



Skript Modul 1

Einführung in die KI-Arbeit

IHK-WEITERBILDUNG

KI-Tool Expert*in (IHK)





INHALTSVERZEICHNIS

Einführung in die KI-Arbeit im Arbeitsalltag	3
Anleitung zur Vertiefungsaufgabe A.....	39
Anleitung zur Vertiefungsaufgabe B.....	43



BUSINESS ACADEMY RUHR

Weiterbildung

KI-Tool Expert*in (IHK)



IHK-Weiterbildung

EINFÜHRUNG IN DIE KI- ARBEIT IM ARBEITSALLTAG

EINFÜHRUNG IN DIE KI-ARBEIT IM ARBEITSALLTAG

Willkommen im ersten inhaltlichen Modul eurer Weiterbildung KI-Tool Expert (IHK). Bevor wir gemeinsam loslegen, eine kurze Einordnung dessen, was euch in dieser Lernwoche erwartet – und was bewusst noch nicht. Dieses Modul ist das **Grundlagenmodul** des Kurses. Es legt die Basis, auf der alle weiteren fünf Module aufbauen. Das bedeutet: Wir steigen hier noch nicht in die konkrete Anwendung ein. Ihr werdet in diesem Modul keine Texte mit KI verfassen, keine Präsentationen bauen, keine Bilder generieren und keine Datenanalysen durchführen. Das alles kommt – und zwar Modul für Modul. Worum es hier geht, ist etwas Grundsätzlicheres: Ihr sollt KI verstehen, einordnen, einschätzen und mit der richtigen Haltung in die Folgemodule starten können. Unser Versprechen an euch: Wenn ihr dieses Modul abgeschlossen habt, ist KI für euch keine geheimnisvolle Black Box mehr. Ihr versteht, was unter der Haube passiert, ihr kennt die wichtigsten Fachbegriffe, ihr könnt Tools fundiert auswählen, ihr seid rechtlich auf der sicheren Seite, und ihr könnt eure ersten Prompts strukturiert formulieren. Damit habt ihr alles, was es braucht, um in ab Modul 2 selbstbewusst in die konkrete Anwendung einzusteigen.



WIE SICH DIE ARBEITSWELT VERÄNDERT

Bevor wir uns Definitionen, Tools und Techniken anschauen, lohnt sich ein Schritt zurück. KI ist nicht einfach eine neue Software, die ihr installiert und benutzt. KI verändert die Art, wie **Wissensarbeit** funktioniert – und das spürt fast jeder Mensch, der in den letzten Jahren beruflich am Schreibtisch gesessen hat. Manchmal als spannende Möglichkeit, manchmal als diffuse Sorge, oft als beides gleichzeitig. Dieses erste Kapitel hat deshalb keinen Tool-Charakter. Es geht darum, dass ihr euch selbst in dieser Veränderung verortet, bevor wir tiefer einsteigen. Schauen wir kurz auf die letzten Jahre. Bis Ende 2022 war KI für die meisten Menschen ein Hintergrundthema. Es gab Sprachassistent*innen auf dem Smartphone, automatische Übersetzungen im Browser, Empfehlungsalgorithmen bei Streamingdiensten. Aber KI war selten etwas, mit dem ihr aktiv gearbeitet habt – sondern etwas, das im Hintergrund mitlief. Mit der Veröffentlichung von **ChatGPT im November**

2022 hat sich das fast über Nacht geändert. Plötzlich konnte jeder mit einem Browser ein leistungsfähiges KI-System direkt ansprechen, auf Deutsch, kostenlos, ohne Programmierkenntnisse. Innerhalb von zwei Monaten hatte ChatGPT über hundert Millionen Nutzer*innen – schneller als jede Anwendung zuvor in der Geschichte des Internets. Für viele Berufstätige war das ein Aha-Moment. Manche probierten es ein paar Mal aus und legten es zur Seite. Andere haben seitdem nicht mehr aufgehört, KI in ihren Alltag einzubauen. Wieder andere, das ist der häufigste Fall, schwanken irgendwo dazwischen: Sie nutzen KI sporadisch, haben aber das Gefühl, längst nicht alles auszuschöpfen, was möglich wäre. Was diesen Moment so besonders gemacht hat, war nicht nur die Technik. Es war die Tatsache, dass eine sehr **breite Gruppe von Menschen** plötzlich erlebt hat, dass eine Maschine ihnen bei Aufgaben helfen konnte, die bis dahin als zutiefst menschlich galten: Texte verfassen, argumentieren, übersetzen, Ideen entwickeln, komplexe Sachverhalte erklären. Seitdem läuft eine Welle, die nicht abreißt. Ständig neue Tools, ständig neue Funktionen, ständig neue Berichte darüber, was angeblich bald alles möglich sein wird.

Diese Welle erzeugt zwei **Reaktionen**, die ihr vielleicht selbst kennt. Die eine ist Begeisterung: Endlich ein Werkzeug, das mir lästige Aufgaben abnimmt und mir mehr Zeit für das Wesentliche gibt. Die andere ist Verunsicherung: Wenn KI bald all das kann, was ich beruflich tue – wo bleibe ich dann? Beide Reaktionen sind nachvollziehbar, und ehrlich gesagt: Beide haben einen wahren Kern. Genau deshalb lohnt es sich, einen genaueren Blick darauf zu werfen, was tatsächlich passiert.

Drei Muster, wie KI Berufe verändert

Wenn man sich anschaut, wie KI heute schon konkret in Unternehmen wirkt, lassen sich drei Muster beobachten. Sie sind hilfreich, weil sie das diffuse Gefühl von Veränderung in greifbare Kategorien sortieren.

Das erste und mit Abstand häufigste Muster ist die **Verschiebung von Aufgaben** innerhalb eines Berufs. Eine Marketingfachkraft schreibt nicht weniger Texte als früher – aber der Anteil des aktiven Schreibens sinkt, weil KI Entwürfe liefert. Dafür steigt der Anteil von Briefing, Redaktion und Qualitätskontrolle. Eine Sachbearbeiterin im Kund*innenservice schreibt nicht weniger E-Mails – aber sie schreibt sie schneller mit KI-Unterstützung und hat mehr Zeit für komplexe Fälle, die Empathie und Erfahrung brauchen. Ein Steuerberater liest nicht weniger Verträge – aber er nutzt KI, um die Standardklauseln zu erfassen, und konzentriert sich auf die kritischen Stellen. In allen drei Fällen bleibt der Beruf, aber die innere Verteilung der Aufgaben verändert sich. Die Routinearbeit schrumpft, die anspruchsvolle Arbeit gewinnt an Gewicht. Wer diese Verschiebung aktiv mitgestaltet, entwickelt sich beruflich weiter, ohne den Beruf zu wechseln.

Das zweite Muster ist die **Entstehung neuer Rollen**. In den letzten Jahren sind Berufsbilder und Teilrollen entstanden, die es vorher schlicht nicht gab. Prompt Engineers, KI-Trainer*innen, KI-Beauftragte in Unternehmen, KI-Ethikberater*innen. Manche davon sind hochspezialisierte Vollzeitstellen in großen Unternehmen, andere sind Teilrollen, die in bestehende Aufgaben hineinwachsen. Wenn ihr in eurem Unternehmen die Person seid, die nach diesem Kurs am meisten Ahnung von KI hat, werdet ihr früher oder später solche Teilaufgaben mitübernehmen – ob das offiziell heißt oder nicht. Das ist eine echte

Karrierechance, gerade in Unternehmen, in denen das Thema noch nicht systematisch aufgegriffen wurde.

Das dritte Muster ist das **Verschwinden bestimmter Routinetätigkeiten** – und in manchen Fällen ganzer Berufsfelder. Hier wird es etwas unbequemer, und es lohnt sich, ehrlich hinzuschauen. Es gibt Tätigkeiten, die durch KI praktisch verschwinden – etwa das händische Erfassen einfacher Rechnungen, das Transkribieren von Interviews oder das einfache Übersetzen standardisierter Dokumente. Bei einzelnen Routinetätigkeiten entstehen meist neue Aufgaben, die mit Qualitätskontrolle, Steuerung und Optimierung der KI zu tun haben. Was wirklich vollständig wegfällt, ist auf der Ebene einzelner Tätigkeiten erstaunlich wenig. Auf der Ebene ganzer Berufsfelder sieht das Bild aber differenzierter aus. Es gibt Berufe, in denen sich heute schon abzeichnet, dass die Nachfrage deutlich sinken wird. Klassische Übersetzungsbüros für Standardtexte, einfache Texterstellungs-Agenturen, Datenerfassungsdienste, bestimmte Bereiche im Telefon-Support, Teile der einfachen Programmierarbeit. Wer in diesen Feldern arbeitet, sollte das ernst nehmen – nicht panisch, aber realistisch. Die Veränderungen passieren nicht über Nacht, aber sie sind bereits messbar. An dieser Stelle hilft ein Blick zurück.

Es ist nicht das erste Mal, dass eine Technologie ganze Berufe verändert oder verschwinden lässt. Mit der **industriellen Revolution im 19. Jahrhundert** gingen Berufe wie der Webstuhlhandwerker, der Lampenanzünder, der Eisbringer oder der Telegraphist nach und nach verloren. Der berühmte Aufstand der Ludditen in England richtete sich gerade dagegen: Textilarbeiter*innen zerstörten mechanische Webstühle, weil sie ihre Lebensgrundlage bedroht sahen. Sie hatten in dem Moment recht – ihre konkreten Berufe gab es bald nicht mehr. Sie hatten aber im Großen unrecht: Die Industrialisierung hat über die Jahrzehnte mehr Beschäftigung geschaffen, als sie zerstört hat. Sie hat allerdings auch zu Verwerfungen geführt, die manche Generationen schwer getroffen haben. Wer im richtigen Moment in den richtigen neuen Beruf gewechselt ist, hat profitiert. Wer am alten festgehalten hat, hatte es schwer. Mit KI sehen wir gerade etwas Ähnliches – mit einigen wichtigen Unterschieden. Der erste Unterschied liegt darin, welche Arbeit betroffen ist. Die industrielle Revolution hat vor allem körperliche und mechanische Arbeit verändert. KI verändert zunächst **geistige und kommunikative Arbeit**, also genau die Wissensarbeit, die in den letzten Jahrzehnten als besonders sicher galt. Das trifft eine andere Bevölkerungsgruppe als damals und kommt für viele Beschäftigte überraschend, weil sie sich auf der vermeintlich sicheren Seite des Arbeitsmarkts wähnten. Der zweite Unterschied: KI bleibt nicht beim Geistigen. In Verbindung mit Robotik dringt sie zunehmend auch in **körperliche und mechanische Tätigkeiten** vor. Humanoide Roboter, autonome Fahrzeuge, intelligente Maschinen in Produktion und Logistik, Pflegeroboter – all das ist nicht mehr nur Science-Fiction, sondern wird in den nächsten Jahren in immer mehr Bereichen praktische Realität. Damit erweitert sich die Welle der Veränderung von den Schreibtischen in die Werkhallen, auf die Straßen, in die Lager und Krankenhäuser. Berufe wie LKW-Fahrer*in, Lagerarbeiter*in, Reinigungskraft, Pfleger*in oder Mitarbeiter*in in der Fertigung werden sich ebenfalls verändern – manche stark, manche weniger, aber niemand bleibt unberührt. Insofern ist die heutige Veränderung umfassender als die industrielle Revolution: Sie betrifft sowohl den Kopf als auch die Hand. Der dritte Unterschied liegt im **Tempo**. Während die industrielle Revolution sich über Generationen erstreckte, vollziehen sich KI-getriebene Veränderungen in Zeiträumen von Jahren, manchmal Monaten. Eine neue

Modellgeneration kann innerhalb weniger Wochen Aufgaben lösen, die vorher als unlösbar galten. Das gibt Beschäftigten weniger Zeit, sich anzupassen, und das macht systematische Weiterbildung wichtiger denn je.

Die unbequeme Seite ist also: Die Veränderung ist breiter, schneller und betrifft mehr Menschen gleichzeitig als frühere technologische Umbrüche. Die gute Seite: Genau wie damals entstehen auch heute **neue Berufe, neue Branchen und neue Bedarfe**, von denen wir heute teilweise noch keine klare Vorstellung haben. Die Wirtschaft als Ganzes hat in solchen Übergängen historisch betrachtet nicht weniger Beschäftigung gehabt, sondern eher mehr. Die Frage ist also weniger, ob es genug Arbeit geben wird, sondern wer die Anpassung schafft und wer zurückbleibt.

Was KI gut kann – und was nicht

Damit ihr nicht in die Falle übertriebener Erwartungen oder unbegründeter Ängste tappt, lohnt sich eine grobe Einordnung dessen, was KI heute wirklich leistet. KI ist hervorragend in allem, was mit **Sprache, Mustererkennung und Rekombination vorhandenen Wissens** zu tun hat. Sie kann Texte zusammenfassen, übersetzen, formulieren, umformulieren, strukturieren. Sie erkennt Muster in großen Datenmengen, generiert Bilder zu Beschreibungen, transkribiert Audio, beantwortet Fragen zu Inhalten, die ihr ihr gebt. In vielen dieser Bereiche ist sie auf einem Niveau, das vor wenigen Jahren als unerreichbar galt. Schwächer ist sie überall dort, wo es um **echtes Verstehen, langfristige Verantwortung, körperliche Präsenz, situatives Urteilsvermögen und vertrauensvolle Beziehungen** geht. Eine KI kann eine medizinische Anamnese formulieren, aber nicht den Patienten beruhigen, wenn er weint. Eine KI kann einen Vertragsentwurf liefern, aber nicht entscheiden, ob ihr diesem Geschäftspartner wirklich vertrauen wollt. Eine KI kann eine Verhandlungsstrategie vorschlagen, aber nicht das schwierige Gespräch mit der Mitarbeiterin führen, deren Leistung nachgelassen hat. Genau in diesen Schnittstellen entsteht der Wert eurer Arbeit in den nächsten Jahren. Nicht im Konkurrieren mit der KI, sondern im **Kombinieren mit der KI**. Ihre Stärken nutzen, ihre Schwächen kennen, die eigenen menschlichen Qualitäten gezielt dort einsetzen, wo sie wirklich zählen. Wer diese Haltung verinnerlicht, sieht KI nicht als Bedrohung, sondern als Verstärker der eigenen Wirksamkeit.

Warum diese Weiterbildung – und warum gerade jetzt?

Vielleicht habt ihr euch bei der Anmeldung gefragt, warum ihr ausgerechnet jetzt einen Kurs zum Thema KI-Tools macht. Drei Gründe sprechen dafür, und sie haben alle mit Timing zu tun. Erstens befindet sich die Technik gerade an einem interessanten Punkt. Die **großen Sprachmodelle** sind reif genug, um im Alltag zuverlässig zu funktionieren, aber noch nicht so etabliert, dass alle anderen schon selbstverständlich damit arbeiten. Wer jetzt einsteigt und systematisch Kompetenz aufbaut, ist denen voraus, die später nachziehen müssen. Der Vorsprung, den ihr euch in den nächsten Wochen erarbeitet, lässt sich später nur mit deutlich mehr Aufwand wieder aufholen. Zweitens entstehen gerade die **Standards und Routinen für KI im Berufsalltag**. In zwei oder drei Jahren werden viele Unternehmen feste Prozesse, Tools und Richtlinien haben. Heute habt ihr die Chance, diese Standards in eurem Umfeld mitzuprägen, statt sie später nur zu übernehmen. Das ist gerade in mittelständischen Unternehmen ein interessanter Hebel, weil dort die Strukturen noch flexibler sind als in Großkonzernen. Drittens ist die

rechtliche Grundlage in Deutschland und Europa gerade in Bewegung. Die DSGVO bleibt der zentrale Datenschutzrahmen, der EU AI Act tritt schrittweise in Kraft, Urheberrechtsfragen werden vor Gerichten geklärt. Wer hier auf dem Laufenden ist, schützt sich und das eigene Unternehmen vor Risiken, die andere unterschätzen. Wir werden diese rechtliche Seite in diesem Modul ausführlich behandeln, weil sie die Voraussetzung für jede ernsthafte KI-Nutzung im beruflichen Kontext ist.

Genau diese drei Ebenen behandeln wir in diesem Kurs: die Technik, damit ihr versteht, was passiert. Die Anwendung, damit ihr produktiv arbeiten könnt. Und das rechtliche Umfeld, damit ihr sicher unterwegs seid. Am Ende habt ihr nicht nur ein IHK-Zertifikat, sondern eine eigene KI-Nutzungsplanung – also ein konkretes Dokument, das euren persönlichen Weg in der KI-Welt beschreibt und das ihr direkt im Arbeitsalltag nutzen könnt.

Übung: Diese Übung ist freiwillig und dient eurer eigenen Reflexion – ihr müsst sie nicht einreichen oder vorzeigen. Nehmt euch ein paar Minuten Zeit und denkt in Ruhe nach.

Schaut auf einen typischen Arbeitstag der letzten Woche. Welche Tätigkeiten haben einen großen Teil eurer Zeit eingenommen? Sortiert sie gedanklich in zwei Kategorien: solche, die eher Routine sind, und solche, die wirklich euer Wissen, eure Erfahrung und euer Urteilsvermögen erfordert haben. Überlegt anschließend, was sich an diesem Tag verändern würde, wenn euch ein guter KI-Assistent für die Routineanteile zur Verfügung stünde – und was ihr mit der gewonnenen Zeit anfangen würdet.

KÜNSTLICHE INTELLIGENZ: BEGRIFFE UND FUNKTIONSWEISEN

Im ersten Kapitel ging es um die große Linie: Wie verändert sich Arbeit durch KI, wo stehen wir, was bedeutet das für euch persönlich. Damit ihr in dieser Veränderung kompetent agieren könnt, braucht ihr aber auch ein **Grundverständnis** dessen, womit ihr da eigentlich arbeitet. Genau das ist das Ziel dieses Kapitels. Wir gehen einmal sauber durch die Begriffe und Funktionsweisen, die euch im KI-Alltag immer wieder begegnen – damit ihr nicht nur Tools bedient, sondern wirklich versteht, was unter der Haube passiert.

Die KI-Familie: Von schwacher KI bis generativer KI

Der Begriff „Künstliche Intelligenz“ ist riesig. Wenn jemand davon spricht, kann fast alles gemeint sein – vom Spamfilter in eurem Postfach bis zum Sprachmodell, das ein Gedicht schreibt. Damit ihr Diskussionen einordnen könnt, lohnt sich ein kurzer Überblick über die wichtigsten Ebenen dieser Familie. Ganz oben steht der **Oberbegriff Künstliche Intelligenz**. Damit meint man jede Form von Software, die Aufgaben löst, für die wir normalerweise menschliche Intelligenz voraussetzen würden. Das ist absichtlich weit gefasst, weil sich das Feld ständig verändert. Innerhalb der KI gibt es eine wichtige Unterscheidung zwischen schwacher KI und starker KI. **Schwache KI** ist alles, was wir heute haben: hochspezialisierte Systeme, die in einem klar umrissenen Bereich beeindruckend gut sind, aber außerhalb davon ratlos. ChatGPT, Bilderkennung, Schachcomputer, Übersetzungssoftware – all das ist schwache KI. **Starke KI** wäre eine Maschine mit menschenähnlichem Bewusstsein und allgemeiner Intelligenz, also einer KI, die selbstständig denkt, fühlt und in beliebigen Kontexten handelt. Die gibt es heute noch

nicht, und es ist umstritten, ob es sie je geben wird. Wenn ihr in Medien von „AGI“ oder „Artificial General Intelligence“ lest – das ist die laufende Debatte um genau diese Frage. Für eure praktische Arbeit reicht es, zu wissen: Alles, womit ihr in diesem Kurs zu tun habt, ist schwache KI.

Innerhalb der schwachen KI ist **maschinelles Lernen** der wichtigste Bereich. Dabei lernt ein Programm aus Daten, statt von Programmierer*innen explizit für jede Situation Regeln eingegeben zu bekommen. Wenn euer E-Mail-Programm gelernt hat, was Spam ist, dann hat es das aus tausenden Beispielen gelernt – nicht aus einer Liste fester Regeln. Eine besondere Form des maschinellen Lernens ist **Deep Learning**. Hier kommen sogenannte neuronale Netze zum Einsatz, die in vielen Schichten aufgebaut sind und sehr komplexe Muster erkennen können. Deep Learning hat die KI-Welt in den letzten zehn Jahren revolutioniert, weil sich damit Aufgaben lösen ließen, an denen klassische Methoden gescheitert waren – etwa das Erkennen von Gesichtern auf Fotos oder das Verstehen gesprochener Sprache. Und schließlich, ganz konkret für euren Alltag relevant: **generative KI**. Das ist die Form von KI, die Inhalte selbst erzeugen kann – Texte, Bilder, Audio, Video, Code. Sie basiert auf Deep Learning und ist mit Tools wie ChatGPT, Claude, Midjourney oder Sora der Bereich, der seit 2022 die größte Aufmerksamkeit bekommen hat. Generative KI ist das, womit ihr in diesem Kurs zu rund 95 Prozent arbeiten werdet. Wenn wir im weiteren Verlauf einfach von „KI“ sprechen, ist meist generative KI gemeint.

Wie ein Sprachmodell denkt: Tokens, Wahrscheinlichkeiten, Kontextfenster

Schauen wir jetzt etwas genauer auf das System, mit dem ihr am meisten zu tun haben werdet: das **Large Language Model**, kurz LLM. Das ist ein riesiges KI-Modell, das auf das Verarbeiten von Sprache spezialisiert ist und die Grundlage für ChatGPT, Claude, Gemini und viele andere bildet. Den Begriff „Large“ hat es nicht ohne Grund: Diese Modelle haben hunderte Milliarden Parameter und wurden mit einer kaum vorstellbaren Menge an Text trainiert.

Damit ihr versteht, wie ein LLM funktioniert, ist ein Begriff zentral: **Token**. Ein Token ist die kleinste Einheit, mit der ein Sprachmodell rechnet. Es ist meistens ein Wort oder ein Wortteil. Das deutsche Wort „Hund“ ist in etwa ein Token, das Wort „Hundefutter“ sind oft schon zwei oder drei Tokens. Auch Satzzeichen und Leerzeichen zählen mit. Wenn ihr eine Anfrage an ChatGPT stellt, zerlegt das System euren Text zunächst in Tokens und arbeitet dann auf dieser Ebene. Das Spannende ist, was danach passiert. Ein Sprachmodell macht im Grunde immer dasselbe: Es schaut sich die bisherigen Tokens an und berechnet, welches Token mit welcher **Wahrscheinlichkeit** als nächstes kommt. Wenn ihr fragt „Was ist die Hauptstadt von Frankreich?“, dann hat das Modell gelernt, dass nach diesem Satz mit sehr hoher Wahrscheinlichkeit das Token „Paris“ folgt. Es spuckt also Paris aus. Wenn ihr fragt „Schreib mir ein Gedicht über Herbst“, gibt es viele plausible nächste Tokens, und das Modell wählt eines aus, dann das nächste, dann das nächste, bis ein vollständiges Gedicht entstanden ist. Das klingt einfacher, als es ist, denn dahinter steckt eine ungeheuer komplexe Maschinerie. Aber das Grundprinzip ist genau das: Token für Token, jedes Mal die wahrscheinlichste oder eine plausible Fortsetzung. Genau hier liegt übrigens der Grund, warum KI manchmal halluziniert – aber dazu kommen wir im nächsten Kapitel.

Ein weiterer wichtiger Begriff ist das **Kontextfenster**. Das ist die Menge an Tokens, die ein Modell bei einer Anfrage gleichzeitig im Blick haben kann. Ihr könnt es euch wie einen Arbeitsspeicher vorstellen: alles, was darin Platz hat, kann das Modell beim Antworten berücksichtigen. Alles, was außerhalb liegt, ist für die aktuelle Antwort unsichtbar. Frühere Modelle hatten Kontextfenster von wenigen tausend Tokens, was etwa zehn DIN-A4-Seiten entspricht. Aktuelle Modelle haben Kontextfenster von hunderttausend Tokens und mehr, also mehrere ganze Bücher. Das ist praxisrelevant, weil ihr damit ganze Verträge, Studien oder Datensammlungen in einer Anfrage verarbeiten lassen könnt. Wenn ein Tool sich plötzlich „nichts mehr merkt“, liegt das oft daran, dass das Kontextfenster überschritten wurde. Ein Detail, das oft Verwirrung stiftet: Sprachmodelle haben kein Langzeitgedächtnis, sondern nur dieses Kontextfenster. Wenn ihr morgen ein neues Gespräch beginnt, hat das Modell vergessen, was ihr gestern besprochen habt – es sei denn, das Tool speichert Gespräche aktiv und legt sie automatisch in den Kontext (manche Tools haben dafür eine eigene Funktion namens „Memory“ oder „Erinnerungen“). Diese Eigenheit ist wichtig zu wissen, weil sie erklärt, warum ihr bei längeren Projekten oft denselben Kontext immer wieder mitliefern müsst.

Wie ein Modell entsteht: Trainingsdaten, Pre-Training und Fine-Tuning

Ein Sprachmodell fällt nicht vom Himmel. Damit es funktioniert, muss es trainiert werden – und dieses **Training** läuft in mehreren Phasen, die ihr verstehen solltet, weil sie viele spätere Phänomene erklären.

Die erste und größte Phase ist das **Pre-Training**. Hier wird das Modell mit gewaltigen Mengen an Text gefüttert: Webseiten, Bücher, Artikel, Foren, Wikipedia, Code, oft mehrere hundert Milliarden Wörter. In dieser Phase lernt das Modell die Struktur von Sprache, die Beziehungen zwischen Wörtern, grammatikalische Regeln, Faktenwissen und auch implizite Muster wie Stil, Tonfall und Argumentationsformen. Pre-Training dauert Wochen oder Monate, läuft auf riesigen Rechenclustern und kostet Millionenbeträge. Das ist auch der Grund, warum es weltweit nur eine Handvoll Unternehmen gibt, die wirklich konkurrenzfähige große Modelle von Grund auf trainieren können. Wichtig zu verstehen: Beim Pre-Training „lernt“ das Modell kein Wissen im engeren Sinne, sondern statistische Zusammenhänge. Es weiß nicht, dass Paris die Hauptstadt von Frankreich ist – es hat nur sehr oft gelesen, dass nach „Hauptstadt von Frankreich“ das Wort „Paris“ folgt. Diese Unterscheidung wird wichtig, sobald wir über Halluzinationen sprechen.

Die Frage, womit ein Modell genau trainiert wird, ist eine der spannendsten und gleichzeitig eine der heikelsten in der ganzen KI-Welt. Vereinfacht gesagt: Ja, ein Großteil der **Trainingsdaten** kommt aus dem Internet – aber nicht aus „dem Internet“ in einem unscharfen Sinn, sondern aus konkreten, sehr großen Datensätzen, die Anbieter wie OpenAI, Anthropic, Google oder Meta zusammenstellen. Die wichtigste Quelle ist meist **Common Crawl**, ein gemeinnütziges Projekt, das seit Jahren systematisch Webseiten archiviert und hunderte Milliarden Seiten umfasst. Daraus werden je nach Anbieter unterschiedliche Auszüge verwendet, oft gefiltert und qualitätsgeprüft. Dazu kommen weitere Quellen: Wikipedia in vielen Sprachen, große Sammlungen von Büchern, wissenschaftliche Artikel, Code aus Plattformen wie GitHub, Foren wie Reddit oder Stack Overflow, journalistische Inhalte, juristische Texte, Patente, Untertitel von Videos und Transkripte von Podcasts. Manche Anbieter lizenzieren auch gezielt bestimmte Inhalte ein

– etwa OpenAI mit Lizenzverträgen mit großen Verlagen wie Axel Springer, Associated Press oder der Financial Times. Die schlichte, aber wichtige Erkenntnis: Wenn euer eigener Text irgendwann frei im Netz stand, ist die Wahrscheinlichkeit hoch, dass er Teil eines Trainingsdatensatzes geworden ist. Damit kommen wir zur juristisch heikelsten Frage der gesamten KI-Welt, die bis heute nicht abschließend geklärt ist: Sehr viel **urheberrechtlich geschütztes Material** ist in den Trainingsdaten der großen Modelle gelandet – Bücher, Artikel, Songtexte, Code, Bilder. Die KI-Anbieter argumentieren, dass dieses Training unter „Fair Use“ beziehungsweise im europäischen Recht unter die Schranken für Text- und Data-Mining fällt. Viele Rechteinhaber sehen das anders, und in den USA, Großbritannien und Deutschland laufen aktuell mehrere Klagen – etwa von der New York Times gegen OpenAI, von Buchautor*innen-Verbänden, von Bildagenturen wie Getty Images oder von Musikverlagen. Erste Urteile gibt es, eine einheitliche Linie aber noch nicht. Wir vertiefen das im rechtlichen Kapitel.

Nach dem Pre-Training folgt eine zweite Phase, das **Fine-Tuning**. Hier wird das Modell auf bestimmte Verhaltensweisen feinjustiert: hilfsbereit antworten, höflich sein, gefährliche Inhalte vermeiden, Anweisungen befolgen. Beim Fine-Tuning werden oft Beispiele genutzt, in denen Menschen vorgeben, wie eine gute Antwort aussehen soll, und das Modell lernt, diesem Stil zu folgen. Eine besondere Variante davon heißt **Reinforcement Learning from Human Feedback**, kurz RLHF. Dabei bewerten Menschen verschiedene Antwortvarianten, und das Modell lernt aus diesem Feedback, welche Antworten als besser gelten. Dieses Verfahren hat ChatGPT seinerzeit den Durchbruch ermöglicht – nicht weil es ein größeres Modell war, sondern weil es durch RLHF deutlich besser auf menschliche Erwartungen abgestimmt war. Was das praktisch bedeutet: Wenn ihr verschiedene Modelle vergleicht und eines „angenehmer“ oder „präziser“ wirkt, liegt das oft nicht an der reinen Größe des Modells, sondern am Fine-Tuning. Anbieter wie OpenAI, Anthropic oder Google investieren enorm viel Arbeit in diese Feinjustierung, und das ist ein wesentlicher Teil dessen, was die Tools voneinander unterscheidet.

Multimodalität: Wenn KI nicht nur Text versteht

Lange Zeit waren KI-Modelle auf eine einzige Art von Input spezialisiert: Texte oder Bilder oder Audio. Multimodalität bedeutet, dass ein Modell **mehrere dieser Modalitäten gleichzeitig** verarbeiten kann. Moderne Modelle wie GPT-4o, Claude 4 oder Gemini sind multimodal: Ihr könnt ihnen ein Foto schicken, es beschreiben lassen, eine Frage dazu stellen, eine Tabelle als Bild hochladen und sie analysieren lassen. Im Alltag heißt das zum Beispiel, dass ihr ein Foto eines handschriftlichen Notizzettels machen und das Modell bitten könnt, daraus eine saubere digitale Notiz zu erstellen. Oder ihr schickt ein Foto eines defekten Geräts und fragt nach möglichen Ursachen. Oder ihr ladet ein PDF eines Vertrags hoch und lasst zentrale Klauseln zusammenfassen. Multimodalität ist einer der Bereiche, in denen sich die Tools rasant weiterentwickeln. Es lohnt sich, regelmäßig zu prüfen, was euer Tool aktuell kann – manchmal sind Funktionen über Nacht dazugekommen, ohne dass ihr es mitbekommen habt.

Reasoning-Modelle: Wenn die KI laut nachdenkt

Eine relativ neue Entwicklung sind sogenannte **Reasoning-Modelle**. Klassische Sprachmodelle antworten direkt: Frage rein, Antwort raus. Reasoning-Modelle arbeiten anders. Sie nehmen sich Zeit zum „Nachdenken“, indem sie intern Schritt für Schritt durch

ein Problem gehen, bevor sie eine Antwort formulieren. Das funktioniert besonders gut bei Aufgaben, die mehrere Denkschritte erfordern – mathematische Probleme, logische Rätsel, komplexe Analysen, mehrstufige Planungsaufgaben. In der Praxis erkennt ihr Reasoning-Modelle daran, dass sie **länger** brauchen, manchmal zehn oder zwanzig Sekunden statt einer Sekunde. Manche Tools zeigen euch dabei sogar an, was das Modell gerade „denkt“. Bekannte Beispiele sind die o-Reihe von OpenAI, der „extended thinking“-Modus von Claude oder Geminis Deep-Think-Funktion. Für einfache Anfragen sind diese Modelle Overkill und liefern nicht zwingend bessere Ergebnisse als klassische Modelle. Für komplexe Probleme, in denen mehrere Schritte sauber durchgegangen werden müssen, sind sie aber ein echter Fortschritt. Praktisch heißt das: Wenn ihr euch das nächste Mal fragt, warum manche Anfragen plötzlich länger brauchen oder warum es bei manchen Tools die Wahl zwischen einem „normalen“ und einem „Thinking“-Modus gibt – jetzt wisst ihr es. Nutzt Reasoning-Modelle gezielt für anspruchsvolle Aufgaben, nicht für Smalltalk.

Agents: Wenn KI selbstständig Aufgaben übernimmt

Der neueste Trend, von dem ihr in den nächsten Jahren noch viel hören werdet, sind Agents. Ein Agent ist eine KI, die nicht nur Antworten generiert, sondern selbstständig Aufgaben ausführt. Während ihr ChatGPT bisher fragt und eine Antwort bekommt, könnt ihr einem Agent eine Aufgabe geben – etwa „buche mir den günstigsten Flug nach München für nächsten Mittwoch“ – und der Agent arbeitet die einzelnen Schritte selbstständig ab: Webseiten aufrufen, Preise vergleichen, Optionen prüfen, Buchung vornehmen, Bestätigung schicken. Damit das funktioniert, braucht ein Agent mehrere Bausteine: ein Sprachmodell als „Gehirn“, die Fähigkeit, Werkzeuge zu nutzen (Browser, E-Mail, Datenbanken), ein Gedächtnis, um den Stand der Aufgabe zu behalten, und eine Steuerung, die zwischen den Schritten plant. Aktuell sind Agents an vielen Stellen schon im Einsatz – als Recherche-Agents, die selbstständig Informationen aus mehreren Quellen zusammentragen, als Daten-Agents, die Tabellen analysieren und aufbereiten, als Buchungs-Agents im Reisebereich, als Coding-Agents, die Programmierarbeit übernehmen. Für euch ist wichtig zu wissen, dass Agents zwar enorm viel Potenzial haben, aber auch neue **Risiken** mit sich bringen. Ein Agent, der selbstständig handelt, kann auch Fehler machen, die nicht so leicht zu korrigieren sind wie eine falsche Textantwort. Wenn er aus Versehen die falsche E-Mail verschickt oder eine Buchung mit dem falschen Datum vornimmt, ist der Schaden real. Deshalb ist beim Einsatz von Agents besondere Sorgfalt nötig, und die meisten seriösen Tools bauen Sicherheitsmechanismen ein, etwa Rückfragen vor wichtigen Aktionen.

Was generative KI nicht ist

Bevor wir dieses Kapitel abschließen, ein wichtiger Realitätscheck. Bei aller Beeindruckung, die KI auslösen kann, ist es sinnvoll, euch ein paar Dinge bewusst zu machen, die generative KI nicht ist.

Generative KI **versteht nicht im menschlichen Sinne**. Sie hat kein Bewusstsein, keine Absicht, keine Gefühle. Wenn ein Chatbot „freundlich“ wirkt, ist das ein Effekt seines Trainings, kein Ausdruck echter Zuwendung. Diese Einsicht ist nicht trivial, weil viele Menschen anfangen, eine emotionale Beziehung zu Chatbots aufzubauen – das passiert leicht, ist aber psychologisch nicht ganz harmlos.

Generative KI ist **keine Faktendatenbank**. Sie speichert nicht, dass Paris die Hauptstadt von Frankreich ist, sondern hat statistische Zusammenhänge gelernt. Das ist ein wichtiger Unterschied, weil es erklärt, warum sie sich manchmal irrt, ohne es zu merken.

Generative KI **rechnet nicht zuverlässig**. Sprachmodelle sind keine Taschenrechner. Bei einfachen Rechnungen kommen sie meist auf das richtige Ergebnis, bei komplexeren Berechnungen können sie unauffällig falsch liegen. Viele moderne Tools schalten deshalb für Rechenaufgaben einen echten Taschenrechner zu (Stichwort „Tool Use“), aber verlasst euch nie blind auf KI-Mathematik ohne Plausibilitätsprüfung.

Und schließlich: Generative KI hat **keinen Zugriff auf die Echtzeitwelt**, es sei denn, das Tool ist explizit dafür ausgerüstet. Ein klassisches Sprachmodell hat einen Wissensstichtag – meist einige Monate vor seiner Veröffentlichung – und weiß nichts über Ereignisse danach. Tools mit Web-Suche oder Browsing-Funktion sind eine Ausnahme, aber das müsst ihr aktiv prüfen, bevor ihr nach aktuellen Themen fragt.

(...)