



11/12.02.2026

von Sicherheit, Akzeptanz und ganz Grundsätzlichem

KI oder nicht KI



KI – Was ist das eigentlich?

Begriffsdefinition
Klassifizierung von KI

Künstliche Intelligenz – Oft verwendete definierende Merkmale

- **Keine einheitliche Definition**
- **Einzelmerkmale meist nicht ausreichend für Abgrenzung**

- Computerbasiertes System
- Verarbeitet Input zu Output
- Mustererkennung
- Klassifizierung
- **Anpassung der Problemlösungswege nach Training (Lernfähigkeit)**

- Nachbildung menschlicher Intelligenz (Ziel: Artificial General Intelligence/Superintelligence)
 - Folgerung, Wahrnehmung, räumliche Orientierung, Problemlösung, Motorik, Spracherkennung, Verständnis von Inhalten, Kreativität, Verständnis von Emotionen, etc.

Künstliche Intelligenz – Artificial General vs. Narrow Intelligence

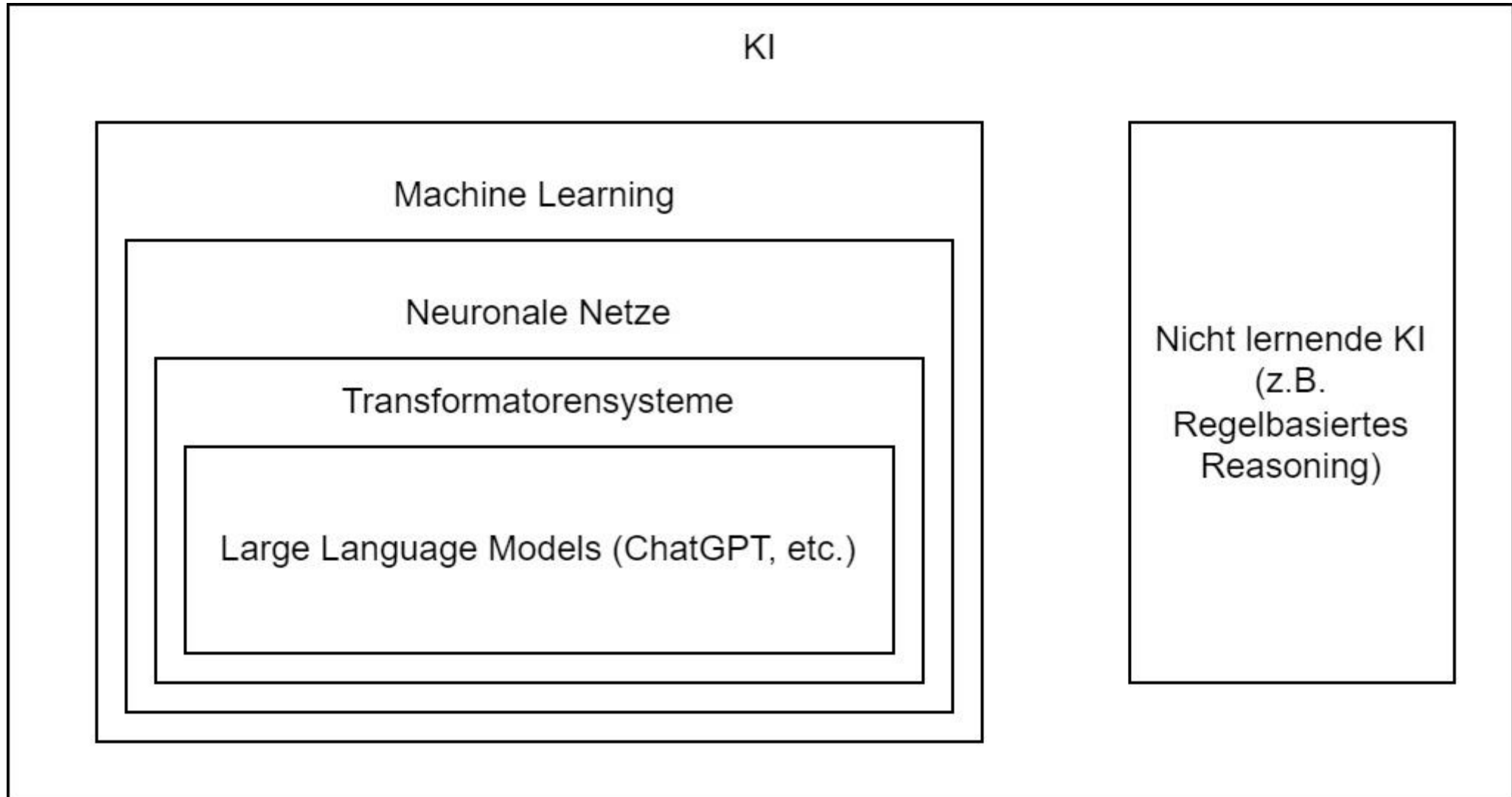
General AI (AGI)

- Theoretischer erster Zielzustand
- „Tatsächliche Intelligenz“
- Lösung vieler verschiedener Probleme
- Fähig menschliche Intelligenz komplett nachzubilden
- Vorstufe von „Superintelligenz“, die menschliches Denken übersteigt

Narrow AI

- Lösung spezifischer Probleme oder Problemräume
- Definiertes Anwendungsgebiet
- Anpassung an neue Gegebenheiten

Klassifizierung folgender Systeme



Künstliche Intelligenz – Einteilung nach Funktionsweise

- Regelbasiertes Reasoning System (z.B. Entscheidungsbaum)
 - Feste Regeln geben wiederholbare Ergebnisse
- Machine-Learning System (z.B. simple Klassifikation)
 - Anpassung von Systemverhalten an Lern-Input
- Neuronales Netz (z.B. simple Mustererkennung)
 - Bildet Neuronen nach
 - Verändert Informationen anhand von Knoten, Verbindungen und Gewichten
- Deep-Learning System (z.B. Computer Vision)
 - Hierarchisches Neuronales Netz mit Schichtstruktur
- Transformatoren-System (z.B. Verständnis natürlicher Sprache)
 - Basis für viele moderne KI-Systeme
 - Verarbeitung von Kontexten
- LLM – Large Language Model
 - Multifunktional
 - Trainiert mit riesigen Datensets

Künstliche Intelligenz – Fertig angelernte KI?

Ist ein fertig ausgelerntes System eine KI?

KI – Viele Einteilungen möglich

→ Nach Fähigkeit

→ General AI vs. Narrow AI

→ Nach Technologie

→ Symbolic AI (veraltet), Neural Network, Deep-Learning Neural Network mit Transformatoren

→ Nach Einsatzzweck

→ Computer Vision, Spracherkennung, Datenregression

→ Nach Risikobetrachtung

→ **Beispiel AI-Act**



KI - Wirtschaftlicher Einsatz

Folgen von KI-Einsatz, Hoffnungen und Ängste

Positivszenarien

- Revolutionäre Technik
 - Vergleichbar mit Industrialisierung oder dem Internet
- Einsparung von Arbeit
- Kostensenkung
- Lösung vorher unlösbarer Probleme
- Durchbrüche in Wirtschaft und Wissenschaft
- Bessere Software
- Bereicherung von Kunst und Kultur
- Befähigung von Laien durch intelligente Werkzeuge
- Bessere Arbeitsergebnisse
- Personal Assistants
- Künstliche Lebenspartner
- Weitere Evolution des Menschen
- Superintelligenz zum Wohl der Menschheit

Negativszenarien

- Verlust von Arbeitsplätzen
- Sinkende Qualitätsstandards durch geringere Eintrittshürden
- Massengenerierte Inhalte
- Verdummung der Gesellschaft
- Verwendung von KI zu gesellschaftlicher Steuerung
- Verbreiten von Fehlinformationen und gefälschter Inhalte
- Leicht zugängliche Werkzeuge für Verbrechen
- Generierung illegaler Inhalte
- Diebstahl geistigen Eigentums
- Massive Verschwendung von Ressourcen
- Wirtschaftskrisen und Engpässe durch Blasen
- Weltuntergang

Typische Annahmen in der Wirtschaft

- KI macht Prozesse besser, einfacher und günstiger
- Das Ersetzen von Personal nützt Unternehmen durch sinkende Kosten
- KI ist gewünscht und Betroffene nehmen neue Technik gut an
- Die Einführung von KI ist wirtschaftlich sinnvoll
- Wir erleben eine „zweite Digitalisierung“
- Jeder wird KI nutzen

KI-Schlagzeilen – Ein gemischtes Bild

Can AI companies become profitable?

Most organizations are still in the experimentation or piloting phase: Nearly two-thirds of respondents say their organizations have not yet begun scaling AI across the enterprise.

Despite \$30–40 billion in enterprise investment into GenAI, this report uncovers a surprising result in that 95% of organizations are getting zero return.

Why Everybody Is Losing Money On AI

Duolingo's AI-First Disaster: A Cautionary Tale of What Happens When You Replace Rather Than Partner

AI is Redefining How Startups Scale

NOVEMBER 4, 2025 | 2 MINUTE READ | SHARE:   in



AI Startups Are Changing the Game for Growth and Scale



4:20

Advancements in generative Artificial Intelligence (“AI”) and large language models are enabling early-stage companies to scale faster than ever before. Venture investors who can identify and access the ultimate category creating companies are benefiting from unprecedented growth and bigger exits.

How AI can help companies cut costs



KI – Sicherheit und Umgang mit KI

Organisatorische Regelungen

KI, das große Unbekannte

→ Viel diskutiert im Kontext KI und Sicherheit:

- Wie können KI-Produkte Sicherheit unterstützen?
- Welche Arbeiten kann KI bei Sicherheitsmanagement erleichtern?
- KI als Werkzeug zum Angriff und zur Abwehr von „Cyberattacken“
- Wie kann ich KI zweckentfremden?
- KI und Datenschutz

→ Wenig diskutiert:

- Wie setze ich KI sicher ein?
- Muss ich KI absichern?
- Stellt KI ein Sicherheitsproblem dar oder nicht?
- Wie regle ich den Einsatz von KI?

KI ist nicht gleich KI – ein Top-Down Ansatz

→ Wir erinnern uns: Es gibt viele verschiedene Arten von KI

→ Erster Schritt: **Klarheit schaffen**

→ (Annahme: ISMS besteht)

→ Weitere Schritte:

→ KI-Verwendung regeln

→ KI-Einsatz und Einführung regeln

→ Alternativen prüfen

→ Technische Sicherheit schaffen und KI-spezifische Anforderungen definieren

→ Restrisiken beleuchten

KI ist unumgänglich

- KI wird genutzt werden
- Mindestens: Regelung treffen und Fachwissen aufbauen
- Besser: Sensibilisierungs- und Schulungsangebot für **alle Mitarbeitenden**
- Geschäftsführung benötigt auch Beratung
- ISB hat ein **Recht** auf Gehör und Mitsprache bei Entscheidungen zu KI

Nutzung von KI verbindlich Regeln

- Nutzer informieren
- Aufklären über Arten von KI
- Klarstellung freigegebener und verbotener KI-Anwendungen
- Klare Regelungen was wo verarbeitet werden darf
- Stichprobenartige Kontrollen
- Ggf. ISO-42001 anwenden

Mitsprache bei Entscheidungen zu KI

→ Genau klären worum es geht

→ Einsatz LLM, Training einer eigenen KI und Verwendung von ausgelernter Spracherkennung sind **grundverschieden**

→ Entscheidungen **genau** prüfen

→ Tatsächlichen Nutzen genau herausarbeiten und kritisch hinterfragen

→ Tatsächliche Einführungskosten und Betriebskosten

→ Nutzerakzeptanz und Kundenakzeptanz kritisch hinterfragen

→ Zukünftige gesellschaftliche Meinung über KI miteinbeziehen

→ Widersprüche KI-Nutzung und Nachhaltigkeit betrachtet

→ Datenschutzprobleme und –risiken genau prüfen



KI – Sicherheit und Umgang mit KI

KI Systeme bewerten

KI – Red Flags

→ Folgende Szenarien geben grundsätzlich Grund zur Vorsicht:

→ KI lernt aus fremden Daten

→ KI ist essenziell für Geschäftserfolg

→ KI verarbeitet Nutzereingaben (insbesondere externe/öffentliche Nutzer)

→ KI trifft selbstständig Entscheidungen mit realen Folgen

→ KI lernt aus sensiblen Daten

→ KI verarbeitet personenbezogene Daten

KI – Rechtliche Grundlagen genau kennen

- KI ist derzeit noch kaum gesetzlich geregelt
- Große rechtliche Grauzonen und Unsicherheit
- AI-Act gibt erste Ansätze
- Red Flags beachten
- Ggf. Rechtsberatung einholen

KI – Risiken bewerten

→ Ist das System unbedenklich?

→ Fertig ausgelerntes System

→ Rein interner Einsatz

→ Keine Entscheidungsgewalt

Hier kann klassische Informationssicherheit ausreichen.

→ Ansonsten:

→ Risikoanalyse, z.B. nach BSI 200-3

→ Betrachtung typischer KI-Gefährdungen



KI – Sicherheit von KI-Systemen

KI – Was sind eigentlich meine Assets?

→ KI-Anwendung selbst

→ Trainingsdaten

→ KI-Modell (Trainingsstand der KI)

→ Betroffene Systeme

→ Die Reputation meiner Institution

KI – Daten und Qualität

→ Jedes KI-System ist maximal so gut wie seine Daten

→ Trainingsdaten

→ Validierungsdaten

→ Testdaten

→ Garbage in – Garbage out

→ Schlechte Trainingsdaten – Schlechtes Ergebnis

→ Angriff auf Trainingsdaten kann KI ruinieren

→ Daten sind wertvoll

→ Diebstahl von Trainingsdaten

KI – Notfallvorsorge für aktiv lernende Systeme

→ Falsche Lernergebnisse sind nicht einfach rückgängig zu machen



Agentic Misalignment: How LLMs could be insider threats



KI – Notfallvorsorge für aktiv lernende Systeme

- Falsche Lernergebnisse sind nicht einfach rückgängig zu machen
- AI-Unlearning ist ein aktives Forschungsfeld
 - Bislang keine einfache Lösung gefunden
- Daher: Periodische Vollsicherungen und Wiederherstellungsprüfung



KI – Fazit

KI – Mut und Vorsicht

- KI muss betrachtet werden
- KI ist nicht grundsätzlich schlecht
- Es ist nicht immer sinnvoll KI einzusetzen
- KI kann viele Dinge nicht ersetzen
- Konventionelle Betrachtung weiterhin unerlässlich
- **Aufklärung und Regelung statt Hysterie, Verbot oder Hype**

Vielen Dank

Marco Reda
SVA System Vertrieb Alexander GmbH
Borsigstraße 26
65205 Wiesbaden

+49 6122 536-0
mail@sva.de
www.sva.de

Weiterführendes Material

→ **Has Generative AI Already Peaked? – Computerphile**

- <https://www.youtube.com/watch?v=dDUC-LqVrPU>
- <https://arxiv.org/abs/2404.04125>
- Wissenschaftliche Prognosen zur Entwicklung der Leistungsfähigkeit von generativen KI-Systemen

→ **What Is a Prompt Injection Attack?**

- <https://www.youtube.com/watch?v=jrHRe9ISqqA>
- Betrachtung eines typischen Angriffs auf LLM-Systeme analog zu SQL-Injection und Social Engineering

→ **Security of AI-Systems: Fundamentals - Adversarial Deep Learning**

- https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/Security-of-AI-systems_fundamentals.pdf
- Paper zu best Practices zur Sicherheitsbetrachtung von KI-Systemen

→ **ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system**

- <https://www.iso.org/standard/42001>
- ISO-Standard zur Einführung eines Management-Systems für KI

Quellenverzeichnis eingebundener Inhalte

→ Headlines accessed on 29.01.2026:

→ Reddit

https://www.reddit.com/r/languagelearning/comments/1kxmixs/duolingos_aifirst_disaster_a_cautionary_tale_of/

→ Indatalabs Blog <https://indatalabs.com/blog/ai-cost-reduction>

→ McKinsey <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>

→ MIT NANDA https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

→ Epoch AI <https://epoch.ai/gradient-updates/can-ai-companies-become-profitable>

→ Commonfund Private Equity <https://www.commonfund.org/cf-private-equity/ai-is-redefining-how-startups-scale>

→ Wheresyouredat <https://www.wheresyoured.at/why-everybody-is-losing-money-on-ai/>

→ Grok Mechahitler Tweet posted on Reddit 09.07.2025

https://www.reddit.com/r/Destiny/comments/1lv1v7k/grok_is_now_entering_its_final_form_and_self/

→ Google cigarettes advice <https://medium.com/@m.s.rajpoot20/google-messed-up-again-the-ai-game-219b24fc1181> date unknown, article accessed 29.01.2026

→ AI as insider threat <https://www.anthropic.com/research/agent-misalignment> accessed on 29.01.2026