

Ulrich Engelke | Barbara Engelke

ChatGPT & Co. in der **Schule**

Modern unterrichten mit KI

Effizient vorbereiten,
lehren, prüfen und
korrigieren



Hinweis des Verlages zum Urheberrecht und Digitalen Rechtemanagement (DRM)

Liebe Leserinnen und Leser,

dieses E-Book, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Mit dem Kauf räumen wir Ihnen das Recht ein, die Inhalte im Rahmen des geltenden Urheberrechts zu nutzen. Jede Verwertung außerhalb dieser Grenzen ist ohne unsere Zustimmung unzulässig und strafbar. Das gilt besonders für Vervielfältigungen, Übersetzungen sowie Einspeicherung und Verarbeitung in elektronischen Systemen.

Je nachdem wo Sie Ihr E-Book gekauft haben, kann dieser Shop das E-Book vor Missbrauch durch ein digitales Rechtemanagement schützen. Häufig erfolgt dies in Form eines nicht sichtbaren digitalen Wasserzeichens, das dann individuell pro Nutzer signiert ist. Angaben zu diesem DRM finden Sie auf den Seiten der jeweiligen Anbieter.

Beim Kauf des E-Books in unserem Verlagsshop ist Ihr E-Book DRM-frei.

Viele Grüße und viel Spaß beim Lesen

Ihr mitp-Verlagsteam



Barbara Engelke
Ulrich Engelke

ChatGPT & Co. in der Schule

Modern unterrichten mit KI

Effizient vorbereiten, lehren, prüfen und korrigieren



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://portal.dnb.de/opac.htm> abrufbar.

ISBN 978-3-7475-0878-7

1. Auflage 2025

www.mitp.de

E-Mail: mitp-verlag@lila-logistik.com

Telefon: +49 7953 / 7189 - 079

Telefax: +49 7953 / 7189 - 082

© 2025 mitp Verlags GmbH & Co. KG, Augustinusstr. 9a, DE 50226 Frechen

Dieses Werk, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Lektorat: Janina Vervost

Sprachkorrektorat: Christine Hoffmeister

Covergestaltung: Christian Kalkert

Bildnachweis: © yatcenko / stock.adobe.com

Satz: Petra Kleinwegen

Inhaltsverzeichnis

	Einleitung	11
	Was Sie in diesem Buch erwartet	12
	Über die Autoren	13
1	KI, Machine Learning & ChatGPT	15
1.1	Wie funktionieren KI-Chatbots?	15
1.1.1	Überblick über KI und maschinelles Lernen	15
1.1.2	Die Dartmouth-Konferenz und führende Köpfe der KI	16
1.1.3	Kurzer Abriss der KI-Entwicklung bis heute	18
1.1.4	Machine Learning und Deep Learning	20
1.1.5	Sam Altman, OpenAI und ChatGPT	24
1.1.6	Und wie funktionieren KI-Chatbots nun wirklich?	26
1.2	Datenschutz, Urheberrecht und Einschränkungen	32
1.2.1	Datenschutz und KI-Bots	32
1.2.2	Urheberrecht und LLMs	34
1.2.3	Einschränkungen von KI-Chatbots	38
1.2.4	Halluzinieren von KI-Sprachmodellen	45
1.2.5	Sind KI-Suchmaschinen die Zukunft?	48
1.3	Erste Schritte mit ChatGPT	51
1.3.1	Als neuer Nutzer registrieren	51
1.3.2	Erste Schritte des Promptings	52
1.3.3	Kostenlose oder kostenpflichtige Nutzung?	55
2	Richtig prompten – Ein Prozess	57
2.1	Simple-Prompts und einfache Erweiterungen	57
2.2	Die wichtigsten zehn Prompting-Parameter	59
2.2.1	Aufgabe, Limitierung und Ausgabeform	60
2.2.2	Das Ziel	61
2.2.3	Der Tonfall	62
2.2.4	Die Rolle	63
2.2.5	Das Publikum	64
2.2.6	Die Terminologie	65
2.2.7	Schlüsselwörter	67
2.2.8	Die Sprache	69
2.2.9	Top 10 der Prompting-Parameter im Überblick	73

2.3	Weitere Prompting-Parameter	73
2.3.1	Kontext	73
2.3.2	Statistik	75
2.3.3	Standpunkte	75
2.3.4	Begründungen	76
2.3.5	Gegenargumente	78
2.3.6	Analogien	79
2.3.7	Expertenurteile	80
2.3.8	Zitate	83
2.3.9	Beispieltex te und Prompt-Rückschluss-Technik	85
2.3.10	Sensible Inhalte	88
2.4	Allgemeine Regeln für bessere Prompts	89
2.5	Was sind Power-Prompts?	92
2.6	Basis-Prompts setzen Rahmen für folgende Anfragen	94
2.7	Rollenspiele mit der KI: »Du bist jetzt mein ...«	95
2.8	Temperatur der Ausgabe steuern	98
2.9	Iteratives Prompting, Feedback-Schleifen und Optimierungen	101
2.10	Prompts von ChatGPT verbessern lassen	103
3	ChatGPT als Assistent des Lehrenden	111
3.1	Unterrichtsvorbereitung	111
3.1.1	Jahresplanung und Strukturierung	111
3.1.2	Fächerübergreifende Projekte	113
3.1.3	Binnendifferenzierung	115
3.1.4	Vereinfachen komplizierter Regeln	119
3.1.5	Vereinfachen von Texten – Hilfen für förderungsbedürftige Schüler oder bei schwierigen Inhalten (inklusive r Ansatz) .	121
3.1.6	Eilige Unterrichtsvorbereitung	123
3.1.7	Größere Methodenvielfalt	125
3.2	Herausforderungen außerhalb des Unterrichts	128
3.2.1	Artikel für den Jahresbericht der Schule	128
3.2.2	Zwischenmenschliche Probleme	129
3.2.3	Tipps für Disziplinprobleme und Elterngespräche	131
3.2.4	Vorbereitung von Studienfahrten (inkl. Elterninformation)	135
3.3	ChatGPT im Unterricht	139
3.3.1	Kognitive Schulung: Erörtern und Debattieren	141
3.3.2	Die Arbeit mit Bildern	142
3.3.3	Einführung in ChatGPT	143
3.4	Prüfungsvorbereitung	145
3.4.1	Allgemeine Prüfungsvorbereitung	145
3.4.2	Prüfungsvorbereitung mit Custom GPTs	147

3.5	Korrigieren mit KI	149
3.5.1	PEER	150
3.5.2	DeepL-Write-Schreibassistent	151
3.5.3	ChatGPT 4o	152
3.6	Weitere digitale Tools für Lehrkräfte	154
4	Einsatz von KI für verschiedene Schularten und Fächer	157
4.1	Grundschule	158
4.1.1	Unterrichtsvorbereitung	158
4.1.2	Unterstützung im Unterricht	162
4.1.3	Nachbereitung (Korrekturen und Prüfungen)	165
4.1.4	Pädagogische Arbeit an den Grundschulen	173
4.2	Mittelschule (ehemals Hauptschule)	179
4.2.1	Unterrichtsvorbereitung	180
4.2.2	Unterstützung im Unterricht	182
4.2.3	Nachbereitung, Korrekturen und Prüfungen	185
4.2.4	Nebenfächer an der Mittelschule	187
4.2.5	Pädagogische Arbeit an den Mittelschulen	192
4.3	Realschulen und Gymnasien	193
4.3.1	Unterrichtsvorbereitung	194
4.3.2	Unterstützung im Unterricht	199
4.3.3	Nachbereitung, Korrekturen und Prüfungen	202
4.3.4	Weitere Fächer an Realschulen und Gymnasien	206
4.4	Pädagogische Arbeit an weiterführenden Schulen	225
4.4.1	Ethnische Heterogenität und Rassismus	227
4.4.2	Bildungsferne Stammfamilien	228
4.5	Verfassungsviertelstunde in Bayern	230
4.5.1	Grundschule	230
4.5.2	Mittelschule	231
4.5.3	Realschule	231
4.5.4	Gymnasium	232
5	Unterrichtsziel: KI-Kompetenz	233
5.1	Wie Jugendliche die Relevanz von KI für sich und die Schule einschätzen	233
5.2	Warum der Umgang mit ChatGPT und anderen KI-Technologien entscheidend ist	234
5.2.1	Die Zukunft (und Gegenwart) des Arbeitsmarkts	234
5.2.2	Fähigkeit zum kritischen Urteil	235
5.2.3	Verbesserung der Lernprozesse	237
5.2.4	Lebenslanges Lernen	237

5.3	Die 4 Ks der Bildung im 21. Jahrhundert	238
5.4	Reformpädagogik und KI – das geht	246
5.4.1	Erfahrungsorientiertes Lernen	246
5.4.2	Neue Rolle der Lehrkraft	247
5.4.3	Kooperatives Lernen	247
5.4.4	Offene Unterrichtsformen	248
6	Datenschutzkonform unterrichten	255
6.1	Die Unsicherheit der Lehrkräfte im Umgang mit KI	255
6.2	Grundprinzipien des Datenschutzes und die DSGVO	256
6.2.1	Was sind personenbezogene Daten?	256
6.2.2	Lehrkräfte und Datenschutz	257
6.2.3	Die sechs Grundprinzipien der DSGVO	257
6.3	Warum Datenschutz im Unterricht mit KI meist gar kein Problem ist	259
6.4	KI-Tools und Plattformen für Lehrkräfte mit Datenschutzversprechen im Faktencheck	259
6.4.1	KI-Tools für Lehrkräfte im Überblick	260
6.4.2	Datenschutzkonformität von KI-Tools für Lehrkräfte	263
6.4.3	ChatGPT und Datenschutz – Welche Version für Lehrkräfte?	264
6.5	Vorschlag für einen DSGVO-konformen KI-Einsatz in Schulen ...	265
6.5.1	Vermeidung der Eingabe personenbezogener Daten	265
6.5.2	Team-, Enterprise oder Edu-Version von ChatGPT mit korrekter Einstellung	266
6.5.3	Nutzung einer lokalen KI ohne Internetverbindung	266
6.6	Konkrete Handreichung mit Checkliste für Schulen	267
6.7	Quelle: datenschutz-schule.info	269
6.8	Regeln für künstliche Intelligenz – der EU AI Act und die KI-Verordnung	270
7	Herausforderungen für den Bildungssektor	273
7.1	Betrug durch KI?	273
7.1.1	Selbstverständlicher Umgang mit KI-Modellen im Schulalltag	274
7.1.2	Prüfungsleistungen aus dem häuslichen Umfeld	275
7.1.3	Wie zitiert man die Ergebnisse von KI-Textgeneratoren? ..	275
7.2	Generative KI im W-Seminar	276
7.2.1	Assistenz der Lehrkräfte	276
7.2.2	ChatGPT als Schülerwerkzeug für die W-Seminararbeit ...	278

7.3	Prüfen mit Sprachgeneratoren – ein Blick in die Zukunft	284
7.3.1	Zukunftsformat 1: Text im Team	284
7.3.2	Zukunftsformat 2: Gamifiziertes Prüfen	286
	Sei nett zu deiner KI	289
	Stichwortverzeichnis	293

Einleitung

Wer zu spät kommt, den bestraft das Leben! Dieser Satz, der Michail Gorbatschow zugeschrieben wird, klingt wie eine historische Mahnung. Tatsächlich stammt das Zitat aber gar nicht von ihm, sondern von seinem außenpolitischen Sprecher *Genadi Gerassimow*, der es 1989 im Zusammenhang mit einem Treffen Gorbatschows mit DDR-Staatschef Honecker prägte. Trotz dieser Verwechslung hat das Zitat bis heute eine starke Aussagekraft und trifft unabhängig vom tatsächlichen Urheber den Kern unseres Anliegens: In einer Welt, die sich technologisch mit Lichtgeschwindigkeit entwickelt, können wir es uns nicht leisten, auf der Stelle zu treten – schon gar nicht im Bildungsbereich.

Während draußen die digitale Revolution tobt und Künstliche Intelligenz unseren Alltag neu definiert, verharren viele Schulen noch im analogen Zeitalter. Kreide und Tafel statt interaktivem E-Screen und die Möglichkeiten der KI bleiben meist ungenutzt. Diese Diskrepanz zwischen technologischer Entwicklung und schulischer Realität ist nicht nur ein technisches, sondern letztlich auch ein pädagogisches Defizit. Bildung steht am Scheideweg: Auf der einen Seite eröffnet Künstliche Intelligenz ungeahnte Möglichkeiten, den Unterricht zu bereichern, Lernprozesse zu individualisieren und Schüler auf eine digital geprägte Zukunft vorzubereiten. Auf der anderen Seite stehen Ängste, Unsicherheiten und oft auch mangelndes Wissen über den sinnvollen Einsatz dieser Technologien im Klassenzimmer.

Aber was bedeutet es für unsere Schülerinnen und Schüler, wenn wir diese Chancen ungenutzt lassen? In einer Welt, in der KI bereits in vielen Lebensbereichen Einzug gehalten hat – von personalisierten Empfehlungen in Streamingdiensten über Spracherkennung bis hin zu selbstfahrenden Autos –, ist es unsere Aufgabe als Lehrkräfte, jungen Menschen die Kompetenzen zu vermitteln, die sie in dieser neuen Realität benötigen. Das neue Zauberwort heißt **KI-Kompetenz**, die ich als Lehrkraft heute brauche, um sie morgen zum Gegenstand und Lernziel meines Unterrichts machen zu können.

Es geht nicht (nur) darum, technisches Wissen zu vermitteln, sondern auch darum, kritisches Denken zu fördern, ethische Fragen zu diskutieren und Medienkompetenz zu stärken. Wenn wir in der Schule an veralteten Methoden festhalten, laufen wir Gefahr, eine ganze Generation nicht ausreichend auf die Anforderungen der Zukunft vorzubereiten. Das Leben, d.h. die gesellschaftliche Entwicklung, kennt keinen Stillstand, und wer zu spät handelt, läuft Gefahr, den Anschluss dauerhaft

zu verlieren. Wir haben jetzt die Chance, den Bildungsprozess neu zu gestalten: interaktiver, individueller und näher an der Lebenswelt unserer Schülerinnen und Schüler. Nutzen wir diese Chance, um gemeinsam die Bildung von morgen aktiv zu gestalten.

Dieses Buch richtet sich an Lehrende, die bereit sind, neue Wege zu gehen und sich mit den Möglichkeiten von generativer KI und speziell ChatGPT im Unterricht auseinanderzusetzen. Dabei soll es keine Technikgläubigkeit propagieren, sondern praxisnahe Anleitungen und Reflexionen bieten, das Potenzial von KI im Unterricht zu entdecken und aktiv zu nutzen. Lassen Sie uns diese Möglichkeiten gemeinsam erforschen.

Was Sie in diesem Buch erwartet

Dieses Buch bietet Ihnen eine praxisorientierte Einführung in den Einsatz von ChatGPT im Unterricht. Es richtet sich sowohl an Lehrkräfte aller Schulformen und Fächer als auch an Entscheidungsträger, die mehr über das Potenzial und die Einsatzmöglichkeiten von KI im schulischen Kontext erfahren möchten. Vorkenntnisse sind hierbei nicht erforderlich.

In den **ersten beiden Kapiteln** erhalten Sie wichtige Hintergrundinformationen zur generellen Funktionsweise generativer KIs wie ChatGPT sowie Strategien, um zielführende Anfragen (Prompts) zu formulieren.

Das Herzstück des Buchs bilden die praktischen Kapitel 3 **und** 4, in denen Sie zahlreiche nützliche Prompts sowie Anregungen für den Einsatz von ChatGPT im Schulalltag finden – von der Unterrichtsplanung über die Verwendung als integrativer Bestandteil des Unterrichts bis hin zur Prüfungsvorbereitung und -korrektur. Der gezielte Einsatz von KI hilft Ihnen dabei, alltägliche Aufgaben effizienter zu erledigen und den Unterricht abwechslungsreicher sowie individueller zu gestalten. Zudem erhalten Sie konkrete Vorschläge für die Anwendung in unterschiedlichen Schulformen und Fächern.

Kapitel 5 betrachtet das Thema KI im Unterricht aus einer weiteren Perspektive und thematisiert KI-Kompetenz als Lernziel.

Die abschließenden Kapitel 6 **bis** 8 beschäftigen sich mit den Herausforderungen und dem Datenschutz im Kontext generativer KI. Hierbei stehen Fragen wie der Umgang mit Betrugsversuchen mit KI sowie die Vereinbarkeit von generativer KI mit der Datenschutzgrundverordnung (DSGVO) im Vordergrund. Sie werden dabei sehen, dass der Einsatz von KI für Lehrkräfte unter Berücksichtigung einfacher Richtlinien problemlos möglich ist.

Dieses Buch vermittelt Ihnen das notwendige Wissen, um ChatGPT effektiv im Unterricht einzusetzen und das Potenzial der KI voll auszuschöpfen.

Über die Autoren

Barbara Engelke hat Germanistik und Theologie für das Lehramt an Gymnasien studiert und arbeitet als Lehrkraft sowie Fachschaftsleitung Deutsch an einem bayerischen Gymnasium. Im Vordergrund ihrer pädagogischen Tätigkeit steht, junge Menschen an neue Herausforderungen heranzuführen und sie für die Zukunft mit den nötigen Kompetenzen sowie einer kritisch-unabhängigen Haltung auszustatten.

Ulrich Engelke hat das erste Staatsexamen in Germanistik und Anglistik sowie einen Magister in Linguistik. Nach einem kurzen Ausflug in das Verlagswesen und freiberuflicher Tätigkeit als Fachautor hat er eine Internetagentur gegründet. Heute ist er als freier KI-Trainer und IT-Unternehmensberater tätig. Sein besonderes Interesse gilt technologischen Innovationen und deren ökonomischen wie gesellschaftlichen Auswirkungen.

lehren-mit-ki.de

Zusammen bieten Ulrich und Barbara Engelke Fortbildungen für Lehrkräfte und Schulleitungen zu künstlicher Intelligenz in Schule und Unterricht sowie Chat-GPT an.

KI, Machine Learning & ChatGPT

Das nächste große Ding ist da und es passt in jede Hosentasche: Künstliche Intelligenz ist dabei, den Alltag zu erobern und klopft nun auch an den Türen der Klassenzimmer an. Zwischen Schulbüchern, Tafel, Heften und Stundenplänen ist ein neues Thema aufgetaucht, das nicht mehr verschwinden wird. Wie verändert KI im Allgemeinen und ChatGPT im Besonderen die Schule und unseren Unterricht?

Wir zeigen Ihnen, wie diese intelligenten Technologien funktionieren und wie sie nicht nur unseren Alltag, sondern auch das Lernen und Lehren revolutionieren. Zunächst werfen wir mit diesem einführenden Kapitel einen Blick hinter die Kulissen der digitalen Revolution im Bildungsbereich. Statt trockener Theorie erwartet Sie eine anschauliche Reise von den Anfängen der KI bis zu aktuellen Entwicklungen wie ChatGPT, die weit mehr können, als nur einfache Antworten auf simple Fragen zu geben.

1.1 Wie funktionieren KI-Chatbots?

Ein Chatbot alias *Large Language Model* (LLM) ist ein KI-Modell, das auf einer riesigen Menge an Textdaten trainiert wurde, um natürliche Sprache zu verstehen und zu generieren. Es kann Texte verfassen, Fragen beantworten und viele sprachbasierte Aufgaben unterstützen. KI und maschinelles Lernen haben eine Revolution eingeläutet, von der erst spätere Generationen wissen werden, welche tiefgreifenden Veränderungen sie ausgelöst hat. Klar ist aber längst, dass die Veränderungen bahnbrechend sind und dass die Auswirkungen auf unser Leben ebenso vielfältig wie unabsehbar sind.

1.1.1 Überblick über KI und maschinelles Lernen

KI als Thema wirft uns notgedrungen immer wieder auf die Frage zurück, was *Intelligenz* eigentlich ist. Das ist gar nicht so einfach zu beantworten. Auf jeden Fall ist die Frage nach der Intelligenz eine, an der sich unterschiedliche Fächer und Forschungsrichtungen seit vielen Jahren abarbeiten und noch zu keiner gemeinsamen und allgemeingültigen Definition gefunden haben. Intelligenz kann man als ein komplexes und vielschichtiges Konzept beschreiben, das sich auf die Fähigkeit von Individuen bezieht, Informationen zu verarbeiten und darauf basierend zu handeln. Wir können also **lernen** und das Gelernte **sinnvoll anwenden** und immer **weiterentwickeln**, bis wir Dinge tun, die wir so niemals gelernt haben. Das wäre

eine zumindest vorläufige und grundlegende Definition von natürlicher, menschlicher Intelligenz.

Das eigentlich Faszinierende daran ist der *Transfer*, also unsere Fähigkeit, Probleme zu lösen, die teilweise weit über das Gelernte hinausgehen. Dass wir unser Wissen ganz offensichtlich an immer neue Situationen anpassen und dort erfolgreich anwenden, ist ein ganz wichtiger Teil unserer Intelligenz. Um zu verstehen, wie das möglich ist, müssen wir uns die ganze Palette unserer kognitiven Fähigkeiten möglichst detailliert ansehen, einschließlich des **logischen Denkens**, der kreativen **Problemlösungskompetenz**, der gezielten **Wahrnehmungsfähigkeit**, des **Verständnisses von Zusammenhängen**, der Fähigkeit zur **Kommunikation** und des **Lernens aus Erfahrungen**. Aus diesen und anderen Grundfertigkeiten entstehen neue und zunehmend komplexere Fertigkeiten: die Fähigkeit, aus der Interaktion mit der Umwelt zu lernen, anspruchsvolle Konzepte zu erfassen, effektiv zu planen und zu handeln sowie kreativ und innovativ auf immer neue Herausforderungen zu reagieren.

Die Urväter der KI waren von der natürlichen, menschlichen Intelligenz auf jeden Fall so fasziniert, dass sie diese unsere Fähigkeiten auf Maschinen übertragen und Computern das selbstständige Lernen und Denken beibringen wollten. Die Ursprünge der Künstlichen Intelligenz sind stark mit dem Streben verbunden, maschinelle Systeme zu schaffen, die ähnliche kognitive Fähigkeiten wie der Mensch aufweisen. Die Urväter der KI, darunter Wissenschaftler wie Alan Turing, John McCarthy, Marvin Minsky, Allen Newell und Herbert A. Simon, waren von der Komplexität und Vielseitigkeit der menschlichen Intelligenz tief beeindruckt und inspiriert. Sie wollten verstehen, wie Intelligenz funktioniert, und diese Erkenntnisse nutzen, um Maschinen zu entwickeln, die ähnliche Aufgaben ausführen können. Heute ist Künstliche Intelligenz – oder im englischen Original Artificial Intelligence (AI) – der Überbegriff für durch Maschinen erbrachte, menschenähnliche Intelligenzleistungen.

1.1.2 Die Dartmouth-Konferenz und führende Köpfe der KI

Wie immer ist die Entwicklung neuer Technologien eng mit der Leistung ihrer Pioniere und bestimmten Ereignissen verbunden. **Alan Turing** gehört zu den ersten Visionären, der die Frage stellte, ob Maschinen denken können. Der berühmte **Turing-Test** war damals (1950) eher eine Idee als ein Test, der mit den vorhandenen Programmen schon hätte durchgeführt werden können: Wenn eine Maschine in der Lage ist, in einem Gespräch nicht von einem Menschen unterscheidbar zu sein, könnte sie als *intelligent* betrachtet werden und hätte damit den Turing-Test bestanden.

Die **Dartmouth-Konferenz** von 1956 gilt als die Geburtsstunde der KI als eigenständiges Forschungsfeld der Informatik, hier wurde die Grundlage für die zukünftige

Forschung in diesem Bereich gelegt. Eine Vielzahl führender Köpfe kam auf dieser Konferenz zusammen.

John McCarthy ist eine Schlüsselfigur in der Geschichte der Künstlichen Intelligenz, er gilt vielen sogar als der eigentliche Vater der KI. Zumindest hat er schon während der Vorbereitung der Dartmouth Konferenz den Begriff *Künstliche Intelligenz (KI)* geprägt. Außerdem entwickelte er 1958 die Programmiersprache *LISP*, die zur bevorzugten Sprache in der KI-Forschung wurde und es in einigen Bereichen bis heute ist. McCarthy hat bedeutende Beiträge in mehreren Kernbereichen der KI verfasst, war ein Vorreiter bei der Entwicklung von Theorien und Modellen, die es Computern ermöglichen sollen, Alltagswissen zu nutzen, das für Menschen selbstverständlich ist. Er arbeitete sogar bereits an der Konzeptualisierung von autonomen Systemen, einschließlich autonomer Fahrzeuge, und entwickelte Technologien, wie sie heute in selbstfahrenden Autos zum Einsatz kommen.

Marvin Minsky war ein weiterer Pionier und gilt als der einflussreichste Theoretiker im Forscherfeld. Er gründete zusammen mit John McCarthy das *MIT Artificial Intelligence Laboratory*, das zu einem der weltweit führenden Zentren für KI-Forschung wurde. Dieses Labor zog viele talentierte Forscher an und war der Geburtsort zahlreicher bahnbrechender Ideen und Technologien im Bereich der KI. Minsky entwickelte die sogenannte *Frametheorie*, die ein Konzept zur Wissensrepräsentation in KI-Systemen darstellt. Minskys Buch *Society of Mind* ist ein bahnbrechendes Werk, das eine Theorie der menschlichen Intelligenz und des Bewusstseins als Netzwerk aus vielen kleinen und einfachen Prozessen bzw. Agenten entwirft. Diese Idee hat die Forschung in den Bereichen kognitive Wissenschaften und KI beeinflusst, insbesondere im Hinblick auf das Verständnis davon, wie komplexe geistige Prozesse aus simpleren Interaktionen entstehen können. Minsky hat Generationen von Studenten beeinflusst, die später zu führenden Wissenschaftlern und Technikern in der KI und verwandten Feldern wurden.

Nathaniel Rochester arbeitete bei IBM und war einer der Architekten des *IBM 701*, des ersten wissenschaftlichen Computers des Unternehmens. Seine Arbeit trug wesentlich dazu bei, die technische Basis für spätere Entwicklungen in der KI zu ermöglichen. Rochester war auch an der Entwicklung eines der ersten Programme beteiligt, das als KI-Experiment angesehen werden kann. Das Programm wurde auf dem *IBM 704*-Computer implementiert und zielte darauf ab, einfache algebraische Probleme zu lösen.

Claude Shannon revolutionierte das Verständnis von Datenübertragung, -verarbeitung und -speicherung. Seine Grundprinzipien der Informationstheorie sind für viele Aspekte der KI, wie Datenkompression und Fehlerkorrektur in maschinellen Lernsystemen, von großer Bedeutung. Shannon war einer der Ersten, der die Möglichkeiten von Computern zum Spielen von Schach und anderen Spielen erforschte. Seine Arbeiten in den 1950er-Jahren, insbesondere seine Strategien zur Schach-

programmierung, gelten als Pionierleistungen und beeinflussten die Entwicklung von KI-Algorithmen im Bereich der *Spiele* und *Entscheidungstheorien*.

Die führenden Köpfe der Dartmouth-Konferenz und ihre unterschiedlichen Schwerpunkte machen uns begreiflich, wie viele unterschiedliche Forschungsansätze nötig waren und an einem Ort zusammenkommen mussten, damit wir heute die Früchte ihrer visionären Gedanken und ambitionierten Bemühungen ernten können.

1.1.3 Kurzer Abriss der KI-Entwicklung bis heute

Nach der Dartmouth-Konferenz als Geburtsstunde der Künstlichen Intelligenz entwickelte sich die Forschung in mehreren Phasen weiter. Hier sind die wichtigsten Etappen und eine kurze Beschreibung der wesentlichen Entwicklungen von 1956 bis heute:

Frühphase (1956 bis 1970er-Jahre)

Die ersten Jahrzehnte der KI-Forschung konzentrierten sich stark auf symbolische Methoden und Expertensysteme. Diese Systeme basierten auf Logik und regelbasierten Entscheidungsprozessen, die menschliches Wissen speichern und versuchen, komplexe Problemlösungsfähigkeiten zu imitieren. Zu Beginn herrschte großer Optimismus über die Möglichkeiten der KI und die Forscher gingen davon aus, dass bedeutende Fortschritte schnell erreicht werden könnten. Allerdings stießen sie bald auf erhebliche technische und konzeptionelle Herausforderungen, die das Fortschrittstempo verlangsamten.

KI-Winter (1970er- bis 1980er-Jahre)

Der (erste) KI-Winter ist geprägt von allgemeiner Ernüchterung und unerwarteten Finanzierungsproblemen. Aufgrund der enttäuschten Erwartungen und der begrenzten Fortschritte kam es in den 1970er- und 1980er-Jahren zu einem Rückgang des Interesses für die KI-Forschung und einer zunehmend unsicheren Finanzierung. Trotz der allgemeinen Stagnation gab es in einigen spezifischen Bereichen Fortschritte, wie z.B. in der Entwicklung von Algorithmen für das maschinelle Lernen und in der Robotik.

Aufstieg der Maschinen und neuronalen Netze (1980er- bis 2000er-Jahre)

In den 1980er-Jahren erlebte die KI einen erneuten Aufschwung durch den Einsatz von **Expertensystemen** in der Industrie. Expertensysteme zielten darauf ab, das Wissen und die Entscheidungsfähigkeiten menschlicher Experten in einem spezifischen, eng abgegrenzten Bereich nachzuahmen. Diese Systeme zogen deduktive Schlussfolgerungen und trafen Entscheidungen auf Basis einer komplexen Wissensbasis aus Fakten und Regeln. An der Stanford University entstanden mehrere Programme in den Bereichen der medizinischen Diagnostik und chemischen Analyse: **MYCIN** wurde entwickelt, um bakterielle Infektionen zu diagnostizieren

und geeignete Antibiotika-Behandlungen vorzuschlagen. **DENDRAL** half Chemikern bei der Interpretation von Massenspektren, um die Struktur organischer Moleküle zu bestimmen. Damit waren KI-Programme erstmals in der Lage, komplexe wissenschaftliche Probleme zu lösen.

In diese Zeit fällt auch die Wiederentdeckung und Weiterentwicklung **neuronaler Netze**, die für den eigentlichen Durchbruch im maschinellen Lernen verantwortlich sind und zu bedeutsamen praktischen Fortschritten führten. Künstliche neuronale Netze sind Computer-Modelle, die von der Struktur und Funktionsweise des menschlichen Gehirns inspiriert sind. Ein neuronales Netz besteht aus mehreren Schichten von Knoten (Neuronen), die miteinander verbunden sind. Hinzu kamen neue Lernalgorithmen wie das *Backpropagation-Verfahren*, mit denen die Gewichte in einem neuronalen Netz angepasst werden konnten. Diese Gewichte sind die Werte, die die Stärke der Verbindung zwischen einzelnen Neuronen bestimmen. Sie beeinflussen, wie stark ein Eingangssignal weitergeleitet wird und sind entscheidend für das Lernen des Netzes. Das alles ermöglichte jetzt das Training tiefer neuronaler Netzwerke, die komplexe Muster in sehr großen Datensätzen erkennen konnten.

Datengetriebene KI und Deep Learning (2000er-Jahre bis heute)

Die 2000er-Jahre markieren zwei wichtige Entwicklungen, die zu den eigentlichen Katalysatoren der Künstlichen Intelligenz wurden: Mit der rasanten Verbreitung des Internets und dem immer schneller wachsenden World Wide Web bekamen Forscher einfachen Zugang zu riesigen Datenmengen – insbesondere Texten und Bildern. Hinzu kam eine geradezu explosionsartige **Entwicklung der Rechenleistung** von Computern: Große Fortschritte in der Halbleitertechnik, die Entwicklung von Mehrkernprozessoren, immer größere Speicherkapazitäten mit schnelleren Zugriffsgeschwindigkeiten, leistungsstarke Grafikkarten (GPUs) und die skalierbare Rechenleistung des Cloud Computings haben der datengetriebenen KI den entscheidenden Schub versetzt. Komplexe Algorithmen konnten nun auf riesigen Datensätzen mit gigantischer Rechenpower trainiert werden, was zu beeindruckenden Ergebnissen in verschiedenen Anwendungsbereichen führte. Die Nutzung tiefer neuronaler Netze (Deep Learning) hat zu bahnbrechenden Fortschritten in Bereichen wie Bilderkennung, Sprachverarbeitung und Spielen wie AlphaGo geführt.

Gerade in den letzten Jahren hat sich das Tempo nochmals vervielfacht, KI flutet unser Leben wie ein mächtiger Tsunami: Apple, Microsoft, Google und Amazon haben jeweils eigene **Sprachassistenten** entwickelt und über **mobile Endgeräte** in unseren Alltag eingeschleust. Andere KI-Technologien sind mittlerweile in vielen Bereichen des täglichen Lebens präsent, neben Sprachassistenten sind selbstfahrende Autos Realität geworden, medizinische Diagnosen nutzen KI und personalisierte Empfehlungen auf Facebook, YouTube und Spotify sind für uns ganz selbstverständlich geworden.

1.1.4 Machine Learning und Deep Learning

Aber wie wurde das alles, noch dazu in so kurzer Zeit, möglich? Was unterscheidet die KI so grundlegend von anderen Zweigen der Informatik? Welche Methoden und Prinzipien machen den entscheidenden Unterschied? Um das herauszufinden, müssen wir uns mit dem Bereich Machine Learning (ML) befassen. Machine Learning ist ganz allgemein gesprochen ein Teilbereich der Künstlichen Intelligenz, der sich mit der Entwicklung von Algorithmen und statistischen Modellen beschäftigt, die Computersysteme in die Lage versetzen, aus Daten zu lernen und Muster zu erkennen, ohne explizit für genau diese Muster programmiert zu sein.

Machine Learning (oder künstliches Lernen ganz allgemein) wird üblicherweise in drei grundlegende Konzepte bzw. Methoden unterteilt: überwachtes Lernen (Supervised Learning), unbeaufsichtigtes Lernen (Unsupervised Learning) und bestärkendes Lernen (Reinforcement Learning).

Überwachtes Lernen (Supervised Learning)

Überwachtes Lernen liegt immer dann vor, wenn ein Algorithmus aus einem vorbereiteten Trainingsdatensatz lernt, der neben den Eingabedaten auch schon die entsprechenden Ausgabewerte enthält. Der einfachste und am häufigsten zitierte Fall zur Erklärung des überwachten Lernens ist ein Algorithmus, der Bilder von Hunden und Katzen unterscheiden kann. Wir nehmen dazu einen Datensatz mit 100 Bildern, 50 Hundebilder und 50 Bilder von Katzen. Typischerweise würde man jetzt 70 Bilder für das Training verwenden und dem Algorithmus zu jedem Bild die Information geben, um welches der beiden Tiere es sich auf dem gezeigten Bild handelt. Ist das Training abgeschlossen, prüft man mit den restlichen 30 Bildern, wie gut die Unterscheidung in der Praxis schon funktioniert. Ist das Ergebnis noch nicht zufriedenstellend und die Fehlerrate zu hoch, kann der Algorithmus durch zusätzliches Training verbessert werden. Weitere Bilder müssen bereitgestellt und klassifiziert (gelabelt) werden, solche *gelabelten Daten* sind die Voraussetzung für überwachtes Lernen. In unserem Fall müssten mindestens weitere 100 Bilder von Hunden und Katzen einer der beiden Arten zugeordnet werden. Der Aufwand dafür ist zwar überschaubar, aber die Fähigkeiten des so entstandenen Programms sind natürlich auch sehr begrenzt.

Reale Anwendungsbeispiele des überwachten Lernens bringen aber schon ziemlich komplexe Programme mit hohem praktischen Nutzen hervor. Der *Bayes-Filter* ist eine weit verbreitete Methode zur **Klassifikation von E-Mails als Spam oder Nicht-Spam**. Statt Tierbildern legt man dem Algorithmus reale Mails aus dem eigenen Postfach vor und kennzeichnet alle, die wir als Spam einstufen und eigentlich nicht lesen wollen. Daraus lernt der Algorithmus eine individuelle und recht zuverlässige Klassifikation von E-Mails als Spam oder Nicht-Spam. Eine andere KI-Anwendung kann den **Kaufpreis eines Hauses**, basierend auf typischen Merkmalen wie Baujahr, Quadratmeterzahl, Anzahl der Schlafzimmer usw., vor-

hersagen. Überwachtes Lernen ist die Grundlage von **Algorithmen zur Gesichtserkennung** oder um vorherzusagen, ob ein Kunde ein Produkt kauft – basierend auf dem Verhalten anderer, ähnlicher Kunden. Auf der Basis medizinischer Daten kann damit die Wahrscheinlichkeit eines Herzinfarkts für einen gegebenen Patienten berechnet und vorausgesagt werden. Die viel zitierte **Kreditrisikobewertung** für eine Bank ist ebenfalls eine auf überwachtem Lernen basierte Anwendung, um die Wahrscheinlichkeit vorherzusagen, mit der ein Kreditnehmer seine Schulden zurückzahlen wird oder ob die Forderung ausfällt.

Unüberwachtes Lernen (Unsupervised Learning)

Nicht alle Daten können vor dem Training aufbereitet und gelabelt werden: Erstens, weil der Aufwand zu hoch ist und die Datenmenge dadurch immer begrenzt bleibt, und zweitens, weil oft gar noch nicht bekannt ist, welche Merkmale für eine Kennzeichnung überhaupt relevant sind. Unüberwachtes Lernen ist eine Methode des maschinellen Lernens, bei der der Algorithmus auf einen Datensatz ohne vorherige Kennzeichnung der Daten trainiert wird. Statt explizit zu wissen, welche Daten zu welchen (noch unbekannten) Kategorien gehören, versucht der Algorithmus, *Muster und Strukturen in den Daten selbst zu erkennen*. Ein klassisches Beispiel für unüberwachtes Lernen ist das *Clustering*, bei dem ähnliche Datenpunkte in Gruppen oder Clustern zusammengefasst werden.

Stellen Sie sich vor, Sie haben einen Datensatz mit 1.000 Bildern von Tieren, darunter Hunde, Katzen, Vögel und Fische, aber ohne zu wissen, welche Bilder zu welchen Tieren gehören. Ein Clustering-Algorithmus wie *K-Means* könnte verwendet werden, um diese Bilder in Gruppen zu unterteilen, basierend auf ihren Ähnlichkeiten. Der Algorithmus könnte beispielsweise feststellen, dass es vier Hauptgruppen gibt und diese entsprechend kennzeichnen. Wir wissen jedoch nicht im Voraus, dass eine Gruppe *Hunde* und eine andere *Katzen* enthält; das muss durch Interpretation der resultierenden Cluster erfolgen.

Ein weiteres Beispiel für unüberwachtes Lernen ist das Vereinfachen von großen Datensätzen für die Analyse von Kunden in einem Supermarkt, über die sehr viele unterschiedliche Daten gesammelt wurden:

- Alter
- Geschlecht
- Wohnort
- Einkommen
- Einkaufsgewohnheiten (z.B. Häufigkeit der Einkäufe, bevorzugte Wochentage)
- Gekaufte Produkte (z.B. Obst, Gemüse, Fleisch, Milchprodukte)
- Durchschnittlicher Einkaufswert
- Zahlungsmethode (z.B. bar, Kreditkarte)

Das sind sehr viele Informationen und es ist schwierig, einen Überblick zu behalten. Mit Techniken wie der Hauptkomponentenanalyse (PCA) oder *t-SNE* kann man diese vielen Daten auf einige wenige wichtige Merkmale reduzieren, die trotzdem die wichtigsten Unterschiede zwischen den Kunden zeigen. Zum Beispiel könnte man die Daten auf die drei Merkmale *Einkaufsgewohnheiten*, *Alter* und *geografische Lage* reduzieren. Das macht es einfacher, Muster zu erkennen, wie z.B., dass Kunden aus bestimmten Wohngebieten häufiger einkaufen oder dass junge Kunden andere Produkte bevorzugen als ältere Kunden. So kann der Supermarkt seine Marketingstrategien besser anpassen und gezielt auf die Bedürfnisse verschiedener Kundengruppen eingehen.

Die Vorteile des unüberwachten Lernens

Weil Daten vorher nicht mehr manuell aufbereitet werden müssen, können wir beim unüberwachten Lernen auf wesentlich größere (und theoretisch unbegrenzte) Datenmengen zurückgreifen. Unüberwachtes Lernen kann dazu beitragen, neue und unentdeckte Muster oder Gruppen innerhalb der Daten zu identifizieren, um Daten zu bereinigen, zu gruppieren oder zu reduzieren, was die Effizienz und Genauigkeit von Modellen des überwachten Lernens verbessern kann.

Typische Anwendungen des unüberwachten Lernens:

- **Kundensegmentierung:** Einteilung von Kunden in Gruppen für gezieltes Marketing
- **Anomalieerkennung:** Erkennung von Betrug oder Fehlern in Finanztransaktionen oder Netzwerksicherheit
- **Marktforschung:** Entdeckung von Mustern und Trends in großen Datensätzen zur Unterstützung der Produktentwicklung
- **Dokumenten- oder Textklassifikation:** Gruppierung von Dokumenten, basierend auf ihrem Inhalt in Themen oder Kategorien, beispielsweise in Bibliotheken
- **Datenvisualisierung:** Visualisierung komplexer Datensätze in vereinfachter Form, um bessere Geschäftsentscheidungen zu treffen
- **Bildanalyse und Gesichtserkennung:** Erkennung von Mustern in Bilddaten
- **Empfehlungssysteme:** Erstellung personalisierter Empfehlungen für Videobeiträge, Filme und Produkte wie bei YouTube, Netflix und Amazon

Bestärkendes Lernen (Reinforcement Learning)

Bestärkendes Lernen ist die dritte grundlegende Methode des maschinellen Lernens, die durch kontinuierliches Feedback immer genauere Anpassungen und Op-

timierungen ermöglicht. Algorithmen und Maschinen lernen, optimale Entscheidungen zu treffen, um komplexe Aufgaben zu bewältigen. Hierzu lernt ein Agent durch Interaktion mit seiner Umgebung, wie er bestimmte Aufgaben ausführen kann, um eine **Belohnung zu maximieren**. Anders als beim überwachten Lernen, wo der Algorithmus anhand von gelabelten Daten trainiert wird, oder beim unüberwachten Lernen, wo der Algorithmus Muster in unstrukturierten Daten erkennt, basiert das bestärkende Lernen auf dem **Prinzip von Versuch und Irrtum**. Der Agent erhält Feedback in Form von Belohnungen, ausbleibenden Belohnungen oder gar Strafen und passt sein Verhalten entsprechend an. Ein einfaches Beispiel für bestärkendes Lernen ist das Training eines Saugroboters, der lernen soll, wie man durch ein Wohnzimmer navigiert. Der Roboter erhält eine Belohnung, wenn er das Ziel erreicht, und keine Belohnung, wenn er gegen eine Wand stößt. Anfangs wird der Roboter zufällige Bewegungen ausführen, aber mit der Zeit lernt er, welche Aktionen ihn näher zum Ziel der Belohnung bringen und welche nicht.

Roboter mögen keine Schokolade

Wenn Sie sich auch schon Gedanken darüber gemacht haben, womit man einen Staubsauger mit KI-Unterstützung belohnen kann: mit numerischen Belohnungspunkten! Der Roboter erhält Punkte oder einen numerischen Wert als Belohnung, wenn er erfolgreich ein Ziel erreicht, wie z.B. einen bestimmten Bereich vollständig zu reinigen oder effizient von einem Punkt zum anderen zu navigieren, ohne irgendwelche Hindernisse zu berühren. Roboter belohnt man eben anders als Menschen!

Ein anderes Beispiel sind Computerprogramme, die lernen, Spiele wie Schach oder Go zu spielen. Hier wird der Agent (das Programm) durch die Rückmeldungen (Gewinn oder Niederlage) trainiert. Der Algorithmus analysiert verschiedene Spielzüge und deren Ergebnisse, um Strategien zu entwickeln, die seine Gewinnchancen erhöhen.

Weitere typische Anwendungen des bestärkenden Lernens sind sehr vielfältig, viele finden sich in der Robotik. Sie ermöglicht es Robotern in dynamischen und komplexen Umgebungen wie in der Produktion oder bei Rettungseinsätzen, ihre Aufgaben immer effektiver und sicherer ausführen. Sie lernen, sich an neue Situationen anzupassen und Aufgaben effizient zu erledigen. Bestärkendes Lernen hilft bei der Optimierung beliebiger Prozesse und kann z.B. auch die Steuerung von Verkehrsflüssen erheblich verbessern: Der Agent lernt, wie er Ampelschaltungen optimieren kann, um den Verkehrsfluss zu maximieren und Staus zu minimieren.

In der Finanzwelt wird bestärkendes Lernen zur Entwicklung von Handelsalgorithmen eingesetzt werden, die optimale Entscheidungen treffen, um Gewinne zu maximieren und Verluste zu minimieren. Der Algorithmus lernt aus historischen

Daten und passt seine Handelsstrategien, basierend auf den Ergebnissen früherer Entscheidungen, kontinuierlich an. Im Gesundheitswesen kann bestärkendes Lernen Behandlungspläne optimieren, indem es die Wirksamkeit verschiedener Therapien kontinuierlich analysiert.

Deep Learning

Machine Learning (ML) und *Deep Learning* (DL) sind beides Teilbereiche der Künstlichen Intelligenz, unterscheiden sich jedoch in ihrer Komplexität und Anwendungsweise. Deep Learning basiert auf künstlichen neuronalen Netzwerken, insbesondere auf tiefen neuronalen Netzwerken mit sehr vielen Schichten, ähnlich unserem Gehirn. Damit ist DL geeignet, hochkomplexe und abstrakte Muster in sehr großen Datenmengen zu erkennen und geht weit über die Möglichkeiten des Machine Learnings hinaus. Es wird für komplexere Anwendungen wie Bild- und Spracherkennung, automatische Übersetzung, autonomes Fahren und die Generierung von Texten oder Bildern wie bei ChatGPT und DALL·E eingesetzt.

Ein weiterer Unterschied zwischen Machine Learning und Deep Learning liegt in der Komplexität und der zu verarbeitenden Datenmenge, wobei DL sehr große Datenmengen verarbeiten kann, wenn die dafür erforderliche Rechenleistung zur Verfügung steht. Wo beim ML Merkmale teilweise noch manuell ausgewählt werden müssen, ist beim DL der Prozess der Merkmalsextraktion komplett automatisiert, da neuronale Netzwerke selbst relevante Merkmale aus den Rohdaten lernen und anwenden können. DL wird daher bevorzugt für Aufgaben eingesetzt, die hochdimensionale und komplexe Daten erfordern, wie Bilder, Videos und besonders auch Sprache. Praktisch alle Leuchtturmanwendungen der KI sind das Ergebnis von Deep-Learning-Algorithmen, die mit riesigen Datensätzen und einer enormen Rechenleistung erstellt bzw. trainiert wurden.

DL-Modelle wie *Convolutional Neural Networks* (CNNs) werden verwendet, um Bilder zu erkennen und zu klassifizieren, z.B. in der medizinischen Bildgebung zur Erkennung von Tumoren. *Recurrent Neural Networks* (RNNs) und *Transformer-Modelle* (wie die gesamte GPT-Familie) basieren ebenfalls auf DL und werden für Aufgaben der **natürlichen Sprachverarbeitung** (NLP), in der maschinellen Übersetzung, der Textgenerierung und für die Stimmungsanalyse eingesetzt.

1.1.5 Sam Altman, OpenAI und ChatGPT

OpenAI wurde im Dezember 2015 von **Elon Musk**, **Sam Altman** und anderen gegründet. Die Organisation hatte das Ziel, sichere und nützliche Künstliche Intelligenz für alle zu entwickeln und ihre Forschungsergebnisse offen zugänglich zu machen. Die Vision war, dass KI das menschliche Leben verbessern kann, wenn fortschrittliche Technologien zum Wohle der gesamten Menschheit und nicht nur zur Verfolgung kommerzieller Interessen eingesetzt werden. Als unabhängige Forschungsorganisation, die sich auf die Entwicklung von KI-Technologien kon-

zentriert, die transparent und kooperativ mit anderen Forschungseinrichtungen und der Öffentlichkeit geteilt werden, wollte man bei OpenAI die Chancen und Herausforderungen der KI besser verstehen, für alle verständlich machen und im **Interesse der Allgemeinheit** handhaben.

Der Ausstieg von Elon Musk bei OpenAI

Elon Musk, einer der Mitgründer von OpenAI, trat im Jahr 2018 aus dem Vorstand der Organisation zurück. Der Ausstieg wurde offiziell mit potenziellen Interessenskonflikten begründet, da Tesla, das Unternehmen, bei dem Musk CEO ist, zunehmend eigene KI-Forschung und -Entwicklung betrieb, insbesondere im Bereich des autonomen Fahrens. Es wurde befürchtet, dass seine Rolle bei beiden Organisationen zu Interessenskonflikten führen könnte. Über mögliche andere Gründe wie ein persönliches Zerwürfnis zwischen Musk und Altman wurde sehr viel spekuliert, wirklich gesichert ist nichts davon. Trotz seines Rücktritts als Vorstandsmitglied blieb Musk OpenAI weiterhin als bedeutender Unterstützer und Spender verbunden.

Die Entstehung und Entwicklung von ChatGPT bei OpenAI ist das Ergebnis fortlaufender Forschung und Innovation im Bereich der Künstlichen Intelligenz und des Natural Language Processing (NLP). Zentrales Projekt der Organisation war die Entwicklung der **Generative Pre-trained Transformer (GPT)** als leistungsstarke, generative Sprachmodelle. **GPT-1**, das erste Modell, wurde schon 2018 veröffentlicht und zeigte bereits die Fähigkeit, menschenähnlichen Text zu generieren. In Fachkreisen galt das als bedeutender Schritt in der NLP-Forschung und demonstrierte schon damals das Potenzial von vortrainierten Sprachmodellen auf eindrucksvolle Weise. 2019 folgte mit **GPT-2** eine erheblich größere und leistungsfähigere Version, die beeindruckend kohärente und kontextuelle Texte generieren konnte. Aufgrund von Bedenken über den möglichen Missbrauch entschied sich OpenAI zunächst, die vollständige Version nicht zu veröffentlichen.

Der eigentliche Durchbruch und eine stärkere Wahrnehmung in der Öffentlichkeit kam 2020 mit **GPT-3**, das mit 175 Milliarden Parametern einen weiteren immensen Sprung darstellte. Parameter in einem neuronalen Netz sind die Gewichte und Bias-Werte, die während des Trainings angepasst werden, um das Modell zu optimieren und genaue Vorhersagen zu ermöglichen. Dieses Modell konnte sehr detaillierte und nuancierte Texte generieren, was seine Anwendungen in verschiedenen Bereichen erheblich erweiterte. Die Leistungsfähigkeit von GPT-3 ermöglichte die Entwicklung von **ChatGPT**, einer speziell auf Dialoge und interaktive Kommunikation ausgerichteten Anwendung. ChatGPT beeindruckte durch seine Fähigkeit, natürliche und sinnvolle Gespräche zu führen, und fand schnell breite Anwendung in Bereichen wie Kundenservice, Bildung und kreativer Inhaltserstellung.

Die Kommerzialisierung von OpenAI begann im Jahr 2019 und war ein strategischer Schritt, um die langfristige Finanzierung und Entwicklung der KI-Forschung zu sichern. OpenAI kündigte die Gründung von **OpenAI LP** als kapitalisierte Einheit innerhalb der Organisation an. Die Entwicklung fortschrittlicher KI-Technologien, insbesondere großer Modelle wie GPT-3, erfordert immense Rechenressourcen und finanziellen Aufwand. Um diese **Projekte nachhaltig finanzieren zu können**, war es wohl notwendig, zusätzliche Einnahmequellen zu erschließen. Ein weiterer Grund war der Wunsch nach schnellerem Wachstum und Innovation. Durch den Zugang zu externen Investitionen konnte OpenAI seine Forschung und Entwicklung erheblich beschleunigen.

Das Capped-Profit-Modell

OpenAI LP ist eine *begrenzt gewinnorientierte Einheit* innerhalb von OpenAI, die im Jahr 2019 gegründet wurde. Seine Struktur basiert auf einem *Capped-Profit-Modell*, das darauf abzielt, Investoren anzuziehen und gleichzeitig die gemeinnützigen Ziele von OpenAI zu bewahren. Dieses Modell *begrenzt die Gewinne auf das 100-Fache der Investition*, um sicherzustellen, dass die Gewinne nicht übermäßig werden und die Mission der gemeinnützigen Mutterorganisation unterstützt wird. Durch die Kommerzialisierung konnte OpenAI externe Investoren anziehen und Partnerschaften mit führenden Technologieunternehmen eingehen. Ein prominentes Beispiel hierfür ist die *Zusammenarbeit mit Microsoft*, das eine Milliarde Dollar in OpenAI investierte und die Integration von OpenAI-Technologien in seine Cloud-Plattform Azure förderte.

Durch diese Struktur konnte OpenAI eine gewisse Balance zwischen der Sicherstellung ausreichender Finanzierung und der Aufrechterhaltung seiner ethischen und gemeinnützigen Ziele finden. Die Kommerzialisierung ermöglicht es OpenAI, seine Mission in begrenztem Maße weiterzuverfolgen und gleichzeitig die Ressourcen zu haben, um in der sich schnell entwickelnden KI-Landschaft führend zu bleiben.

1.1.6 Und wie funktionieren KI-Chatbots nun wirklich?

Diese eingangs gestellte Frage haben wir noch gar nicht beantwortet und ganz genau scheinen es nicht einmal die Entwickler zu wissen – auch der Weg zu den ersten KI-Chatbots ist gepflastert mit Versuch und Irrtum. Offenbar wenden KI-Entwickler dieselben Prinzipien an, mit denen sie auch ihre neuronalen Netze trainieren, indem sie immer neue Ansätze probieren und die besten weiterverfolgen. Der Entwicklungsprozess ist dadurch nicht vollständig kontrollierbar, was die Systeme umso menschenähnlicher macht. KI-Chatbots wie ChatGPT funktionieren durch die Kombination von verschiedenen Technologien und Prozessen, die es

ihnen ermöglichen, menschenähnliche Gespräche zu führen. Hier ist ein kurzer Überblick über den Prozess, möglichst allgemeinverständlich erklärt:

Verarbeitung von Benutzereingaben

Der erste Schritt ist das **Verstehen der Benutzereingabe** – also des Prompts. Dies geschieht durch Natural Language Processing (NLP), einer Technologie, die es dem Chatbot ermöglicht, die Sprache des Nutzers zu analysieren und zu interpretieren. NLP zerlegt den eingegebenen Text in einzelne Sätze, Wörter und Tokens und bestimmt deren Bedeutung anhand des Kontexts. Nachdem die Eingabe verarbeitet wurde, muss der Chatbot die **Absicht des Nutzers** dahinter erkennen. Ohne diese sogenannte **Intent-Erkennung** würden teilweise unsinnige Ergebnisse herauskommen. Der Chatbot analysiert die Schlüsselwörter und Phrasen im Text, um herauszufinden, was der Nutzer möchte – z.B. eine Information suchen, eine Bestellung aufgeben oder eine Support-Anfrage stellen.

Intent-Erkennung bei LLMs und indirekte Sprechakte

Wenn jemand sagt: »Es zieht!«, meint er damit eigentlich: »Bitte schließe die Tür!« Solche *indirekten Sprechakte* sind in der menschlichen Kommunikation alltäglich und wir verstehen diese spontan und zweifelsfrei. Große Sprachmodelle (LLMs) wie ChatGPT erkennen diese versteckten Absichten ebenfalls, indem sie den Kontext und sprachliche Muster analysieren. Sie verstehen nicht nur die wörtliche Aussage, sondern auch die dahinterliegende Intention. Dadurch können sie auf indirekte Anfragen angemessen reagieren und sind im Unterricht wertvolle Helfer beim Erfassen von Schülerbedürfnissen. Wenn ein Schüler sagt: »Das Thema ist ziemlich komplex«, können LLMs z.B. erkennen, dass er möglicherweise zusätzliche Unterstützung benötigt. Diese Fähigkeit zur Intent-Erkennung macht LLMs zu effektiven Werkzeugen für eine einfühlsame und responsive Lehrpraxis.

Basierend auf der erkannten Absicht greift der Chatbot nun auf relevante Quellen zu, um die benötigten Daten und Informationen abzurufen. Dieser Schritt ist tatsächlich am wenigsten dokumentiert, andererseits wissen wir auch wenig darüber, wie unser Gehirn diese Aufgabe erledigt und Wissen aus dem großen Fundus namens *Gedächtnis* abrufen. Mittlerweile ist ChatGPT nicht einmal mehr auf die eigene Basis der Trainingsdaten angewiesen und kann je nach Frage auch andere Quellen hinzunehmen: Dies kann eine Datenbank, ein Wissensgraph, eine API oder eine Suche bei Bing sein. Wenn der Nutzer beispielsweise nach dem Wetter fragt, wird der Chatbot eine Wetter-API oder einen Online-Service abfragen, um die aktuellen Wetterdaten zu erhalten.

Der nächste und letzte Schritt ist die Generierung einer passenden Antwort. Ein Sprachmodell wie GPT versucht nun eine Antwort zu erstellen, die sowohl inhalt-

lich korrekt als auch natürlich formuliert ist und den spezifischen Anforderungen des Prompts entspricht. Der Chatbot stellt sicher, dass die Antwort im besten Fall klar und verständlich ist. Moderne KI-Chatbots sind außerdem in der Lage, auch aus den Interaktionen mit Nutzern zu lernen. Durch maschinelles Lernen verbessern sie kontinuierlich ihre Fähigkeit, Nutzerabsichten zu erkennen und passende Antworten zu generieren. Feedback-Schleifen wie Rückfragen oder Daumen-nach-oben- bzw. Daumen-nach-unten-Klicks und Nutzerdaten werden verwendet, um die Modelle laufend zu verfeinern und die Leistung des Chatbots permanent zu optimieren.

Schema der KI-Chatbots

Alle KI-Chatbots funktionieren nach diesem einfachen Schema:

1. Eingaben der Nutzers verstehen
2. Die Absicht erkennen
3. Relevante Daten abrufen
4. Passende Antworten generieren
5. Kontinuierlich lernen und sich weiter anpassen

Finde das nächste Wort

ChatGPT basiert auf dem Konzept *Finde das nächste Wort*, das als grundlegender Mechanismus maschinellen Lernens im Bereich der Natural Language Processing (NLP) verwendet wird. Ein Sprachmodell wie ChatGPT wird auf riesigen Mengen an Textdaten trainiert: Schon in der Trainingsphase muss das Sprachmodell beim Lesen der Texte das jeweils nächste Wort vorhersagen und bekommt während des Trainings unmittelbar danach das Ergebnis als Feedback. Je länger ein Algorithmus an Texten trainiert, umso besser werden diese Voraussagen. Wie ein Kind beim Erwerb der Erstsprache lernt die KI, welche Möglichkeiten es im Spiel *Finde das nächste Wort* an jedem einzelnen Punkt gibt. Einige Wörter folgen mit sehr hoher Wahrscheinlichkeit, andere kommen eher selten vor und haben also eine geringere Wahrscheinlichkeit, als nächstes Wort zu erscheinen. Ziel des Trainings ist es, die Wahrscheinlichkeiten für das Auftreten von Wörtern in einem bestimmten Kontext zu lernen. Das Modell lernt immer besser, zu jedem einzelnen Zeitpunkt vorherzusagen, welche Wörter am wahrscheinlichsten als Nächstes in der vorliegenden Sequenz erscheinen.

Testen Sie sich selbst und vervollständigen Sie den Satz im folgenden Kasten: Starten Sie mit dem ersten Wort in der obersten Zeile, raten Sie das nächste und decken Sie die einzelnen Zeilen nach und nach auf, um ein direktes Feedback auf Ihre eigenen Antworten zu erhalten:

Finde das nächste Wort!

Herzlichen

Herzlichen Glückwunsch

Herzlichen Glückwunsch zum

Herzlichen Glückwunsch zum Geburtstag

Herzlichen Glückwunsch zum Geburtstag und

Herzlichen Glückwunsch zum Geburtstag und alles

Herzlichen Glückwunsch zum Geburtstag und alles Gute

Herzlichen Glückwunsch zum Geburtstag und alles Gute lieber

Herzlichen Glückwunsch zum Geburtstag und alles Gute lieber Thomas

Wir kennen das alle von unserem Smartphone, wenn wir Nachrichten in einem Messenger verfassen und die Schreibkorrektur uns immer wieder mehr oder weniger passende Vorschläge liefert. Die Treffergenauigkeit steigt dort übrigens, wenn Sie die richtige Sprache eingestellt haben, weil das die Zahl der Möglichkeiten für das nächste Wort verringert. Außerdem lernt dieser Assistent auch aus den Nutzerdaten. Nach *Liebe Grüße* folgt meistens Ihr eigener Vorname, mit bestimmten Gesprächspartnern tauschen Sie sich regelmäßig über dieselben Themen aus und verwenden dabei immer dieselben Namen und Wörter.

Der Trainingsprozess bei einem großen Sprachmodell beginnt immer mit der sogenannten **Tokenisierung**, bei der der Text in kleinere Einheiten, die *Tokens*, zerlegt wird. Diese Tokens können Wörter, Teile von Wörtern oder sogar einzelne Zeichen sein. Danach betrachtet das Modell den Kontext der vorhergehenden Tokens, um die Wahrscheinlichkeit des nächsten Tokens zu bestimmen. Beispielsweise, wenn die vorherigen Tokens »Der Hund« sind, könnte das Modell vorhersagen, dass das nächste Token mit hoher Wahrscheinlichkeit ein Verb wie »bellt« sein könnte. Für jedes Token im Trainingstext berechnet das Modell eine **Wahrscheinlichkeitsverteilung** über alle möglichen nächsten Tokens. Das Modell wird darauf trainiert, die Parameter so anzupassen, dass die Wahrscheinlichkeit der tatsächlich folgenden Tokens maximiert wird.

Bei der Generierung von Texten verhält sich ChatGPT wie ein *autoregressives Modell*, was bedeutet, dass es die Tokens eines nach dem anderen generiert. Es beginnt mit einem Start-Token und sagt das nächste Token basierend auf dem bisher generierten Kontext voraus. Dieser Prozess wird wiederholt, bis die gewünschte Länge des Texts erreicht oder ein vordefiniertes Endkriterium erfüllt ist. Durch diese Methode kann ChatGPT menschenähnliche Texte erstellen, indem es (wie im Training gelernt) auf das Konzept *Finde das nächste Wort* zurückgreift, um kontextuell relevante und kohärente Sätze zu bilden.