hype oder Substanz? Welche Fragen im KI-Markt weiterhelfen

Wir kennen inzwischen die Technologie und wissen, warum KI Wahrscheinlichkeiten liefert statt Wahrheiten. Jetzt richten wir den Blick nach außen. Denn Führung bedeutet auch, das Kräftefeld zu

kennen, in dem Entscheidungen getroffen werden. Marktanteile und Rankings veralten in diesem Markt innerhalb von Monaten. Was nicht veraltet, sind die Kräfte, die dahinter wirken. Um die geht es hier in diesem Kapitel.

Im Grunde wird der KI-Markt momentan von zwei Arten von Anbietern dominiert. Es stehen sich zwei grundverschiedene Spielertypen gegenüber. Auf der einen Seite die Frontier Labs. Unternehmen wie OpenAI oder Anthropic, die ausschließlich KI-Modelle entwickeln. Sie treiben die technologische Spitze, ihr gesamtes Geschäft steht und fällt mit der KI. Auf der anderen Seite die Plattformkonzerne, wie Microsoft, Google, Meta, SAP und andere. Sie verdienen ihr Geld längst mit etablierten Produkten und bauen KI dort ein, wo Millionen Menschen ohnehin täglich arbeiten, in E-Mails, WhatsApp, Besprechungen, Geschäftsprozessen.

Die Labs haben die besten Modelle, aber keinen direkten Zugang zu Unternehmenskunden. Die Plattformen haben diesen Zugang, kaufen aber die Modelle häufig ein. Für Führungskräfte hat das eine praktische Konsequenz: Wer das Werkzeug eines Plattformkonzerns einführt, kauft Bequemlichkeit und Integration und bindet sich zugleich enger an dessen Ökosystem. Und weil fast alle täglich mit denselben Programmen arbeiten, gewinnt am Ende selten das beste Modell, sondern das am tiefsten integrierte. Jede Tool-Entscheidung ist damit auch eine stille Wette darauf, welche Plattform sich durchsetzt. Neben der Differenzierung zwischen den beiden großen Gruppen der Frontier Labs und der Plattform Player kommen nun viele KI-Anbieter in den Markt mit ganz unterschiedlichen Wertschöpfungstiefen. Viele KI-Produkte basieren im Hintergrund auf den gleichen Modellen. Es lohnt sich daher ein Blick hinter die Kulissen, um eine gute Entscheidung zu treffen.

Der KI-Markt wirkt unübersichtlich. Tausende Produkte, jede Woche neue Anbieter, jedes Werkzeug bewirbt sich »jetzt mit KI«. Dabei ist es wichtig zu erkennen, was wirklich eine neue Innovation ist. Unter der Haube arbeitet meist eines der wenigen großen

Modelle aus den Frontier Labs. Die wenigsten Anbieter entwickeln nämlich eigene KI. Sie bauen eine Hülle um ein eingekauftes Modell, etwa eine ansprechende Oberfläche mit vorformulierten Anweisungen oder eine Anbindung an bestimmte Daten. In der Branche nennt man das einen Wrapper. Also eine Verpackung um einen Motor, den jeder andere Anbieter genauso einkaufen kann.

In jede Anbieterprüfung gehören darum drei Fragen. Welches Modell arbeitet im Kern, und was geschieht dort mit unseren Daten? Worin besteht der Mehrwert jenseits des Modells? Und was bleibt davon übrig, wenn die nächste KI-Modellgeneration diese Fähigkeit von Haus aus mitbringt? Solche Verschiebungen sind in der KI-Welt keine Ausnahme, sondern die Regel. Funktionen, für die heute ein Spezialanbieter bezahlt wird, stecken morgen im Basismodell. Wer das weiß, verhandelt kürzere Laufzeiten und schaut sehr genau hin, wofür er eigentlich bezahlt.

Damit kommen wir zu einer der meistdiskutierten Fragen rund um KI: Handelt es sich um die nächste Blase? Die Frage ist jedoch zu pauschal, um sie mit einem einfachen Ja oder Nein zu beantworten. Man muss zwei Dinge auseinanderhalten. Die Kategorie KI ist real. Hinter ihr stehen zahlende Unternehmenskunden, echte Umsätze, und im Gegensatz zu früheren Hypes, wie Blockchain oder Metaverse, ist KI an kein Produkt und keine Branche gebunden. Sie lässt sich in fast jede Tätigkeit einbetten, in der mit Sprache, Dokumenten, Zahlen oder Entscheidungen gearbeitet wird. Damit revolutioniert sie vor allem die Wissensarbeit.

Die einzelne Investition ist damit noch lange nicht gerechtfertigt. Die Summen, die derzeit in Infrastruktur fließen, übersteigen die heutigen Erträge um ein Vielfaches, und in vielen Organisationen klafft eine Lücke zwischen Nutzung und Wirkung. Ein beeindruckender Chatbot verändert noch keinen Prozess, eine ausgerollte Lizenz schafft noch keine Produktivität. Diese Lücke gilt es noch zu schließen.

Dann bleibt die lauteste Frage der Branche: Wann erreicht KI den Menschen? Das wird als die sogenannte AGI bezeichnet, Artificial General Intelligence. Bemerkenswert daran ist, dass sich nicht einmal die Menschen einig sind, die diese Systeme bauen. Einige Industrieführer erklären AGI für erreicht oder greifbar nah, renommierte Forscher halten den heutigen Ansatz für grundsätzlich ungeeignet, echte Intelligenz hervorzubringen. Wenn die Schöpfer keine gemeinsame Definition haben, taugt der Begriff nicht als Planungsgrundlage.

Gefährlich ist die Frage aus einem anderen Grund. Sie wirkt wie strategische Weitsicht, funktioniert aber wie eine Vertagung. »Bis das kommt, haben wir noch Zeit.« So kann aus informierter Vorsicht ein Grund werden, heute nicht zu handeln.

Wer Signale beobachten will, findet konkretere: Agenten – KI-Systeme, die nicht mehr nur antworten, sondern selbstständig mehrschrittige Aufgaben erledigen, Dateien bearbeiten, Vorgänge anstoßen. Je mehr ein System handeln darf, desto dringender werden Berechtigungen, Kontrolle und Nachvollziehbarkeit – das ist bereits heute eine Governance-Aufgabe. Und Multimodalität – KI versteht zunehmend Dokumente, Bilder und Ton im Zusammenhang und rückt damit in Prozesse vor, die bisher als nicht automatisierbar galten.

Beides sind keine Zukunftsspekulationen, sondern Entwicklungen mit Ansage. Und beide führen zum selben Schluss: Der Utility Gap schließt sich nicht durch bessere Modelle. Ob KI im Unternehmen etwas verändert, entscheidet sich an Prioritäten, Prozessen, Daten, Kompetenzen, Vertrauen und Verantwortung – also an Führung. Wie daraus Mehrwert entsteht, ist das Thema des nächsten Kapitels.