Glossar Künstliche Intelligenz

Die wichtigsten Begriffe rund um KI – kurz und verständlich erklärt

Dieses Handout fasst zentrale Fachbegriffe aus Grundlagen, Technik, Anwendung und Recht zusammen. Die Begriffe sind nach Themen gruppiert; am Ende finden Sie Praxistipps für den Umgang mit KI-Werkzeugen.

Grundbegriffe

Begriff	Erklärung
Künstliche Intelligenz (KI / AI)	Oberbegriff für Systeme, die Aufgaben lösen, für die normalerweise menschliche Intelligenz nötig ist – z. B. Sprache verstehen, Bilder erkennen, Entscheidungen treffen.
Algorithmus	Eindeutige Abfolge von Schritten zur Lösung eines Problems. KI-Systeme bestehen aus vielen Algorithmen.
Modell	Das Ergebnis eines Trainings: ein mathematisches Gebilde, das aus Eingaben Ausgaben berechnet (z. B. ein Sprachmodell).
Maschinelles Lernen (ML)	Teilgebiet der KI: Systeme lernen Muster aus Daten, statt fest programmiert zu werden.
Deep Learning	Maschinelles Lernen mit tiefen neuronalen Netzen (viele Schichten) – Grundlage der meisten modernen KI.
Neuronales Netz	Rechenmodell, lose vom Gehirn inspiriert: verbundene „Neuronen“, deren Gewichte beim Training angepasst werden.
Schwache / starke KI	Schwache KI ist auf bestimmte Aufgaben spezialisiert (alle heutigen Systeme). Starke KI wäre eine allgemeine, menschenähnliche Intelligenz.
AGI	Artificial General Intelligence: hypothetische KI, die praktisch jede geistige Aufgabe so gut wie ein Mensch lösen kann.

Sprachmodelle & generative KI

Begriff	Erklärung
Generative KI	KI, die neue Inhalte erzeugt: Text, Bilder, Audio, Video oder Code.
LLM (Large Language Model)	Großes Sprachmodell, trainiert auf riesigen Textmengen; versteht und erzeugt Text. Beispiele: Claude, GPT, Gemini, Llama, Mistral.
Transformer	Architektur hinter modernen Sprachmodellen; arbeitet mit dem Attention-Mechanismus.
Attention	Mechanismus, mit dem ein Modell gewichtet, welche Teile der Eingabe für die aktuelle Ausgabe wichtig sind.
Token	Texteinheit, mit der ein Modell arbeitet – oft ein Wortteil. Im Deutschen entspricht ein Wort meist mehr als einem Token.
Kontextfenster	Maximale Textmenge (in Tokens), die ein Modell auf einmal „im Blick“ hat – Eingabe plus Ausgabe.
Prompt	Eingabe bzw. Anweisung an ein KI-Modell.
Prompt Engineering	Gezieltes Formulieren von Prompts, um bessere Ergebnisse zu erhalten.

Begriff	Erklärung
Systemprompt	Grundanweisung im Hintergrund, die das Verhalten des Modells für die ganze Unterhaltung festlegt.
Zero-Shot / Few-Shot	Aufgabe ohne Beispiele (Zero-Shot) bzw. mit einigen Beispielen im Prompt (Few-Shot).
Chain of Thought	Das Modell denkt Schritt für Schritt, bevor es antwortet – verbessert oft komplexe Ergebnisse.
Reasoning-Modell	Sprachmodell, das vor der Antwort intern länger „nachdenkt“; stärker bei Logik, Mathematik und Planung, dafür langsamer.
Halluzination	Das Modell erzeugt plausibel klingende, aber falsche oder erfundene Informationen.
Temperatur	Parameter für die Zufälligkeit der Ausgabe: niedrig = vorhersehbar, hoch = kreativer.
Multimodal	Modell, das mehrere Datenarten verarbeitet, z. B. Text, Bild und Audio.
Chatbot	Anwendung, mit der man per Dialog mit einer KI kommuniziert.
Copilot / KI-Assistent	In Software eingebettete KI, die bei Aufgaben unterstützt, z. B. Microsoft 365 Copilot.

Bild, Audio & Video

Begriff	Erklärung
Diffusionsmodell	Verfahren zur Bilderzeugung: Aus zufälligem Rauschen wird schrittweise ein Bild „herausgerechnet“.
Text-to-Image / -Video	Erzeugung von Bildern bzw. Videos aus einer Textbeschreibung.
Spracherkennung (Speech-to-Text)	Umwandlung gesprochener Sprache in Text, z. B. für Transkripte und Untertitel.
Sprachsynthese (Text-to-Speech)	Umwandlung von Text in natürlich klingende Sprache; inzwischen auch Stimmklone möglich.
Computer Vision	KI zur Analyse von Bildern und Videos (Objekterkennung, Texterkennung/OCR, Gesichtserkennung).

Training & Daten

Begriff	Erklärung
Trainingsdaten	Die Daten, aus denen ein Modell lernt.
Training	Prozess, in dem die Parameter eines Modells anhand von Daten angepasst werden.
Parameter / Gewichte	Interne Zahlenwerte eines Modells; große Sprachmodelle haben Milliarden davon.
Pretraining	Erstes, sehr aufwendiges Training auf großen allgemeinen Datenmengen.
Fine-Tuning	Nachtrainieren eines vortrainierten Modells für einen bestimmten Zweck.
RLHF	Reinforcement Learning from Human Feedback: Training mit menschlichen Bewertungen, damit das Modell hilfreicher und sicherer antwortet.

Begriff	Erklärung
Überwachtes / unüberwachtes Lernen	Lernen mit beschrifteten Daten („Bild zeigt Katze“) bzw. selbstständiges Finden von Mustern in unbeschrifteten Daten.
Bestärkendes Lernen	Reinforcement Learning: Lernen durch Belohnung und Bestrafung für Aktionen.
Overfitting	Das Modell lernt die Trainingsdaten „auswendig“ und verallgemeinert schlecht auf neue Daten.
Bias (Verzerrung)	Systematische Schieflagen in Ergebnissen, oft durch unausgewogene Trainingsdaten.
Synthetische Daten	Künstlich erzeugte Daten (oft von KI selbst), die echte Daten ergänzen oder ersetzen.
Knowledge Cutoff	Zeitpunkt, bis zu dem die Trainingsdaten reichen. Spätere Ereignisse kennt das Modell ohne Websuche nicht.
Benchmark	Standardisierter Test, mit dem die Leistung verschiedener Modelle verglichen wird.

Anwendung & Technik

Begriff	Erklärung
Inferenz	Nutzung eines fertig trainierten Modells, also das Erzeugen von Antworten.
Grounding	Antworten werden an überprüfbare Quellen gebunden (z. B. eigene Dokumente, Websuche), um Halluzinationen zu verringern.
RAG	Retrieval-Augmented Generation: Das Modell holt vor der Antwort passende Informationen aus Dokumenten oder Datenbanken und stützt sich darauf.
Embedding	Darstellung von Text, Bildern o. Ä. als Zahlenvektor; ähnliche Inhalte liegen nahe beieinander.
Vektordatenbank	Datenbank für Embeddings, ermöglicht Ähnlichkeitssuche; oft für RAG genutzt.
API	Programmierschnittstelle, über die Software ein KI-Modell ansprechen kann.
KI-Agent	System, das selbstständig mehrere Schritte plant und ausführt und dabei Werkzeuge nutzt (Websuche, Dateien, Programme).
Tool Use / Function Calling	Fähigkeit eines Modells, externe Werkzeuge oder Funktionen aufzurufen.
MCP (Model Context Protocol)	Offener Standard, mit dem KI-Anwendungen an externe Datenquellen und Tools angebunden werden.
Lokales Modell	KI-Modell, das auf dem eigenen Rechner läuft statt in der Cloud, z. B. über LM Studio oder Ollama.
Open-Weight-Modell	Modell, dessen Gewichte öffentlich verfügbar sind und das man selbst betreiben kann.
Quantisierung	Verkleinerung eines Modells durch geringere Rechengenauigkeit – damit läuft es auch auf normaler Hardware, mit leichtem Qualitätsverlust.
GPU / NPU	Grafikprozessor bzw. Neural Processing Unit: Spezialchips, die KI-Berechnungen stark beschleunigen.

Begriff	Erklärung
NLP	Natural Language Processing: maschinelle Verarbeitung natürlicher Sprache.

Ethik, Recht & Sicherheit

Begriff	Erklärung
AI Act (KI-Verordnung)	EU-Gesetz zur Regulierung von KI nach Risikoklassen; seit August 2024 in Kraft, die Pflichten gelten gestaffelt.
KI-Kompetenz	Laut AI Act müssen Organisationen, die KI einsetzen, für ausreichende KI-Kenntnisse ihrer Mitarbeitenden sorgen.
Risikoklassen	Einteilung im AI Act: verbotene Praktiken, Hochrisiko-KI, KI mit Transparenzpflichten, geringes Risiko.
Kennzeichnung	KI-generierte Inhalte wie Deepfakes und der Einsatz von Chatbots müssen erkennbar gemacht werden.
Datenschutz (DSGVO)	Besonders relevant bei der Eingabe personenbezogener Daten in KI-Dienste.
Urheberrecht	Klärt, wem KI-Ergebnisse gehören und ob Trainingsdaten genutzt werden durften – rechtlich vielfach noch offen.
Deepfake	Täuschend echte, KI-generierte Bilder, Videos oder Stimmen realer Personen.
Alignment	Ausrichtung von KI-Systemen an menschlichen Werten und Absichten.
Guardrails	Schutzmechanismen, die verhindern sollen, dass eine KI schädliche oder unerwünschte Inhalte ausgibt.
Prompt Injection	Angriff, bei dem versteckte Anweisungen in Texten oder Webseiten eine KI manipulieren sollen.
Jailbreak	Versuch, die Sicherheitsregeln eines Modells durch geschickte Prompts zu umgehen.
Erklärbarkeit (XAI)	Explainable AI: Nachvollziehbarkeit, wie eine KI zu ihrem Ergebnis kommt.
Human in the Loop	Ein Mensch prüft oder bestätigt KI-Ergebnisse, bevor sie wirksam werden.

Praxistipps für den Umgang mit KI

1. **Kontext geben:** Wer bin ich, wofür brauche ich das Ergebnis, wer liest es?
2. **Konkret werden:** Länge, Format, Tonfall und Zielgruppe angeben.
3. **Beispiele zeigen:** Ein Muster des gewünschten Ergebnisses verbessert die Antwort deutlich.
4. **Schrittweise arbeiten:** Große Aufgaben in Teilschritte zerlegen und nachfragen.
5. **Ergebnisse prüfen:** Fakten, Zahlen und Quellen immer kontrollieren – KI kann sich irren.
6. **Datensparsam sein:** Keine personenbezogenen oder vertraulichen Daten in öffentliche KI-Dienste eingeben.