

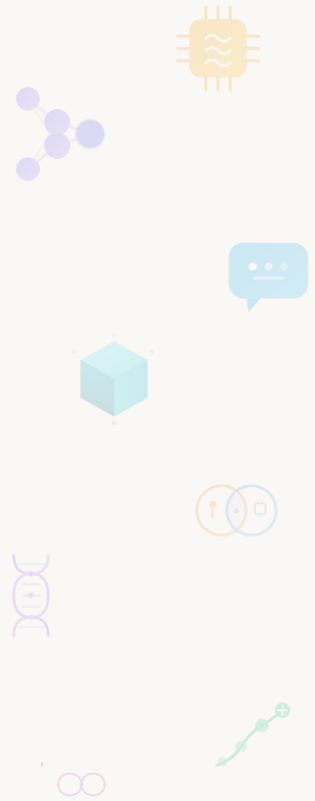


bildungssprache.net

KI Glossar

55 wichtige Begriffe
von A bis Z

LLM, Token, Halluzination, RAG, Prompt Engineering,
Slop, Grounding und mehr. Verständlich erklärt.



Über dieses Glossar

Wer über Künstliche Intelligenz mitreden will, stolpert schnell über Fachbegriffe: LLM, Token, Halluzination, RAG, Prompt Engineering. Was bedeutet das alles?

Dieses Glossar erklärt die wichtigsten Begriffe rund um ChatGPT, Claude und Co. Verständlich, ohne unnötigen Ballast. Ja, ist viel Englisch dabei. Und weil sich in der KI-Welt ständig etwas ändert, wächst diese Liste mit.

Die Einträge sind alphabetisch sortiert. Jeder Begriff wird in ein bis drei Sätzen erklärt, ergänzt um Kontext und Einordnung. Fachsprache, wo nötig, Klartext, wo möglich.

Dieses PDF basiert auf dem Online-Glossar von bildungssprache.net und wurde im April 2026 aktualisiert.

AGI (Artificial General Intelligence)

Eine hypothetische KI, die jede intellektuelle Aufgabe ausführen kann, die ein Mensch bewältigt. Noch Zukunftsmusik, aber erklärtes Ziel von OpenAI und anderen. Wann es so weit sein soll, hängt davon ab, wen man fragt. Die Schätzungen reichen von »in drei Jahren« bis »womöglich nie«.

Agenten (KI-Agenten)

KI-Systeme, die eigenständig Aufgaben ausführen, Entscheidungen treffen und mit ihrer Umgebung interagieren. Sie können Tools nutzen, im Web suchen oder Code ausführen, ohne dass der Nutzer jeden Schritt anweisen muss. Im Kern ist ein Agent eine KI, die nicht nur antwortet, sondern handelt.

Algorithmus

Eine Abfolge von Regeln, die ein Computer befolgt, um eine Aufgabe zu lösen. Klingt trocken, ist aber die Grundlage von allem, was KI tut. Jedes Sprachmodell, jede Bilderkennung, jeder Empfehlungsalgorithmus bei Netflix basiert darauf.

AI Literacy (KI-Kompetenz)

Die Fähigkeit, KI-Systeme in ihren Grundzügen zu verstehen, ihre Ergebnisse kritisch einzuordnen und sie sinnvoll einzusetzen. Dazu gehört das Wissen um Stärken, Grenzen und Risiken wie Halluzinationen oder Bias. Gilt als digitale Grundbildung des 21. Jahrhunderts. Ob das Bildungssystem da mitkommt, steht auf einem anderen Blatt.

Alignment

Die Ausrichtung einer KI an menschlichen Werten und Absichten. Ziel ist es, dass das Modell hilfreich, ehrlich und ungefährlich handelt. Eng verknüpft mit RLHF (siehe unten). Das Problem: »menschliche Werte« sind sich nicht immer einig.

API (Application Programming Interface)

Eine Schnittstelle, über die Programme miteinander kommunizieren. Die APIs von OpenAI oder Anthropic ermöglichen Entwicklern, ChatGPT oder Claude in eigene Anwendungen einzubauen. Wer die API nutzt, zahlt pro Token und hat dafür mehr Kontrolle als im Chatfenster.

B



Benchmark

Ein standardisierter Test, mit dem die Leistung von KI-Modellen verglichen wird. Bekannte Benchmarks messen Mathematik, Programmierung, logisches Denken oder Sprachverständnis. Die Ergebnisse klingen beeindruckend, sagen über die Praxistauglichkeit aber oft wenig aus. Hersteller optimieren gern auf Benchmarks, weil sich die Zahlen gut in Pressemitteilungen machen.

Bias (Verzerrung)

Systematische Verzerrungen in KI-Ergebnissen, die aus den Trainingsdaten stammen. Wenn historische Daten Vorurteile enthalten, etwa bei Geschlecht, Herkunft oder Alter, reproduziert die KI diese Muster. Das ist kein Bug, sondern ein Spiegel der Daten, mit denen sie gefüttert wurde.

Bots

Automatisierte Programme, die Aufgaben im Internet ausführen. Chatbots sind eine spezielle Form, die auf Texteingaben reagiert. Nicht zu verwechseln mit den Spam-Bots, die deine Kommentarspalte zumüllen.

C



Chain-of-Thought (CoT)

Eine Prompting-Technik, bei der man die KI auffordert, Schritt für Schritt zu denken, bevor sie antwortet. Verbessert die Ergebnisse bei Logik, Mathematik und komplexen Aufgaben erheblich. Im Prinzip das, was ein guter Lehrer auch verlangen würde: erst denken, dann schreiben.

Computer Vision

Der Bereich der KI, der sich mit dem Verstehen und Interpretieren von Bildern und Videos befasst. Ermöglicht Gesichtserkennung, Objekterkennung und mehr. Wenn dein Smartphone ein Gesicht im Foto erkennt, ist das Computer Vision bei der Arbeit.

Constitutional AI

Ein von Anthropic entwickelter Trainingsansatz. Statt das Modell ausschließlich durch menschliches Feedback zu formen, bekommt es einen Satz von Prinzipien, an denen es sich selbst ausrichtet. Eine Art Grundgesetz für KI-Verhalten. Ob das in der Praxis besser funktioniert als RLHF allein, ist Gegenstand der Forschung.

Context Window (Kontextfenster)

Die maximale Menge an Text, die ein Sprachmodell gleichzeitig verarbeiten kann. Gemessen in Tokens. Größere Kontextfenster ermöglichen längere Gespräche und die Analyse umfangreicherer Dokumente. Claude unterstützt bis zu 200.000 Tokens, in der erweiterten Version bis zu 1 Million (Stand 03/2026).

D



Data Literacy (Datenkompetenz)

Die Fähigkeit, Daten kritisch zu lesen, zu interpretieren und fundierte Schlüsse daraus zu ziehen. In einer Welt voller KI-generierter Statistiken und Visualisierungen unverzichtbar. Wer Zahlen nicht einordnen kann, wird von ihnen eingeordnet.

Deep Learning

Ein Teilbereich des maschinellen Lernens, der auf neuronalen Netzwerken mit vielen Schichten basiert. Grundlage moderner Sprachmodelle und Bildgeneratoren.

Deep Research

Eine Funktion, die KI-Modelle befähigt, selbstständig und iterativ im Web zu recherchieren, Quellen auszuwerten und Ergebnisse zusammenzufassen. ChatGPT, Claude, Gemini und Perplexity bieten eigene Varianten. Nützlich, wenn man nicht weiß, wo man anfangen soll. Die Ergebnisse muss man trotzdem prüfen.

Deskilling

Der schleichende Verlust von Fähigkeiten durch die Auslagerung kognitiver Aufgaben an KI. Wer Lösungen nur noch generieren lässt, ohne die Lösungswege zu verstehen, verlernt das eigenständige Denken. Das Gegenstück ist Newskilling.

Diffusion

Das Verfahren hinter den meisten Bildgeneratoren wie Midjourney oder DALL-E. Das Modell lernt, aus verrauschten Daten schrittweise klare Bilder zu rekonstruieren, und kann so aus Textbeschreibungen neue Bilder erzeugen.

E



Embedding

Eine mathematische Darstellung von Wörtern, Sätzen oder Dokumenten als Zahlenvektoren. Ermöglicht es KI-Systemen, semantische Ähnlichkeiten zu erkennen, dass also »König« und »Monarch« ähnliche Bedeutungen haben.

Explainable AI (XAI)

Ansätze, die Entscheidungen von KI-Systemen transparent und nachvollziehbar machen. Wichtig für Vertrauen und regulatorische Anforderungen, etwa im Bildungsbereich oder in der Medizin.



Fine-Tuning

Das Anpassen eines vortrainierten Modells auf spezifische Aufgaben oder Datensätze. Macht generelle Modelle zu Spezialisten. In der Praxis oft aufwendig und teuer, weshalb viele Anwendungen eher auf gute Prompts setzen.



GANs (Generative Adversarial Networks)

Ein Modell, bei dem zwei neuronale Netze gegeneinander arbeiten: eines erzeugt Inhalte, das andere bewertet sie. War die Grundlage vieler früher Bildgeneratoren, wurde aber weitgehend von Diffusionsmodellen abgelöst.

Generative KI

KI-Systeme, die neue Inhalte erzeugen: Texte, Bilder, Musik, Code, Video. ChatGPT, Claude, DALL-E, Midjourney und Kling sind Beispiele. Der Unterschied zu früherer KI: Sie antwortet nicht nur, sie produziert.

GPT (Generative Pre-trained Transformer)

Die Modellarchitektur hinter ChatGPT. »Generative« bedeutet, dass es Inhalte erzeugt; »Pre-trained«, dass es auf großen Datenmengen vortrainiert wurde; »Transformer« beschreibt die zugrunde liegende Architektur. Das aktuelle Flaggschiff ist o3, das erste Reasoning-Modell von OpenAI mit multimodalen Werkzeugen (Stand 03/2026).

Grounding

Techniken, die KI-Antworten an verifizierbare Fakten und Quellen binden. Websuche, Dokumentenzugriff und RAG sind Formen von Grounding. Ziel: weniger Halluzinationen, mehr Substanz.

Guardrails

Sicherheitsmechanismen, die das Verhalten von KI-Modellen einschränken. Sie verhindern, dass die KI gefährliche, illegale oder anderweitig problematische Inhalte erzeugt. Mal sinnvoll, mal übervorsichtig. Die Balance ist schwer zu treffen.

H



Halluzination

Wenn eine KI falsche Informationen generiert und diese überzeugend präsentiert. Erfundene Fakten, falsche Zitate, nicht existierende Quellen. Ein bekanntes Problem aller Sprachmodelle, das sich verringert hat, aber nicht verschwunden ist.

I



Inferenz

Der Prozess, bei dem ein trainiertes Modell auf neue Eingaben reagiert und Ausgaben erzeugt. Jede Antwort von ChatGPT ist das Ergebnis einer Inferenz. Das Training ist der teure Teil, die Inferenz der Teil, den du bezahlst.

K



Künstliche Intelligenz (KI)

Das Feld der Informatik, das sich mit der Entwicklung von Systemen befasst, die Aufgaben ausführen, die normalerweise menschliche Intelligenz erfordern.

L



LLM (Large Language Model)

Große Sprachmodelle, trainiert auf riesigen Textmengen. GPT-4, Claude, Gemini und Llama sind LLMs. Sie können Texte verstehen, generieren, übersetzen und zusammenfassen. Die Modelle werden schnell besser und schnell billiger, was den gesamten Markt in Bewegung hält.



Machine Learning (Maschinelles Lernen)

Ein Teilgebiet der KI, bei dem Algorithmen aus Daten lernen, statt explizit programmiert zu werden. Die Grundidee: Man zeigt dem System genug Beispiele, und es findet die Muster selbst.

MCP (Model Context Protocol)

Ein von Anthropic entwickeltes offenes Protokoll. MCP erlaubt KI-Modellen mit externen Datenquellen und Tools zu kommunizieren. Ermöglicht Integrationen mit Datenbanken, APIs und anderen Systemen. Vergleichbar mit einem USB-Standard für KI-Werkzeuge.

MoE (Mixture of Experts)

Eine Modellarchitektur, bei der nicht alle Parameter gleichzeitig aktiv sind. Das Modell aktiviert je nach Anfrage nur einen Teil seiner »Experten«. Macht große Modelle effizienter und günstiger. DeepSeek nutzt dieses Prinzip.

Multimodal

KI-Modelle, die verschiedene Medientypen verarbeiten können: Text, Bilder, Audio, Video. Die meisten aktuellen Modelle sind multimodal. Sie »sehen« Bilder, verstehen deren Inhalt und können darüber sprechen. Manche generieren auch Bilder, Code und Audiodateien in einem Arbeitsschritt.



Neuronales Netzwerk

Ein Computermodell, das lose von der Struktur des menschlichen Gehirns inspiriert ist. Besteht aus Schichten von »Neuronen«, die Eingaben verarbeiten und Muster erkennen. »Lose inspiriert« ist wichtig: Mit echten Gehirnen hat das ungefähr so viel zu tun wie ein Papierflugzeug mit einer Boeing.

Newskilling

Ein von der Wissenschaftlerin Doris Weßels geprägter Begriff. Beschreibt den Prozess, durch die Zusammenarbeit mit KI neue Fähigkeiten zu entwickeln, statt bestehende zu verlieren. Die Idee: Die Entlastung von Routinearbeit bewusst nutzen, um anspruchsvollere Denkaufgaben anzugehen. Das Gegenteil von Deskillling.

NLP (Natural Language Processing)

Natürliche Sprachverarbeitung. Der Bereich der KI, der sich mit dem Verstehen und Erzeugen menschlicher Sprache befasst. Alles, was ChatGPT tut, ist im weitesten Sinne NLP.



Open Source

Software, deren Quellcode öffentlich zugänglich ist. Llama (Meta) und DeepSeek sind Open-Source-Sprachmodelle; GPT und Claude sind es nicht. Die Debatte, ob offene oder geschlossene Modelle sicherer sind, wird leidenschaftlich geführt. Einigkeit ist nicht in Sicht.

Overfitting

Wenn ein Modell die Trainingsdaten zu gut lernt und dadurch bei neuen Daten schlecht abschneidet. Es hat auswendig gelernt statt verstanden. Kommt dir bekannt vor? Sollte es auch.



Parameter

Die internen Einstellungen eines Modells, die während des Trainings gelernt werden. Mehr Parameter bedeuten oft, aber nicht immer, mehr Fähigkeiten. Die genauen Zahlen verraten die Hersteller ungern.

Prompt

Die Eingabe, die du einer KI gibst. Kann eine Frage, eine Anweisung oder ein komplexer Auftrag sein. Die Qualität des Prompts beeinflusst die Qualität der Antwort erheblich. Wer schlampig fragt, bekommt schlampige Antworten.

Prompt Engineering

Die Kunst, effektive Prompts zu formulieren. Umfasst Techniken wie Few-Shot-Learning (Beispiele geben), Chain-of-Thought (schrittweises Denken anfordern) und Rollenzuweisung. Weniger Zauberei als gesunder Menschenverstand, aber mit überraschend großem Effekt.

R



RAG (Retrieval-Augmented Generation)

Eine Technik, bei der ein Sprachmodell zusätzliches Wissen aus externen Quellen abrufen, bevor es antwortet. Reduziert Halluzinationen und ermöglicht aktuelle Informationen. Die meisten KI-Suchfunktionen basieren auf einer Variante von RAG.

Reasoning

Die Fähigkeit eines Modells, schrittweise zu denken, bevor es antwortet. Modelle wie OpenAI o3 oder DeepSeek-R1 zeigen ihren Denkprozess offen. Verbessert die Ergebnisse bei Mathematik, Logik und komplexen Aufgaben. Der Preis: längere Wartezeiten und höhere Kosten.

Reinforcement Learning (RL)

Eine Lernmethode, bei der ein Modell durch Belohnungen und Bestrafungen lernt. RLHF (Reinforcement Learning from Human Feedback) ist entscheidend für die Feinabstimmung von Chatbots. Im Grunde dasselbe Prinzip wie bei der Hundeerziehung, nur mit Mathematik statt Leckerlis.

Schwache KI (Narrow AI)

KI-Systeme, die auf bestimmte Aufgaben spezialisiert sind. Alle heutigen KI-Systeme sind schwache KI, auch wenn sie in ihren Spezialgebieten übermenschliche Leistungen bringen. Starke KI (AGI) bleibt vorerst Theorie.

Slop

Schlick, Schlonz, hingeklatschtes Zeug. Der englische Begriff für minderwertige, generische KI-Ausgaben. Texte, die technisch korrekt, aber inhaltslos, austauschbar und sofort als maschinell erkennbar sind. Das Gegenteil von dem, was man sich von einem guten Prompt erhofft. Das Internet füllt sich damit.

System Prompt

Eine versteckte Anweisung, die das Verhalten eines Chatbots steuert. Definiert Persönlichkeit, Einschränkungen und Verhaltensregeln. Der Nutzer sieht ihn normalerweise nicht, er wirkt aber bei jeder Antwort mit.

Temperature

Ein Parameter, der steuert, wie kreativ oder vorhersagbar eine KI antwortet. Niedrige Temperature (nahe 0) erzeugt gleichförmige, sichere Antworten. Hohe Temperature macht die Ausgabe variantenreicher, aber auch unzuverlässiger. Für Faktenabfragen niedrig halten, für kreative Texte hochdrehen.

Token

Die Einheit, in der Sprachmodelle Text verarbeiten. Ein Token ist oft ein Wort oder Wortteil. »Bildungssprache« könnte je nach Modell 2 bis 4 Tokens sein. Die Tokenanzahl bestimmt Kosten und Kontextgrenzen. Wer die API nutzt, denkt in Tokens statt in Wörtern.

Training

Der Prozess, bei dem ein Modell aus Daten lernt. Die großen Sprachmodelle wurden auf Milliarden von Textseiten trainiert: Bücher, Websites, Artikel, Code. Das Training kostet Millionen, manchmal hunderte Millionen Dollar. Die Ergebnisse nutzt du für ein paar Euro im Monat.

Transformer

Die Architektur hinter modernen Sprachmodellen. Entwickelt 2017 von Google-Forschern in einem Paper mit dem bescheidenen Titel Attention Is All You Need. Ermöglicht effiziente Verarbeitung von Textsequenzen und bildet die Grundlage für GPT, Claude, Gemini und praktisch alle anderen.

Turing-Test

Ein von Alan Turing vorgeschlagener Test: Kann eine Maschine einen Menschen in einem Gespräch täuschen? Moderne Chatbots bestehen diesen Test oft, was weniger über die Intelligenz der Maschinen aussagt als über die Grenzen des Tests.

V



Vibe Coding

Programmieren durch Beschreiben statt durch Tippen von Code. Du sagst der KI, was du willst, sie schreibt den Code. Der Begriff stammt von Andrej Karpathy und beschreibt einen Paradigmenwechsel in der Softwareentwicklung. Ob das Ergebnis funktioniert, muss man allerdings immer noch selbst prüfen.

Z



Zero-Shot Learning

Ein Modell löst eine Aufgabe, für die es nie explizit trainiert wurde. Im Gegensatz dazu steht Few-Shot Learning, bei dem einige Beispiele gegeben werden. Die Tatsache, dass das funktioniert, ist im Grunde das Erstaunlichste an modernen Sprachmodellen.

Sternenvogel Verlagsgesellschaft mbH

Marktplatz 1
09557 Flöha

Vertreten durch Sven Lennartz als Geschäftsführer
verwaltung@sternenvogelpoesie.de

Umsatzsteuer ID: DE453657101

Handelsregister: HRB 86141494597, Amtsgericht Chemnitz

bildungssprache.net/ki-von-a-bis-z-wichtige-begriffe-erklaert/