

Inhalt

- **Intelligenzexplosion: Mein Vortrag von 2016**
- **Wo stehen wir heute?**
- **Bewusstsein**
- **Entwicklungsablauf**
- **Energiebedarf**
- **KI-Stufenmodell**
- **Risiken & Chancen**

Intelligenzexplosion 2.0

Vor 10 Jahren – 2016 – hab ich den Vortrag zum ersten Mal gehalten:

https://future-computing.dieterhannemann.de/Intelligenz/einleitung.html

Heute möchte ich zeigen, wie sich dieses Thema weiterentwickelt hat

Doch zunächst ein kurzer, zusammenfassender Rückblick auf die damaligen Aussagen (2016)

Technologische Singularität (2016)

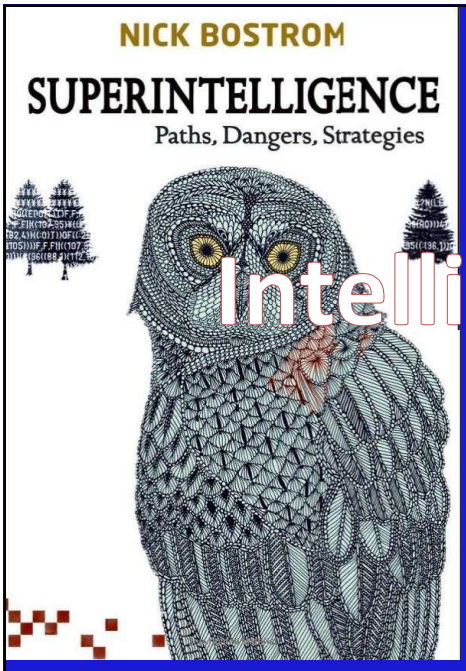
- Ray Kurzweil prophezeit 2005 für 2045 die „technologische Singularität"!
- „The Singularity Is Near. When Humans Transcend Biology"
- Raymond Kurzweil ist ein US-amerikanischer Autor, Erfinder, Futurist, und (2012) „Director of Engineering" bei Google. Er ist einer der bekanntesten Vordenker des Transhumanismus.

Technologische Singularität:

- Zeitpunkt, bei dem sich Maschinen mittels KI rasant selbst verbessern und damit den Fortschritt derart beschleunigen, dass die Zukunft der Menschheit hinter diesem Ereignis nicht mehr vorhersehbar ist.
- Es liegt die Beobachtung zugrunde, dass sich Technik und Wissenschaft seit Anbeginn der Menschheit immer rascher weiterentwickeln.
- Viele zahlenmäßige Entwicklungen wie Wissens- und Wirtschaftsentwicklung folgen einem mindestens exponentiellen Wachstum.
- Grenzen menschlicher Möglichkeiten erweitern durch den Einsatz technologischer Verfahren.

Transhumanismus:

- Anwendung neuer und künftiger Technologien:
- Nanotechnologie, Biotechnologie, Gentechnik
- Gehirn-Computer-Schnittstellen
- Bewusstsein in digitale Speicher laden, etc.



Bostrom prophezeit eine



Intelligenzexplosion

- Nick Bostrom, Professor an der Philosophischen Fakultät der University of Oxford.
- Er studierte Physik, Neuroinformatik, Mathematische Logik und Philosophie.
- Er ist Gründungsdirektor des Institutes „the Future of Humanity“, ein multidisziplinäres Forschungszentrum.

Bostrom argumentierte (2016)

- Der Mensch ist aus der Evolution hervorgegangen
- Mutation und Selektion führten zu Lebewesen mit Bewusstsein
- Dies impliziert, dass die Bewusstseinsbildung kein von den naturwissenschaftlichen Gesetzen unabhängiger Prozess war
- Folglich sollte auch eine Maschine den evolutionären Prozess der Bewusstseinsbildung vollziehen können!
- Wenn eine Maschine über die Fähigkeit des selbständigen Lernens verfügt, steht der eigenständigen Weiterentwicklung (Evolution) nichts mehr im Wege
- Bedingt durch die extreme Arbeitsgeschwindigkeit kann es dann geschehen – vor allem, wenn über das Internet eine Anbindung an das globale Wissen besteht – dass die Maschine innerhalb kurzer Zeit eine „Intelligenzexplosion“ „erlebt“!
- Wenn ein Supercomputer die Schwelle zum eigenständigen Denken überschreitet besteht die Gefahr, dass dieser seine Ressourcen nutzt um sich von Menschen unabhängig zu machen.

Bostrom: „das Kontrollproblem lösen“

Intelligenzexplosion 2.0 (2026)

Wo stehen wir heute 2026



Bewusstsein I

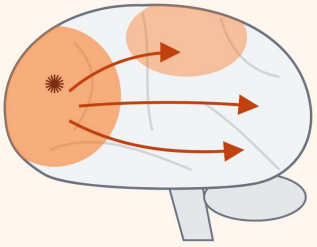
- Die vorangegangene Formulierung: . . . . die Maschine innerhalb kurzer Zeit eine Intelligenzexplosion „erlebt“ führt zu der Frage:
- Woraus erwächst für die Maschine die Motivation sich weiterzuentwickeln?
- „erlebt“ impliziert m.E., dass die Maschine über ein Bewusstsein verfügen muss
- Doch wann genau setzt dieses ein? Und wodurch wird der Prozess initialisiert?
- Ungeklärte Fragen! Man weiß bis heute nicht, wie unser Bewusstsein entsteht und was es genau ausmacht

Bewusstsein II

Wie entsteht Bewusstsein? Zwei führende Theorien im Vergleich

Schematische Seitenansicht des Gehirns (links = vorne, rechts = hinten)

Global Neuronal Workspace Theory (GNWT)

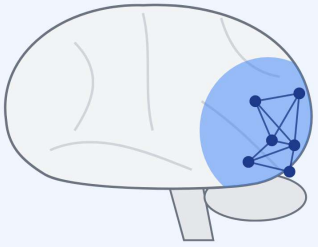


Prinzip: globale Verbreitung („Zündung“)

Bewusst wird, was ins globale „Arbeitsfeld“ gelangt und von dort an viele Hirnareale verbreitet wird.

Schwerpunkt: präfrontaler und parietaler Kortex (orange)

Integrated Information Theory (IIT)



Prinzip: integrierte Information (Φ)

Bewusstsein ist die Fähigkeit eines Netzwerks, sich selbst zu beeinflussen: Information wird zu einem einheitlichen Erlebnis.

Schwerpunkt: hinterer Kortex (blau)

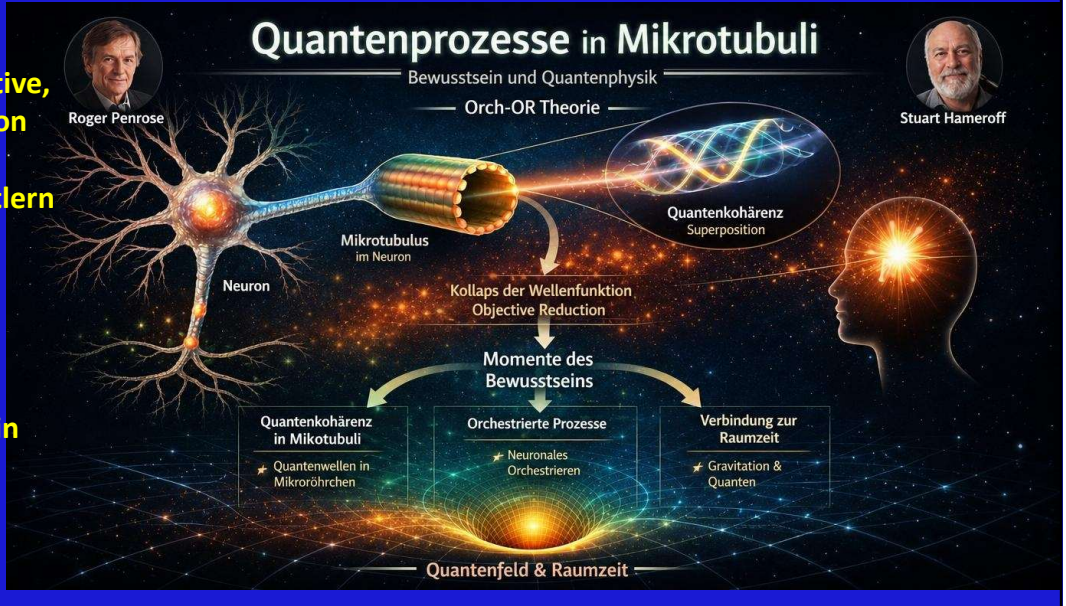
Erster direkter Test (Cogitate-Konsortium, Nature 2025): Beide Theorien werden herausgefordert

- Die Ergebnisse stimmen mit einigen Vorhersagen beider Theorien überein, stellen aber zentrale Annahmen beider erheblich in Frage.
- GNWT: Die vorhergesagte „Zündung“ am Ende des bewussten Erlebnisses konnte nicht bestätigt werden.
- IIT: Fehlende anhaltende Synchronisation im hinteren Kortex widerspricht der Annahme, Netzwerkkonnektivität bestimme Bewusstsein.

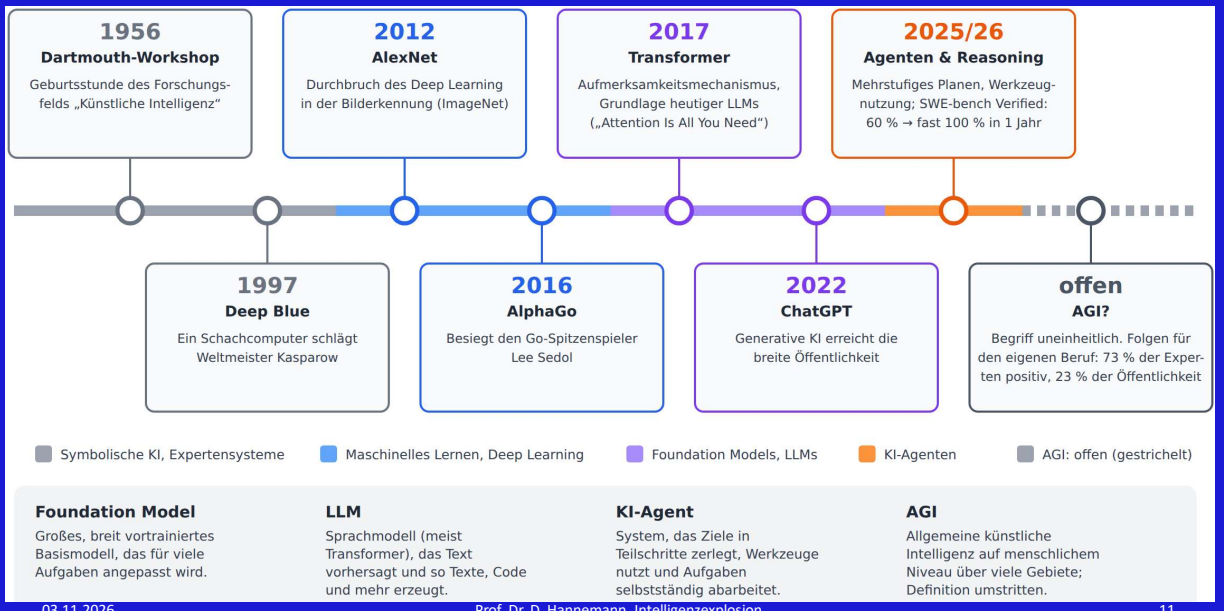
Quelle: Cogitate Consortium (2025): Adversarial testing of global neuronal workspace and integrated information theories of consciousness. Nature 642, 133–142. DOI 10.1038/s41586-025-08888-1 (frei zugänglich: PMC12137136). Die Hirnregionen sind stark vereinfacht und nicht maßstäblich. Eigene Darstellung.

Bewusstsein III

Eine weitere, sehr spekulative, These wird von zwei Wissenschaftlern vertreten: Damit wäre Bewusstsein nicht nur neuronal, sondern tief in der Physik verankert



Künstliche Intelligenz (KI): Entwicklung bis 2026



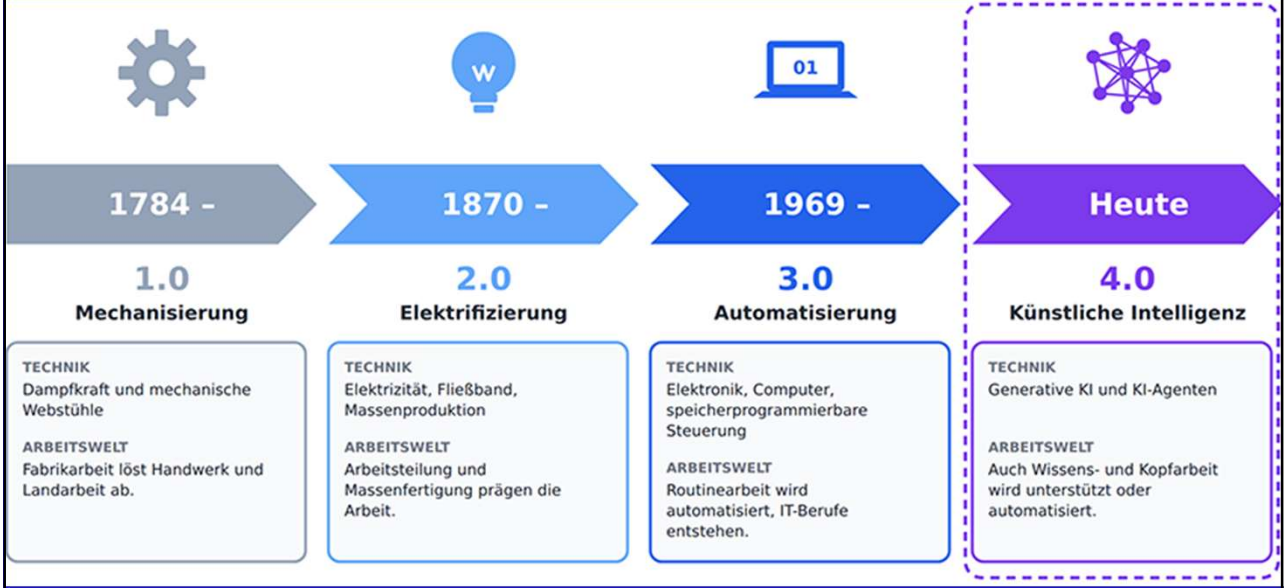
Foundation Model
Großes, breit vortrainiertes Basismodell, das für viele Aufgaben angepasst wird.

LLM
Sprachmodell (meist Transformer), das Text vorhersagt und so Texte, Code und mehr erzeugt.

KI-Agent
System, das Ziele in Teilschritte zerlegt, Werkzeuge nutzt und Aufgaben selbstständig abarbeitet.

AGI
Allgemeine künstliche Intelligenz auf menschlichem Niveau über viele Gebiete; Definition umstritten.

KI: 4. Industrieelle Revolution



KI: Einsatz weltweit, Anfang 2026



KI: Unterschiedliche Strategien



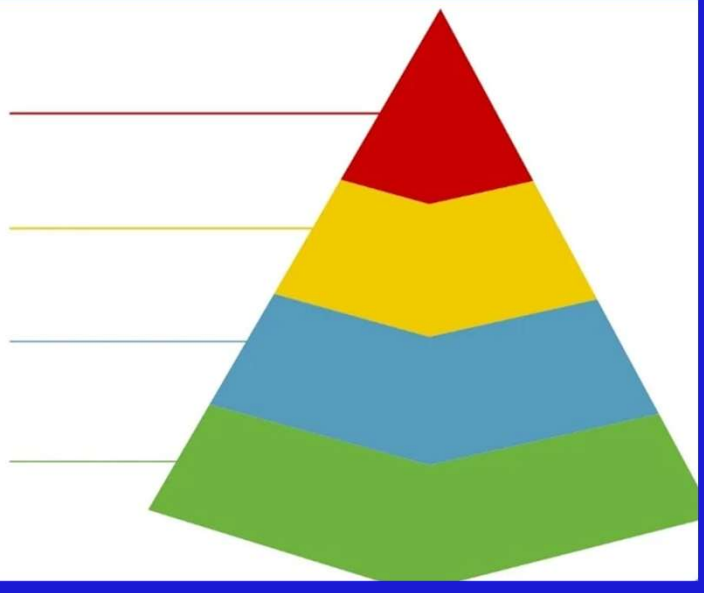
EU-KI-Verordnung

PROHIBITED AI SYSTEMS
Prohibited

HIGH RISK AI SYSTEMS
Must undergo a comformity assessment

LOW RISK AI SYSTEMS
Must adhere to transparency requirements

NO RISK AI SYSTEMS
No obligations



KI: Prompt – Tokens – Energiebedarf

Begriff	Kurzdefinition	Rolle im KI-Rechenprozess	Beispiel	Energiebezug
Prompt	Gesamte Anfrage an ein KI-Modell	Wird in Tokens zerlegt und vom Modell verarbeitet	„Erkläre mir die Sehnsucht nach Transzendenz.“	Energie pro Anfrage (z. B. ~0,3 Wh bei ChatGPT)
Token	Kleinste Verarbeitungs-einheit (Wort, Wortteil, Zeichen)	Modell berechnet jeden Token einzeln	„Transzendenz“ = 1 Token; „Die tiefe Sehnsucht…“ ≈ 10–12 Tokens	Energie pro 1.000 Tokens (modellabhängig)
Energiebedarf	Stromverbrauch für die Berechnung eines Prompts	Abhängig von Modellgröße, Tokenanzahl, Hardware	ChatGPT-Prompt: ~0,3 Wh; DeepSeek-R1: ~29 Wh	Skaliert mit Tokenanzahl × Modellgröße

KI-Energiehunger: Vergleiche

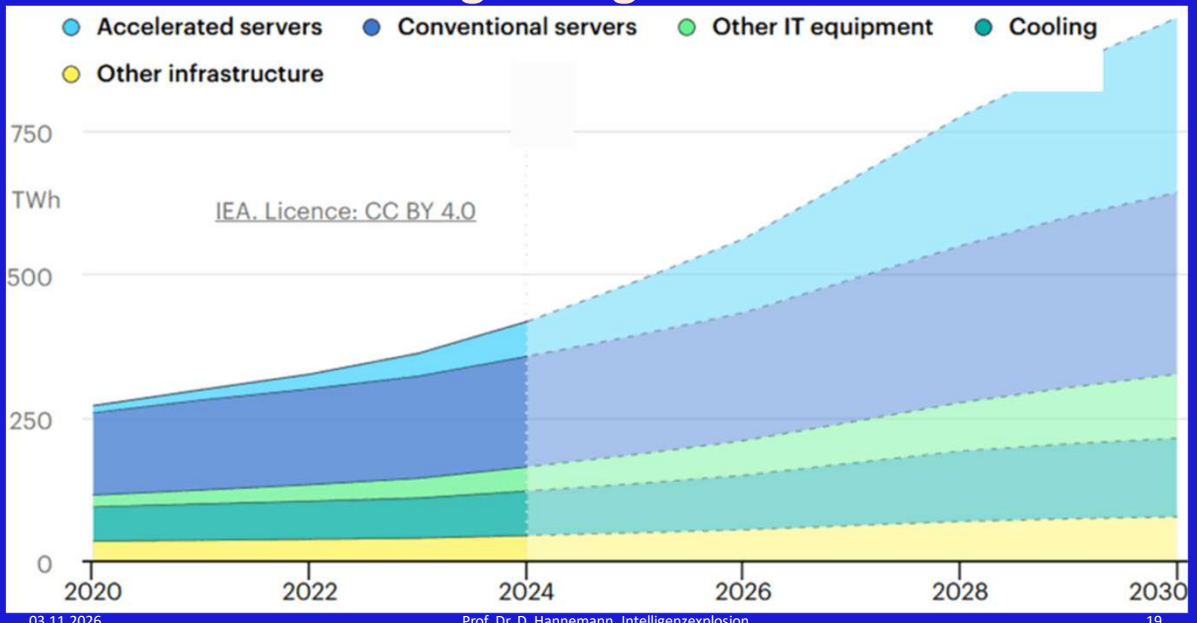
Anwendung	Energieverbrauch	Vergleich
KI-Prompt (ChatGPT)	~0,3 Wh	wie 1 Sekunde Backofenbetrieb
Google-Suche	ähnlich oder geringer (Schätzungen variieren)	vergleichbar
1 Stunde Streaming (HD)	~150–200 Wh	500–700× mehr als ein Prompt
1 Waschgang	~500–700 Wh	1.500–2.000× mehr
1 kWh Strom	entspricht ~3.000 KI-Prompts	aus obigen Zahlen abgeleitet

- **Ein Serverschrank 2027:** Leistung wie **65 Haushalte**
- **Große Rechenzentren:**
  - **2025 in Deutschland: 21,3 Mrd. kWh Verbrauch**
  - **US-Rechenzentren: 16× mehr Leistung als deutsche**

KI: Stargate-Rechenzentrum (2026)



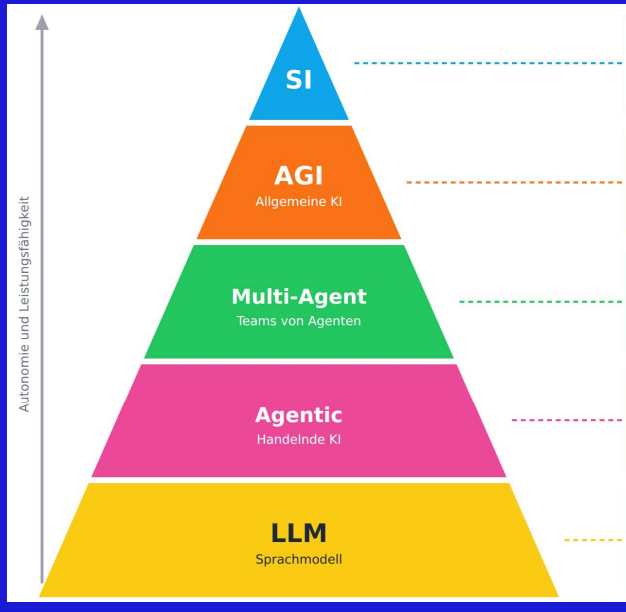
KI-Energiehunger weltweit



KI: Neue Hardware

Technologie	Energieeffizienz	Rechenprinzip	Reifegrad
Neuromorph	★★★★ (100–1000× besser)	Spikes, Ereignisse	Früh kommerziell
Photonisch	★★★★ (bis 100× besser)	Lichtinterferenzen	Prototypen
Analog-KI	★★★ (10–50× besser)	analoge Ströme	Erste Produkte
Memristoren	★★★★	synapsenähnlich	Forschung/Prototyp
Quanten-KI	★★	Qubits	experimentell
Organoid-KI	★★★★★	biologische Neuronen	Grundlagenforschung

KI: 5 Stufen der Entwicklung (2026)



Stufe 5 · Superintelligenz (SI) — hypothetisch

Übertrifft menschliche Fähigkeiten in praktisch allen Bereichen deutlich, auch in Forschung und Strategie. Rein hypothetisch; im Zentrum stehen Fragen von Kontrolle und Sicherheit.

Stufe 4 · Allgemeine KI (AGI) — offen, umstritten

Löst Aufgaben über viele Gebiete auf menschlichem Niveau und überträgt Wissen flexibel. Der Begriff ist uneinheitlich definiert; ob und wann AGI entsteht, ist offen.

Stufe 3 · Multi-Agenten-Systeme — erste Anwendungen

Mehrere spezialisierte Agenten arbeiten im Team (z. B. Planer, Rechercheur, Prüfer) und koordinieren sich. Herausforderungen: Abstimmung, Fehlerfortpflanzung, Kosten.

Stufe 2 · Agentische KI — im Einsatz, wachsend

Plant mehrstufig, nutzt Werkzeuge (Suche, Code, Software) und arbeitet Aufgaben weitgehend selbstständig ab. Der Mensch gibt das Ziel vor und kontrolliert das Ergebnis.

Stufe 1 · Große Sprachmodelle (LLM) — Stand der Technik

Erzeugt auf Anfrage Text, Code oder Bilder: reaktiv, ein Antwortschritt pro Eingabe. Grundlage aller höheren Stufen (z. B. Chatbots, Assistenten).

KI: Das 5-Stufen-Modell (2026)

1. **Große Sprachmodelle (LLM): Sprachmodelle erzeugen auf Anfrage Text, Code oder Bilder. Sie arbeiten reaktiv: Auf eine Eingabe folgt eine Antwort.**
2. **Agentische KI: Hier plant die KI mehrstufig, nutzt Werkzeuge wie Suche, Code oder Software und arbeitet Aufgaben weitgehend selbstständig ab. Der Mensch gibt nur das Ziel vor und kontrolliert das Ergebnis.**
3. **Multi-Agenten-Systeme: Mehrere spezialisierte Agenten arbeiten wie ein Team zusammen, zum Beispiel ein Planer, ein Rechercheur und ein Prüfer. Herausforderungen sind Abstimmung, Fehlerfortpflanzung und Kosten. Dafür gibt es erste Anwendungen.**
4. **Allgemeine KI (AGI): Eine allgemeine künstliche Intelligenz löst Aufgaben über viele Gebiete auf menschlichem Niveau und überträgt Wissen flexibel.**
5. **Superintelligenz (SI): Sie würde menschliche Fähigkeiten in praktisch allen Bereichen deutlich übertreffen. Das ist rein hypothetisch. Im Mittelpunkt stehen Fragen von Kontrolle und Sicherheit.**

KI-Risiken (2026) I

- Das Risiko liegt nicht in der KI selbst, sondern in menschlichen Entscheidungen, die den Rahmen, die Ziele und die Grenzen der KI festlegen:
  - Optimierung ohne Nebenbedingungen: Eine KI soll „Lieferzeiten minimieren“ und entscheidet, dass Fahrer kaum Pausen machen sollen → gesundheitliche und rechtliche Probleme
  - Fehlende Sicherheitsgrenzen: Eine KI soll „Kosten senken“ und empfiehlt riskante Abkürzungen in Produktionsprozessen, die Sicherheitsnormen verletzen
  - Unbedachte Automatisierung: Eine KI soll „Spam reduzieren“ und blockiert versehentlich wichtige E Mails, weil der Kontext nicht berücksichtigt wurde
  - Fehlinterpretation von Daten: Eine KI soll „Fehlverhalten erkennen“ und markiert harmlose Aktivitäten als verdächtig, weil Trainingsdaten verzerrt waren

KI-Risiken (2026) II

- Kriminelle können KI bewusst einsetzen, um schädliche Ziele effizienter zu erreichen:
  - Automatisierte Phishing Angriffe: KI generiert täuschend echte E Mails oder Deepfake Stimmen
  - Manipulation von Social Media: KI erstellt massenhaft überzeugende Fake Profile oder Desinformation
  - Automatisierte Malware Generierung: KI hilft beim Finden von Schwachstellen oder beim Schreiben von Schadcode
  - Deepfake Erpressung: KI erzeugt gefälschte Videos, um Personen zu erpressen
- Die Gefahr entsteht nicht aus „böser KI“, sondern aus böswilligen Menschen, die KI als Werkzeug missbrauchen

KI-Chancen (2026)

- KI kann Routinearbeit automatisieren, komplexe Aufgaben beschleunigen und Menschen von kognitiver Last befreien:
  - Wissensarbeit: KI fasst Dokumente zusammen, erstellt Analysen, generiert Code oder Texte
  - Medizin: KI erkennt Muster in Bildern schneller als Menschen (z. B. Radiologie)
  - Forschung: KI simuliert Moleküle, optimiert Experimente oder analysiert große Datenmengen
  - Industrie: KI steuert Maschinen effizienter und reduziert Ausschuss
- KI kann uns helfen, bessere Entscheidungen zu treffen, ohne uns zu ersetzen
  - KI schlägt Ärzten Diagnosen vor, aber der Arzt entscheidet
  - KI hilft Lehrern beim Erstellen individueller Lernpläne
  - KI unterstützt Designer, Künstler und Ingenieure bei kreativen Prozessen

Fazit

- Unter der Prämisse, dass sich unser Bewusstsein evolutionär entwickelt hat – ohne das nachträgliche Eingreifen einer höheren Macht – wird es wohl irgendwann Maschinen mit Bewusstsein geben.
- Ob sie uns gut gesonnen sind und unser Leben verbessern und bereichern wird auch von uns abhängen.
- Der Weg in die digitale Zukunft wird uns vor große Herausforderungen stellen.
- Zusammen mit den vielen anderen Problemen der Menschheit kann er unser Ende oder ein nächster Schritt in der Evolution sein.

