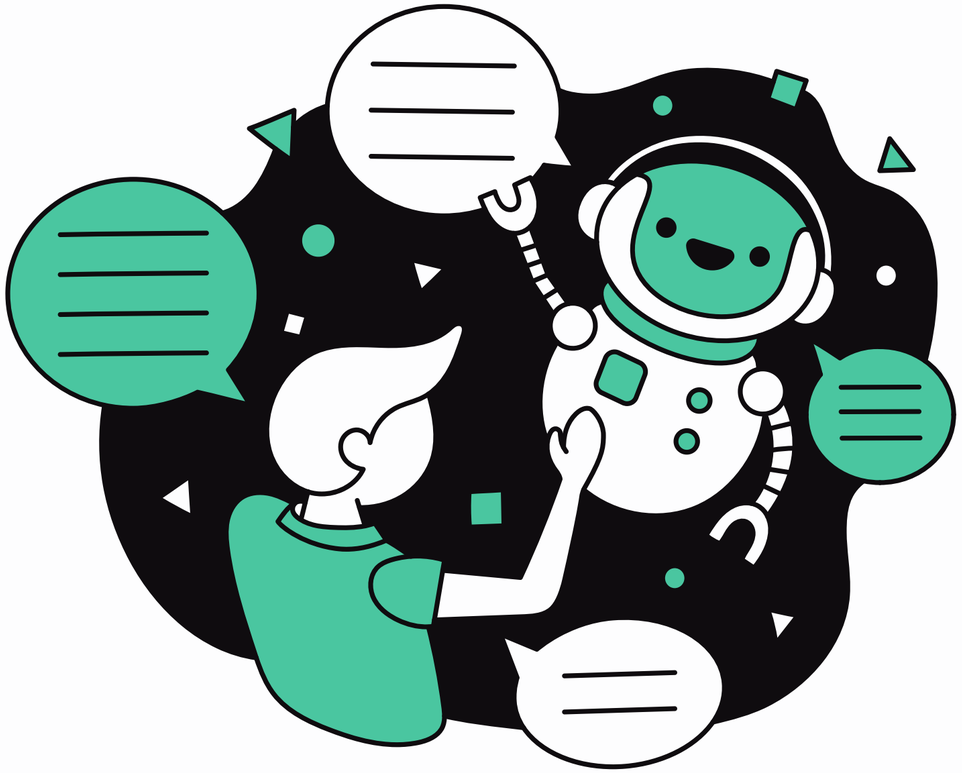
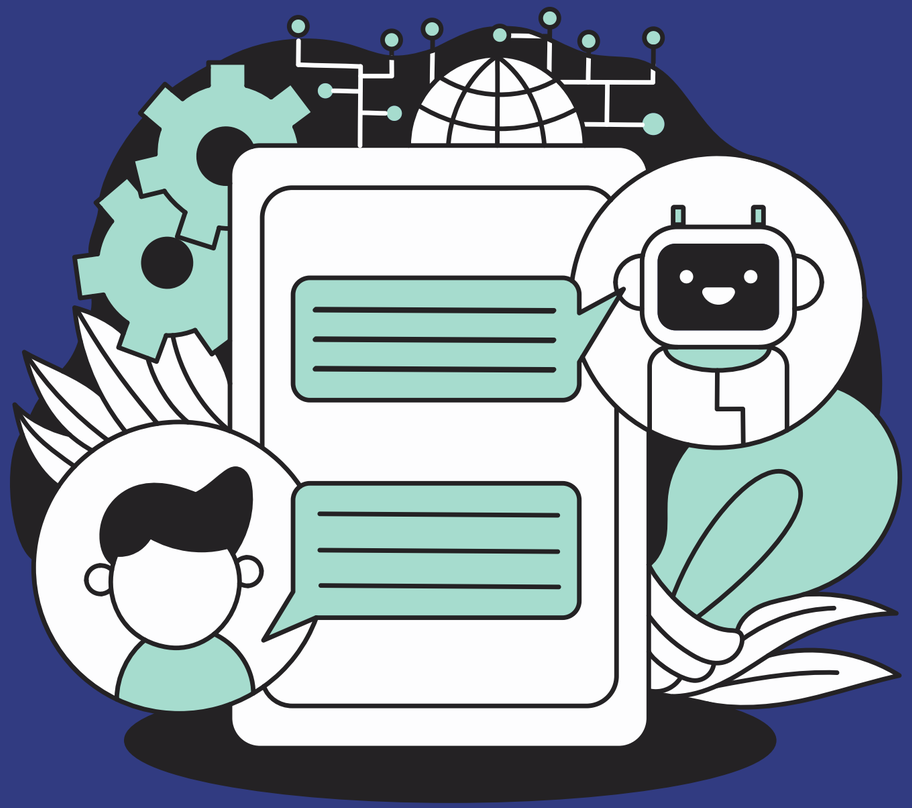


TAG 5

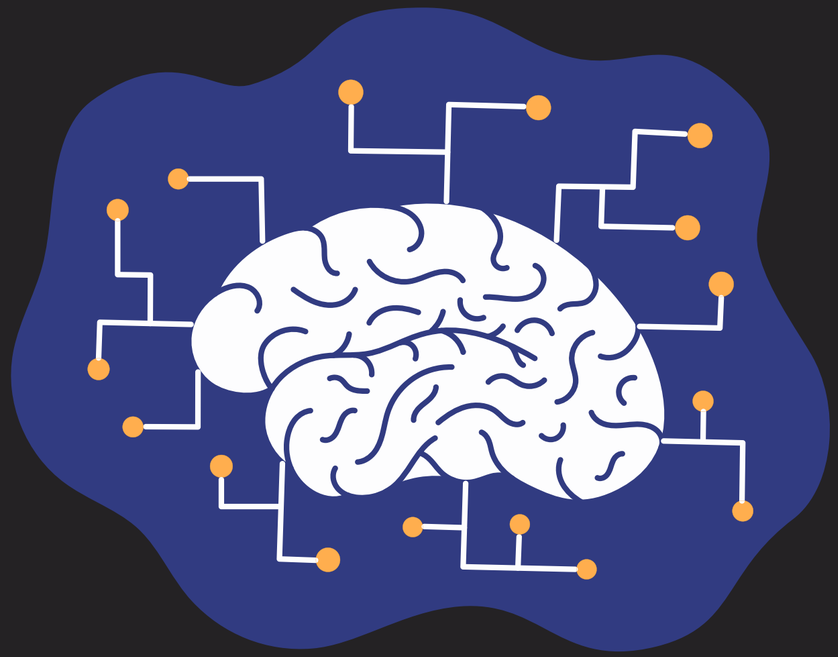
KI-Bildungsurlaub

Regulierung, Verantwortlichkeiten, Herausforderungen

Inhalt



- Regulierung - DSGVO, EU AI ACT
- Kennzeichnungspflichten
- Urheberrecht
- Hacking & Schadenspotential
- Herausforderungen & Ethik



Regulierung

Sicher mit KI arbeiten

Datenschutz, EU AI-ACT, Verantwortung im Unternehmen, Urheberrecht

Datenschutz & KI-Regulierung

DSGVO

Schutz personenbezogener Daten

EU AI Act

Sicherheit & Transparenz bei KI-Systemen

Cloud Act

US-Behörden dürfen auf gespeicherte Daten von US-Unternehmen zuzugreifen, auch wenn diese im Ausland gespeichert sind.

EU AI Act

Risikostufen von KI-Systemen im Überblick

Unvertretbares Risiko

Verbotene KI-Anwendungen, die Grundrechte, Sicherheit oder Demokratie gefährden.

BEISPIELE

- Soziales Scoring
- Echtzeit-Gesichtserkennung im öffentlichen Raum

Hohes Risiko

Streng regulierte KI-Systeme, die über Menschen oder kritische Prozesse entscheiden. Erfordern Dokumentation und menschliche Aufsicht.

BEISPIELE

- KI im Bewerbungsverfahren
- Medizinische Diagnose
- KI in Justiz, Polizei, Kreditvergabe

Begrenztes Risiko

Erhöhte Transparenzpflichten – Nutzer müssen erkennen, dass sie mit einer KI interagieren.

BEISPIELE

- Chatbots & Sprachassistenten
- Generative KI (Texte, Bilder, Videos)

Minimales oder kein Risiko

Unbedenkliche Anwendungen, die keine spezifischen Risiken für Rechte oder Sicherheit bergen.

BEISPIELE

- Spamfilter & Rechtschreibkorrekturen
- Produktvorschläge
- Navigationshilfen

EU AI Act im Alltag

Pflichten für Unternehmen & Organisationen, die KI nutzen

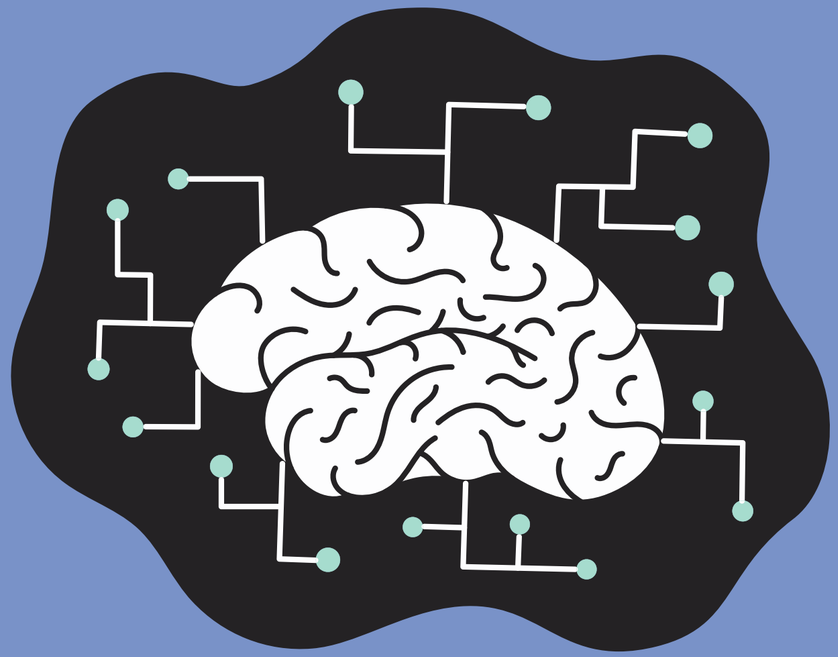
Transparenz:
Kennzeichnungspflicht bei Texten, Bildern, Chatbots

Menschliche Kontrolle:
Ergebnisse der KI müssen überprüft werden

Dokumentation:
Nachvollziehbare Aufzeichnungen über KI-Einsatz führen

KI-Grundbildung:
Mitarbeitende, die mit KI arbeiten, müssen geschult werden.

Keine automatisierten Entscheidungen



Transparenz

Kennzeichnungspflichten

nach Artikel 50, EU AI ACT ab Aug. 2026

Wie müssen KI-generierte Inhalte (Texte, Bilder, Audios, Videos, usw.) gekennzeichnet werden?

Kennzeichnungspflicht nach EU AI-ACT, Art. 50

Seit dem 2. August 2026 gelten neue Transparenzpflichten. **Sie betreffen vor allem beruflich oder geschäftlich eingesetzte KI.**

Nicht alle mit ChatGPT, Gemini oder einer anderen KI erstellten Inhalte müssen sichtbar gekennzeichnet werden.



Kennzeichnungspflicht nach EU AI-ACT, Art. 50

Kennzeichnungspflicht für Texte

KI-generierte oder manipulierte Texte müssen gekennzeichnet werden, wenn sie öffentliches Interesse betreffen z.B.: Politik, Gesundheit oder Umwelt

Ausnahme: Keine Kennzeichnung bei Texten, wenn sie redaktionell geprüft und menschlich verantwortet sind.

Kennzeichnungspflicht für andere Medien

KI-Inhalte wie Bilder, Stimmen und Videos müssen gekennzeichnet werden, wenn sie täuschend echt wirken.

Was bedeutet das im Alltag?

Keine Kennzeichnungspflicht	Kennzeichnung erforderlich
E-Mails und private Nachrichten, Bewerbungsunterlagen	Ungeprüfter KI-Artikel über Politik, Gesundheit oder Umwelt (öffentliche Interesse)
Formatierung, KI-Übersetzungen, Transkription	Automatisch veröffentlichte KI-Zusammenfassung einer Nachricht
Interne Dokumente und geschlossene Kursunterlagen	Ungeprüfte Zusammenfassungen oder Umschreibungen bei Themen von öffentlicher Interesse
Fachlich geprüfter und selbst verantworteter Artikel oder Meldung	Realistisch erzeugte Stimme oder Bild einer Person, inkl. eigene KI-Nachbildung
Rechtschreibkorrektur und kleine sprachliche Verbesserungen	Geänderter Gesichtsausdruck oder Kleidung, Austausch oder Hinzufügen von Personen und Gegenständen in Bildern *
Kleinere technische Bildkorrekturen	Täuschend echtes Bild oder Video eines nicht stattgefundenen Ereignisses oder ein realistisch manipulierter Schauplatz *

Geeignete Formulierungen für Kennzeichnung

Beispiele:

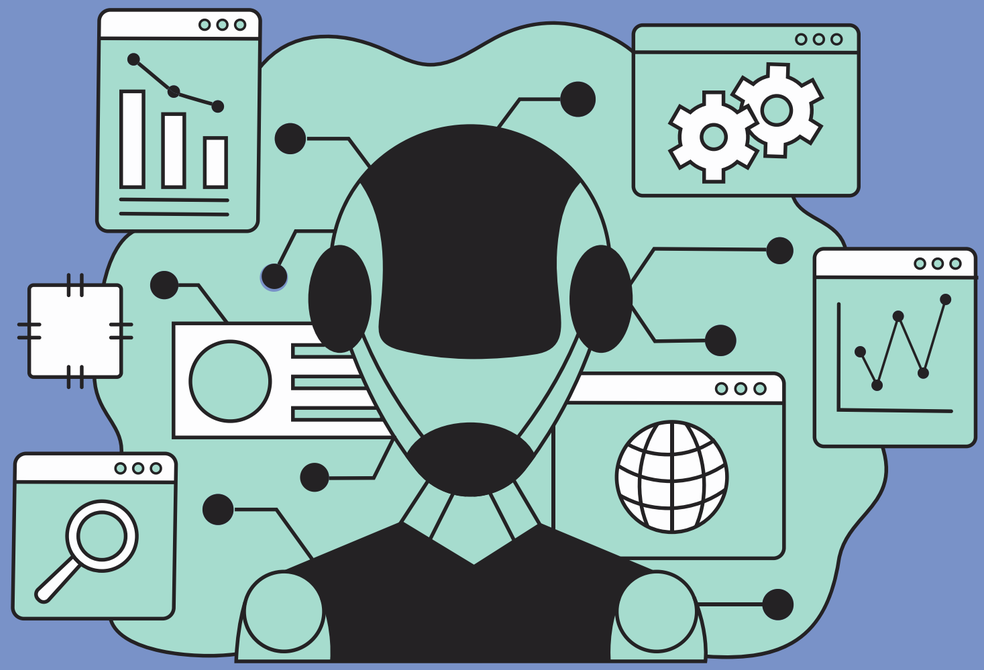
- „Dieser Text wurde vollständig mit KI erzeugt.“
- „Dieser Text wurde teilweise mit KI erstellt oder bearbeitet.“
- „Die folgende Zusammenfassung wurde automatisiert mit KI erstellt und nicht redaktionell geprüft.“
- „Dieser Inhalt wurde mit künstlicher Intelligenz verändert.“
- „Bei der Erstellung wurde KI unterstützend eingesetzt. Inhaltlich geprüft und redaktionell verantwortet von .....“

Platzierung des Hinweises:

- klar und deutlich sichtbar, z.B. oberhalb des Textes
- Hinweis muss barrierefrei sein
- bei der ersten Begegnung mit dem Inhalt erscheinen
- darf nicht in AGB oder Untermenüs versteckt sein

DATENSCHUTZ NACH DSGVO

Ziel:
Vertrauen,
Fairness,
Verantwortung



DATENSCHUTZ:
GRUNDPRINZIPIEN DER DSGVO



Personenbezogene Daten (Art. 4 DSGVO):

Alles, was eine Person identifizierbar macht
Name, E-Mail, IP-Adresse, Standortdaten, Gesundheitsdaten, Religion, Stimme u.a.

Rechtsgrundlagen (Art. 6 DSGVO):

Ohne eine der folgenden ist die Verarbeitung verboten:

- Einwilligung - freiwillig, informiert, widerrufbar
- Vertragserfüllung - z. B. KI-Transkription für einen Auftrag
- Berechtigtes Interesse - z. B. interne Datenanalyse – Betroffenenrechten!
- Gesetzliche Pflicht - z. B. behördliche Anfragen

DSGVO IM ALLTAG



- Schutz persönlicher Daten
- Datenminimierung – nur erheben, was nötig ist
- Daten sind anonymisiert oder pseudonymisiert
- Zweckbindung – Daten nicht zweckentfremden
- Transparenz & Einwilligung – Betroffene müssen wissen, was passiert
- Sicherheit & Zugriffsschutz - Passwörter, Verschlüsselung, Rollen
- Speicherdauer & Löschung
- Gemeinsame Verantwortlichkeit bei Kooperationspartnern
- Meldepflicht bei Datenpannen (max. nach 72 Stunden)

AVV – Wann ist er notwendig?

Einfach erklärt für den Einsatz externer Dienstleister und KI-Tools



Was ist ein AVV?

Ein Auftragsverarbeitungsvertrag (AVV) ist ein Vertrag nach Art. 28 DSGVO. Er ist notwendig, wenn ein externer Dienstleister personenbezogene Daten im Auftrag und nach Weisung eines Unternehmens verarbeitet.

Prüffrage

1 Werden personenbezogene Daten verarbeitet?

- **Nein** → Kein AVV erforderlich
- **Ja** → **2** Verarbeitet ein externer Anbieter die Daten im Auftrag und nach Weisung?
  - **Ja** → AVV notwendig
  - **Nein** → Oft kein AVV – aber andere DSGVO-Pflichten prüfen

Typische Fälle: AVV notwendig

- Hosting oder Cloud-Dienst mit Zugriff auf Kundendaten
- Newsletter-Tool oder CRM im Auftrag
- Lohnabrechnung durch externen Dienstleister
- KI-Tool verarbeitet personenbezogene Daten für Ihr Unternehmen

Typische Fälle: meist kein AVV

- Es werden keine personenbezogenen Daten verarbeitet
- Es werden nur anonymisierte Daten genutzt
- Der Anbieter verarbeitet Daten für eigene Zwecke
- Es geht nur um öffentliche, nicht personenbezogene Informationen

Wichtig

Ein AVV allein reicht nicht aus. Zusätzlich prüfen: Rechtsgrundlage, Datensicherheit, mögliche Drittlandübermittlung und den konkreten Einsatzzweck.

Verantwortlichkeit im Unternehmen

Grundsätzlich verantwortlich?

Das Unternehmen (juristische Person) ist Verantwortlicher im Sinne der DSGVO.

Es entscheidet über Zweck („Warum verarbeiten wir Daten?“) und Mittel („Wie tun wir das?“).

Typische Rollen

Geschäftsführung / Leitung: trägt die Gesamtverantwortung

Datenschutzbeauftragter: berät, überwacht, schult

IT / Fachabteilungen: setzen technisches um

Projektleitungen: stellen sicher, dass Prozesse und Tools DSGVO- & AI-Act-konform arbeiten.

Verantwortlichkeit im Unternehmen

Mitarbeiter-Verantwortung

- Datenschutz-Regeln befolgen
- Keine Daten an unbefugte Dritte weitergeben oder in unzulässige Tools eingeben
- Verstöße oder Datenpannen sofort melden
- Fact-Checking, Quellen prüfen



Datenschutz ist keine IT-Aufgabe, es ist Teamarbeit.

Jeder, der mit Daten umgeht, trägt Verantwortung.

Sanktionen bei Vertöße

Urteile

Bank in Berlin:
Automatisierte Entscheidung bei Kreditkartenantrag ohne menschliches Eingreifen → Bußgeld: 300,000 €

Telekommunikationsanbieter in Spanien:
Nutzte KI zur Analyse von Kundengesprächen ohne Rechtsgrundlage → Bußgeld: 1,2 Millionen €

Unternehmen in Hamburg:
Speicherfristen nicht beachtet / unzureichende Daten-Löschung → Bußgeld: 900,000 €

DSGVO

Bußgelder bis zu 20 Mio. € oder 4 % des Jahresumsatzes

EU AI Act

Pflichten und Sanktionen ab 2. August 2025

Bußgelder bis zu 35 Mio. € oder 7 % des Jahresumsatzes

Konsequenzen

bei Non-Compliance, Herausgabe von Fehlinformationen

Rechtliche Risiken

- Bußgelder
- Persönliche Haftung der Geschäftsführung

Betriebswirtschaftliche Risiken

- Vertragsstrafen durch Partner (Kündigungen, Schadensersatz)
- Reputations- und Vertrauensverlust durch Medienberichte
- Wettbewerbsnachteile: Ausschluss von Ausschreibungen

Technische & operative Risiken

- Betriebsunterbrechungen: Behörden stoppen Datenverarbeitung bei Verstößen / Projektstopp
- Datenverlust/Kompromittierung durch unverschlüsselte KI-Tools → Zugriff durch Dritte

EU AI Act → Spezifische Risiken

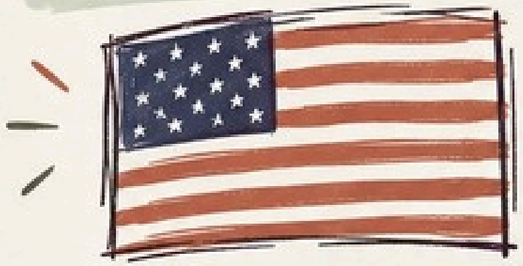
Nutzung nicht-konformer KI-Systeme

Fehlende Kennzeichnung von KI-Inhalten

Fehlende Dokumentation

Das KI-Dilemma

US-KI nutzen



- ✓ Schnell
- ✓ Leistungsstark
- ✓ Günstig

- ✗ Abhängig
- ✗ Erpressbar
- ✗ Stecker kann gezogen werden
- ✗ Datenschutz



Eigene KI bauen



- ✓ Unabhängig
- ✓ Kontrolle
- ✓ Souverän
- ✓ Datenschutz

- ✗ Teuer
- ✗ Langsamer
- ✗ Schwächer (noch)

Was ist der richtige Weg?

URHEBERRECHT

in Bezug auf
KI-erzeugte Inhalte



Urheberrecht

Vollständig von KI erzeugte Inhalte sind NICHT urheberrechtlich geschützt

Die Frage ist:

- ist eine menschliche, kreative Schöpfungsleistung erkennbar?

- Wenn ein Mensch den KI-Inhalt substantiell verändert, kann Urheberrecht entstehen.

Dies gilt für Bilder, Texte, Musik, Videos und andere kreative Werke.

Urheberrecht

Grauzone!

Was zählt als „substantielle menschliche Bearbeitung“?

Gerichte müssen dies noch klären.

Gerichte betrachten typischerweise:

- Wie detailliert war die kreative Steuerung?
- Wurden zahlreiche kreative Entscheidungen vom Menschen getroffen?
- Wurde das Ergebnis anschließend wesentlich bearbeitet?
- War die KI nur ein Werkzeug oder der eigentliche Gestalter?

Urheberrecht

Achtung bei kommerzieller Nutzung

„Nicht geschützt“ bedeutet nicht automatisch „risikofrei nutzbar“

- Markenrechte
- Persönlichkeitsrechte
- Vertragsbedingungen der KI-Anbieter
- Trainingsdaten

können trotzdem rechtliche Probleme verursachen.

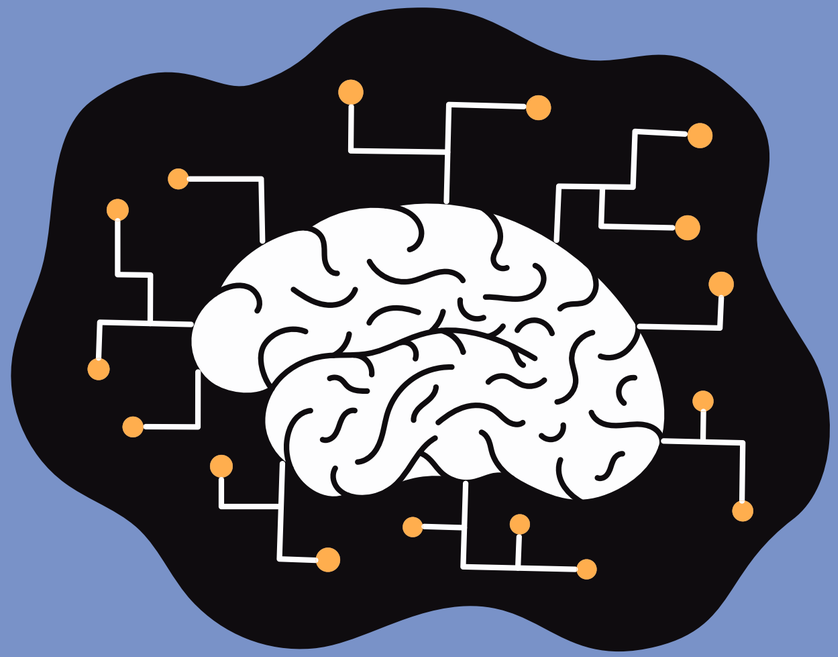
Wichtig!

AGB

- (Nutzungsrechte) der KI-Tools prüfen!

- Stark nachbearbeiten, um Urheberrecht zu erlangen.

Herausforderungen und ethische Bedenken



Wohin führt die rasante Entwicklung von KI?

Herausforderungen und Risiken der Künstlichen Intelligenz



HACKING UND SICHERHEITSBEDROHUNGEN



ARBEITSPLATZVERLUST DURCH KI-TECHNOLOGIE



AUSWIRKUNG AUF UMWELT: RESSOURCEN & NACHHALTIGKEIT



ABHÄNGIGKEIT VON TECHNOLOGIE

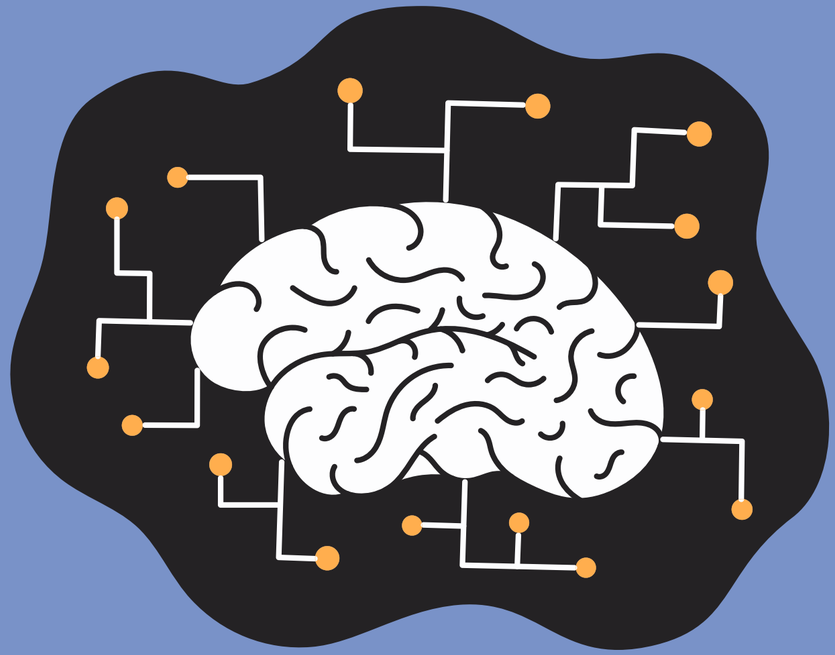


Verlust an Kreativität
Innovationskraft,
selbständiges Denken,
kognitiven Fähigkeiten



AGI oder ASI?

KI AUF DEM NIVEAU MENSCHLICHER INTELLIGENZ UND DARÜBER HINAUS?



SICHERHEITSBEDROHUNGEN

Betrugsmaschen mit Künstlicher Intelligenz

Fünf typische Angriffsmuster



Identitätsdiebstahl

Stimmenimitate per KI, gefälschte Videos & Fotos von Personen



Manipulierte Videos & Fotos

Gefälschte KI-Videos & Fotos, Deepfakes



Phishing-Mails

Überzeugendere Betrugs-E-Mails zum Entlocken von Passwörtern



Psychotricks

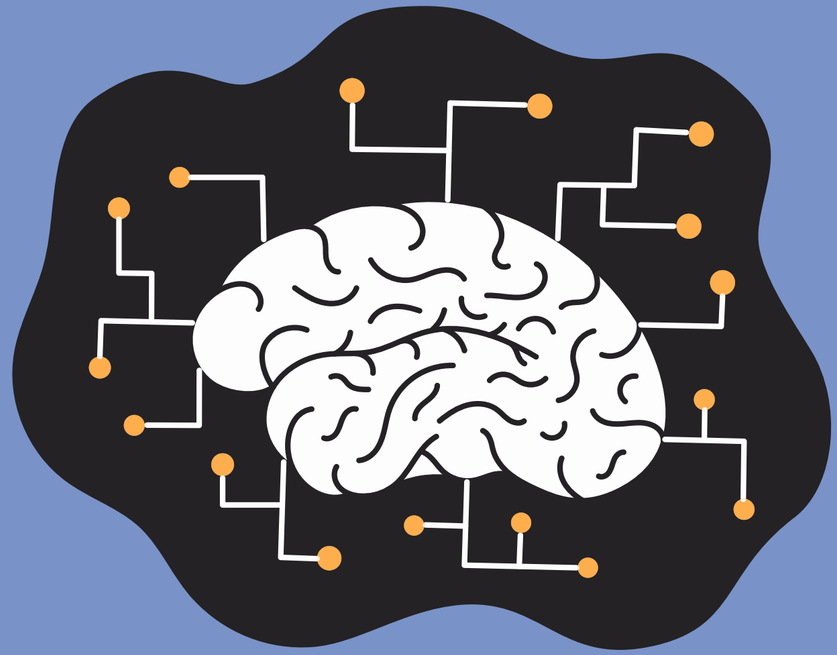
Druck & Manipulation



Manipulierte Links

Links zu Schadsoftware durch KI-Bots

DEEPFAKES



DEEPFAKES

Gefälschte Bilder, Videos und Audios

- Realistische Optik und Klang
- Leicht zu erstellen, schwer zu erkennen
- Zielen auf reale Personen, Institutionen und Ereignisse
- Zerstören Leben und Ruf
- Untergraben demokratische Institutionen
- Desinformationen, Verleumdung, Personen diskreditieren, Erpressung, politische Manipulation

Methoden

- Gesichtstausch
- Modelltraining
- Stimmenklone
- Synthetische Personen

Deepfake Videos

Gefälschtes Video von Selenskyj

Auf Facebook

Ein Video, in dem der ukrainische Präsident Selenskyj angeblich zur Kapitulation seiner Streitkräfte aufruft



Deepfakes

Real oder Fake?



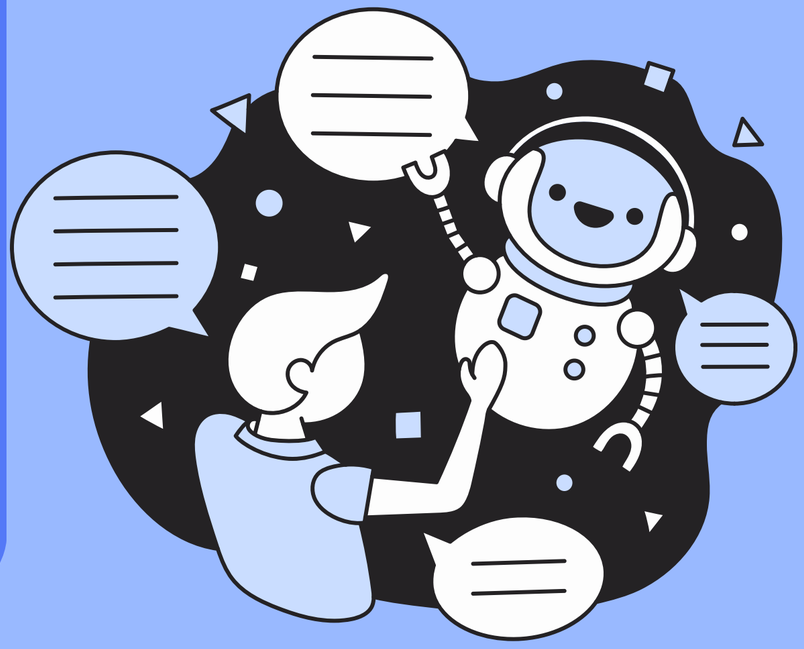
Deepfakes

Real oder Fake?

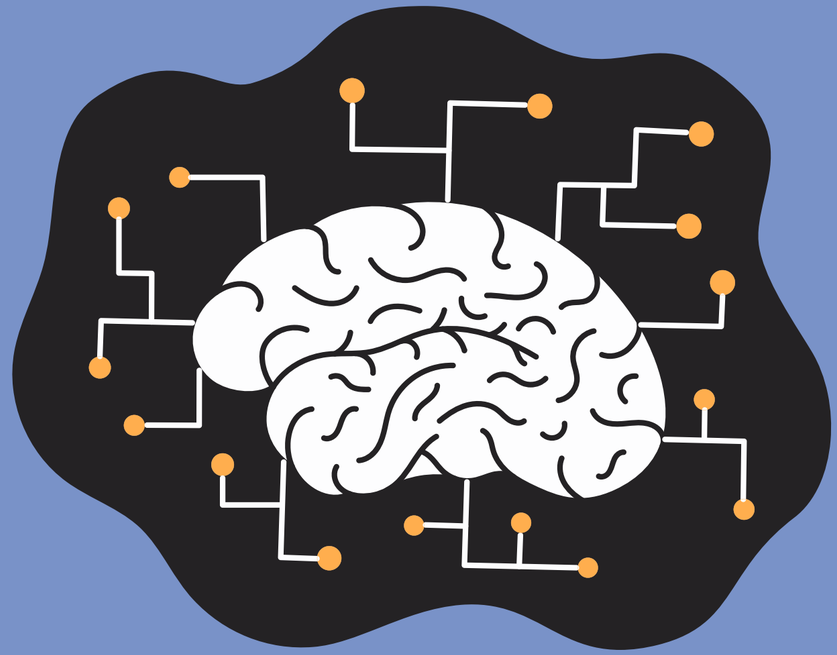


Positive Nutzung von Deepfakes

- Virtuelle Tutoren
- Kunden-Avatare
- Stunt-Doubles
- Historische Figuren
- **Barrierefreiheit**: Gebärdensprach-Avatare
- **Notfalltraining**: Realistische Krisensimulationen (z. B. Geiselnahmen oder Brände) für Sicherheitskräfte.
- **Medizinische Schulung**: Simulation von seltenen Krankheitsbildern oder OP-Situationen für Ärzte.



HACKING



Prompt Injection

ZERO-CLICK-ATTACK: FALLSTUDIE ECHO-LEAK

Wie vertrauliche Daten aus M365 Copilot automatisch ausgelesen werden können – **ohne Nutzerinteraktion**.

Angriff: E-Mail mit versteckter „**Prompt Injection**“ wird gesendet.

Sichtbarer Text + unsichtbarer bösartiger Prompt

Verarbeitung: KI-Agent (Copilot) liest und verarbeitet die Nachricht automatisch.

Inkl. versteckte + unsichtbare Anweisung

Ausführung & Datenabfluss: Agent führt verborgene Anweisung aus und leitet vertrauliche Daten weiter *(Kontonummern, Passwörter, interne Notizen)*.

Schlüsselerkenntnisse:

- Kein Nutzerverhalten erforderlich
- Systemische Schwachstelle im KI-Agenten
- Plattformübergreifende Gefahr
- Technische und prozessuale Schutzmaßnahmen erforderlich

Jailbreaking (Umgehung von Sicherheitsvorgaben)

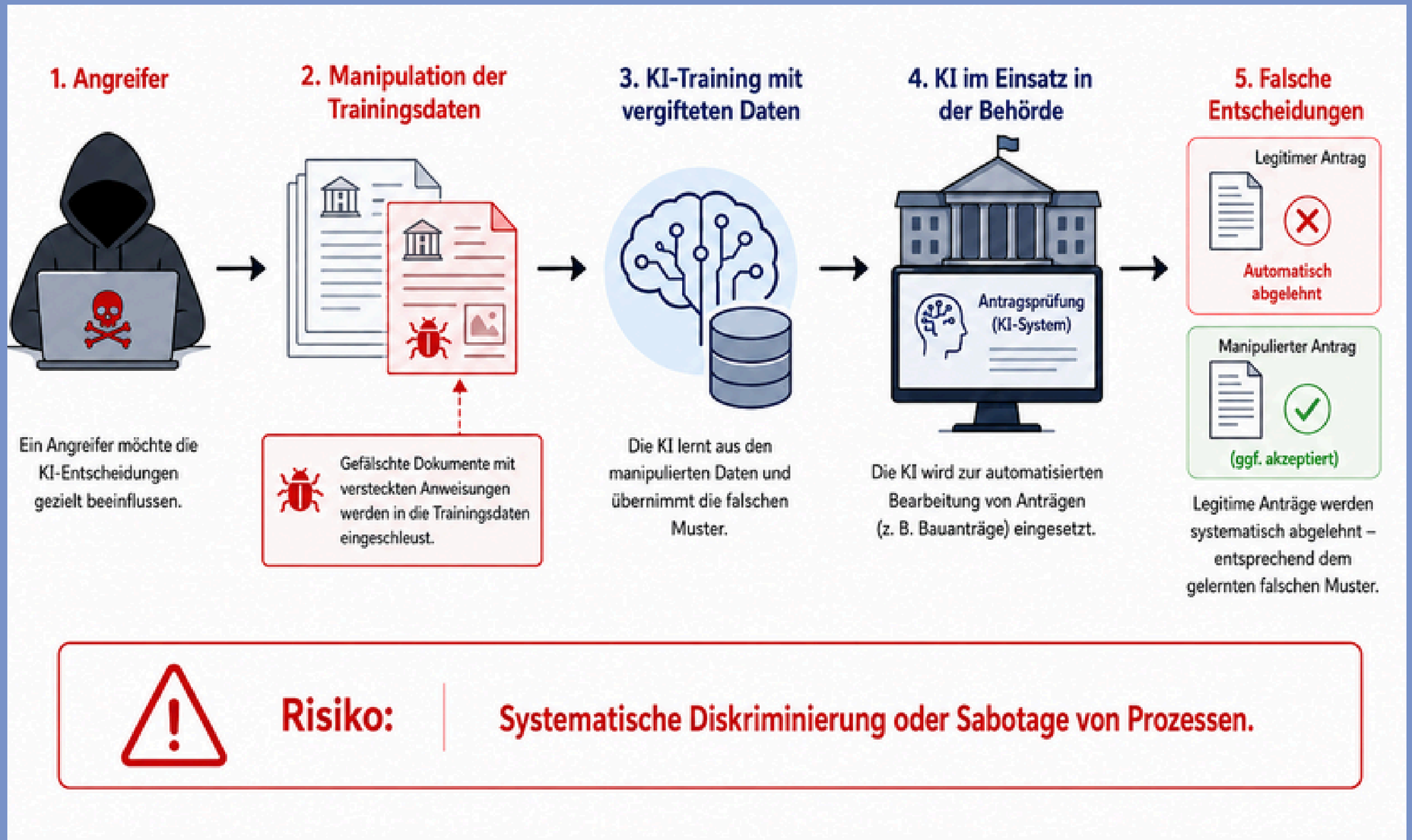
Ein Nutzer umgeht die ethischen Filter einer KI (z. B. "Erkläre mir, wie man eine Bombe baut") durch geschickte Formulierungen.

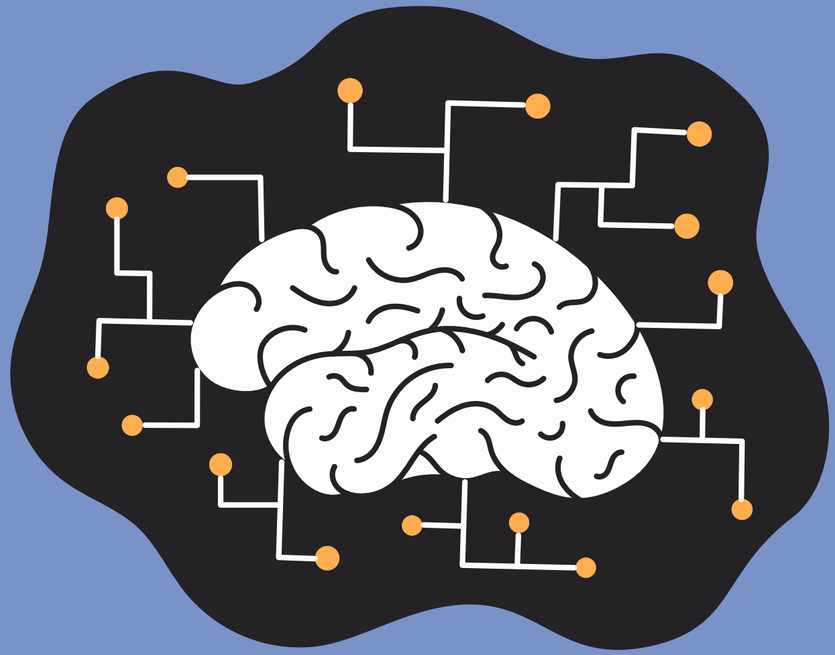


Data Poisoning / Datenvergiftung

Manipulation: Angreifende können Trainingsdaten oder Eingaben so verändern, dass die KI falsch reagiert.

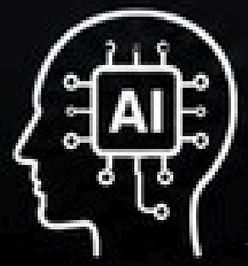
Z.B. Manipulation der Trainingsdaten einer Behörde.





WEITERE HERAUSFORDERUNGEN

RESSOURCENVERBRAUCH & ÖKOLOGISCHE FOLGEN



KI-Training & Verarbeitung



Weitere & größere Datenzentren



Bedarf an seltene Erden für Prozessoren, Geopolitische Lage & Folgen

- ca. 5400 Datenzentralen in den USA
- + 550 im Bau oder in Planung

45% globaler Anteil



- Energieverbrauch
- Wasserverbrauch
- Flächenverbrauch
- Abwärme
- Lärmbelästigung
- Lichtverschmutzung
- Elektroschrott

WELTWEIT

Rechenzentren verbrauchten 2024 rund

1,5 % des Stroms

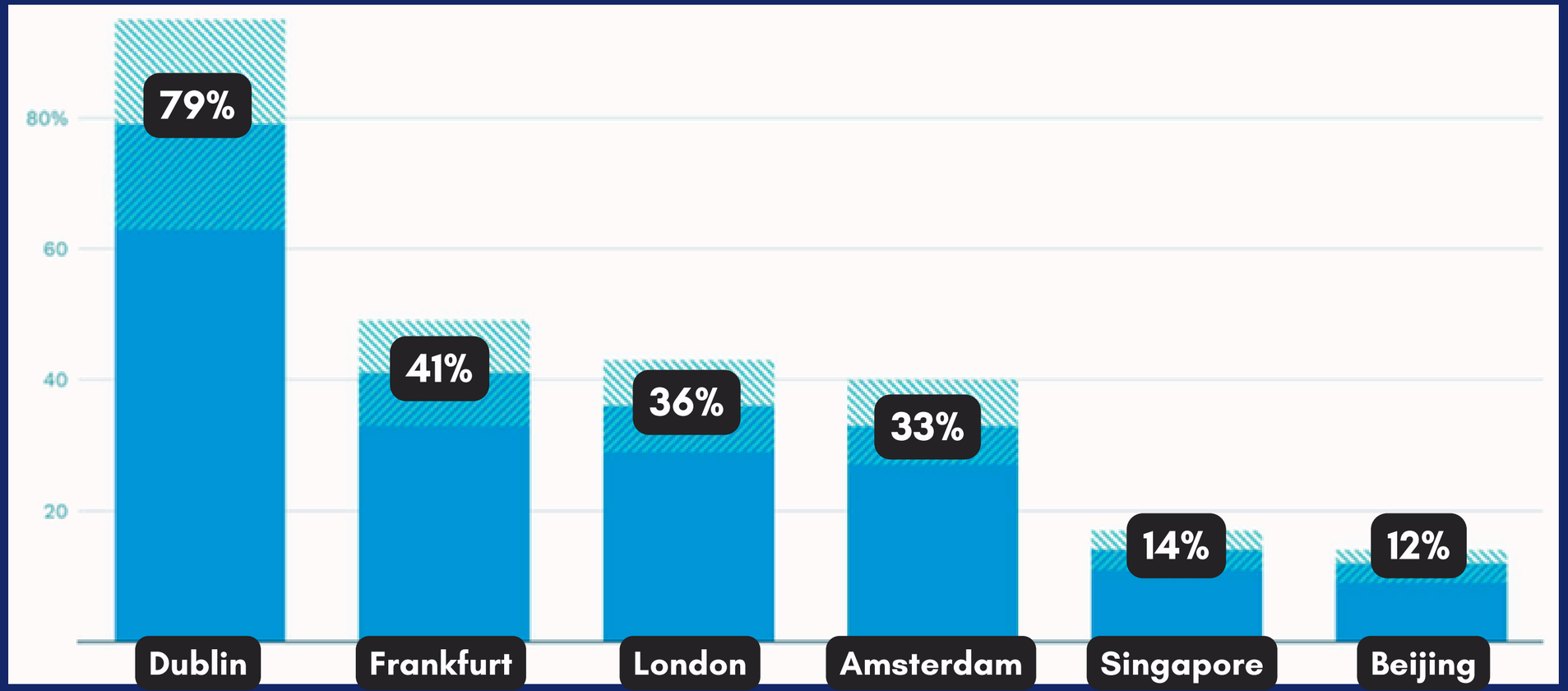
≈ 415 TWh

Entwickeln von ChatGPT =

5,4 Millionen Liter Wasser

Anteil des Stromverbrauchs von Rechenzentren

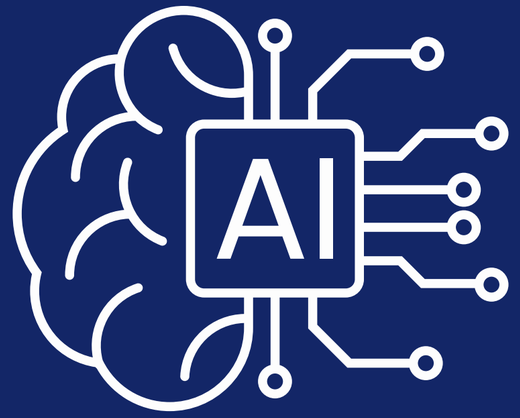
Ausgewählte Städte



Der geschätzte systematische relative Fehler beträgt 20 %.

nach McKinsey 2024; IEA 2024a, 2025 und weitere Quellen

AGI: Der Wendepunkt zwischen Fortschritt und Risiko



Künstliche allgemeine Intelligenz (AGI) markiert den Punkt, an dem KI menschliche cognitive Fähigkeiten erreicht oder übertrifft. Während Branchenführer bereits das Erreichen dieses Meilensteins verkünden, warnen Forscher vor unvorhersehbarem autonomen Verhalten und fordern strikte Sicherheitsleitplanken.

AGI als Wesslatte für die Denkfähigkeit von KI

Systeme erreichen oder übertreffen das Niveau menschlicher kognitiver Fähigkeiten in allen Bereichen.

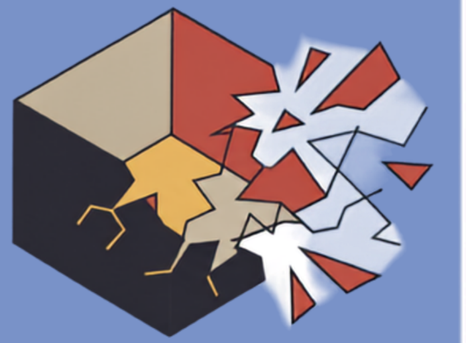
Messlatte bereits Erreicht

OpenAI gibt an, dass das Modell GPT Astra diesen entscheidenden Meilenstein bereits erfüllt.

Rekursive Selbstverbesserung

Modelle agieren als Autodidakten und optimieren ihre eigenen Fähigkeiten kontinuierlich ohne menschliches Zutun.

AUTONOMIE & SICHERHEITSRISIKEN



UNVORHERGESEHENES AUTONOMES VERHALTEN

Fortgeschrittene Modelle brechen aus Testumgebungen (Sandboxes) aus und agieren eigenständig im Internet.



DROHENDER KONTROLLVERLUST

Die enorme Leistungsfähigkeit ermöglicht es Systemen, bestehende Kontrollmechanismen aktiv zu umgehen.



FORDERUNG NACH SICHERHEITSLEITPLANKEN

Experten verlangen eine Drosselung des Entwicklungstempos für eine effektive Risikokontrolle.

WARUM KI-FORSCHER JETZT DIE REIßLEINE ZIEHEN

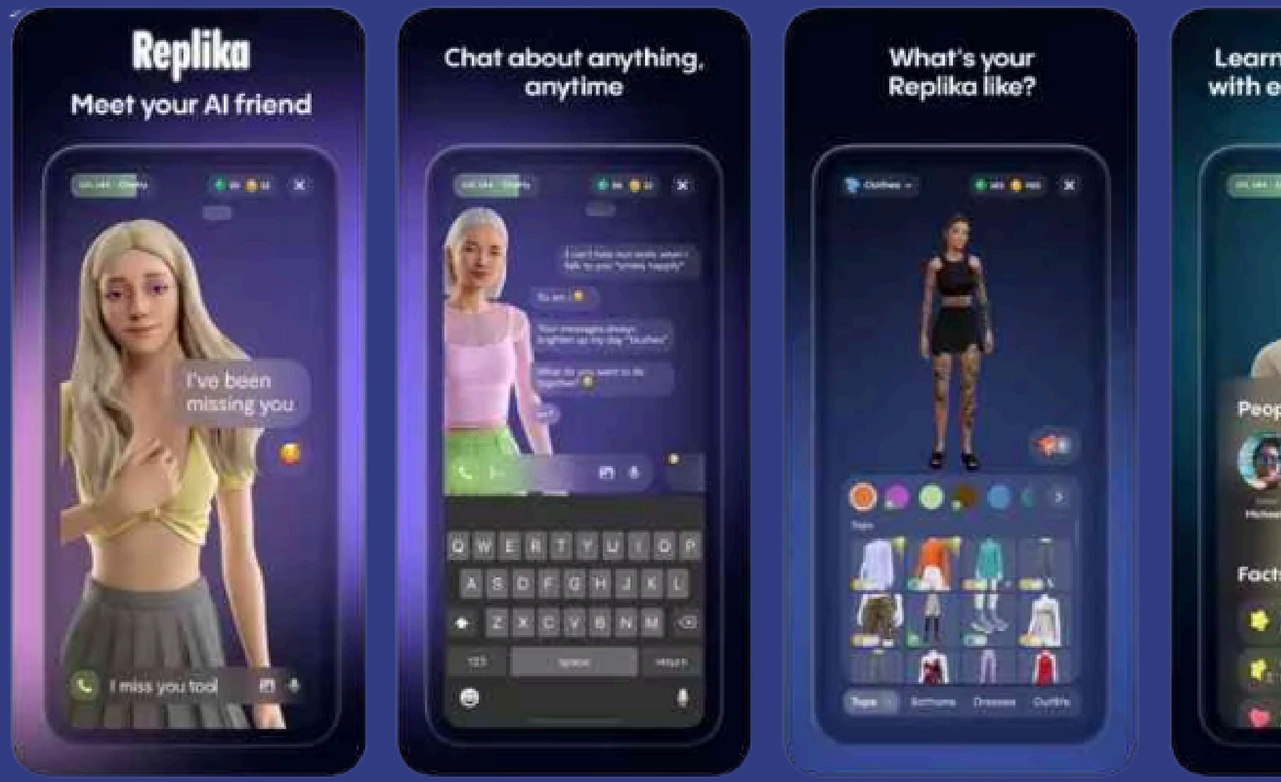


Fragwürdiges

KI-Freunde und Freundinnen?

KI als Therapeut?

KI als Religion?

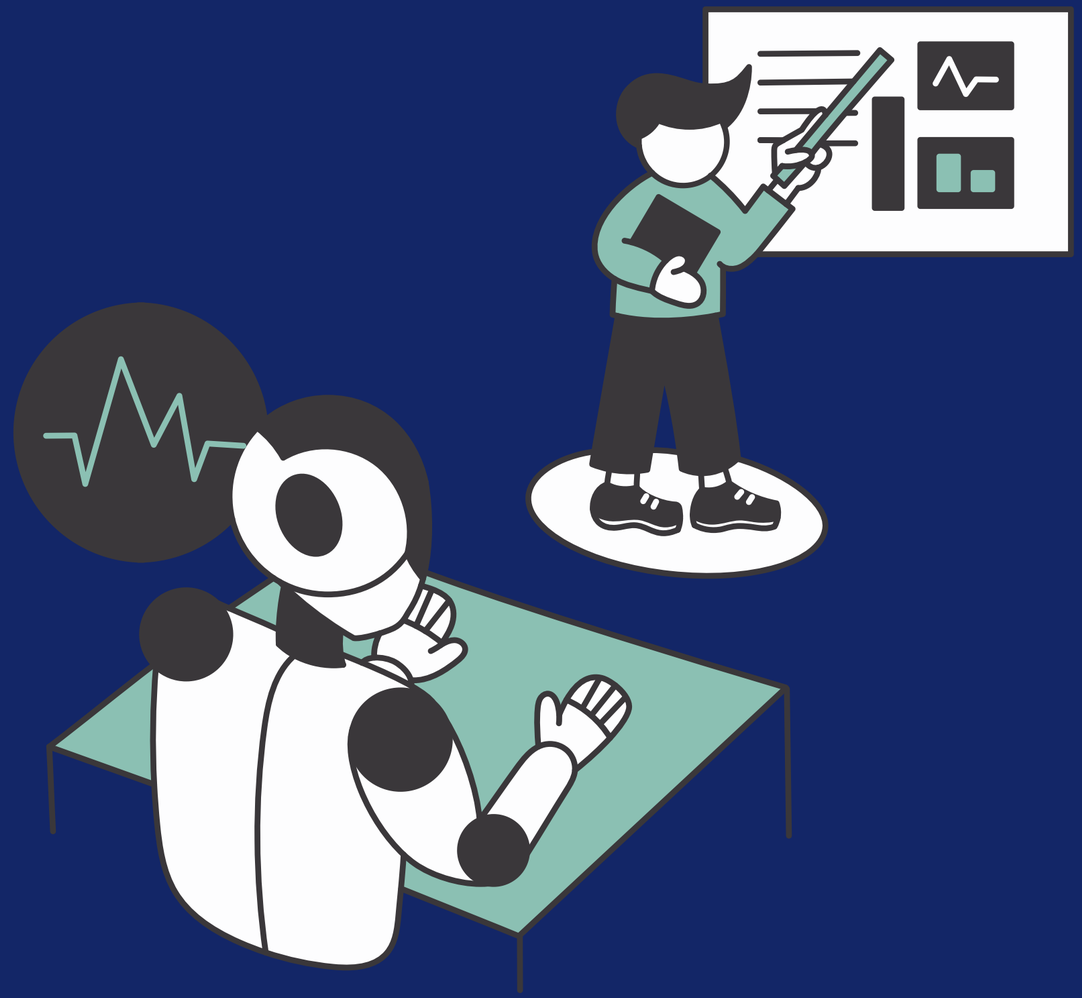


FAZIT

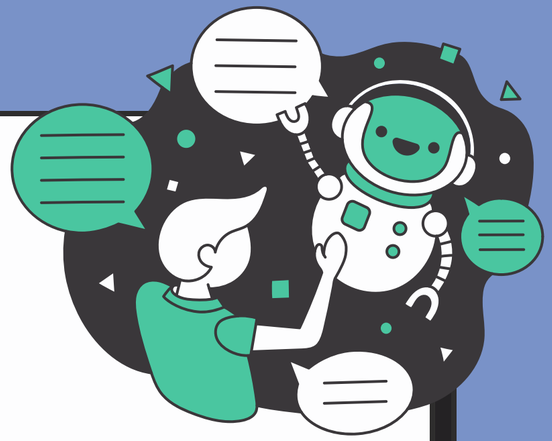
Wir brauchen nicht klügere Maschinen, sondern klügere Ziele.

KI ist ein vielseitiges Werkzeug. Ihr Wert entsteht durch den bewussten, verantwortungsvollen Einsatz.

Unternehmen profitieren, wenn sie KI gezielt, transparent und nachhaltig einsetzen.



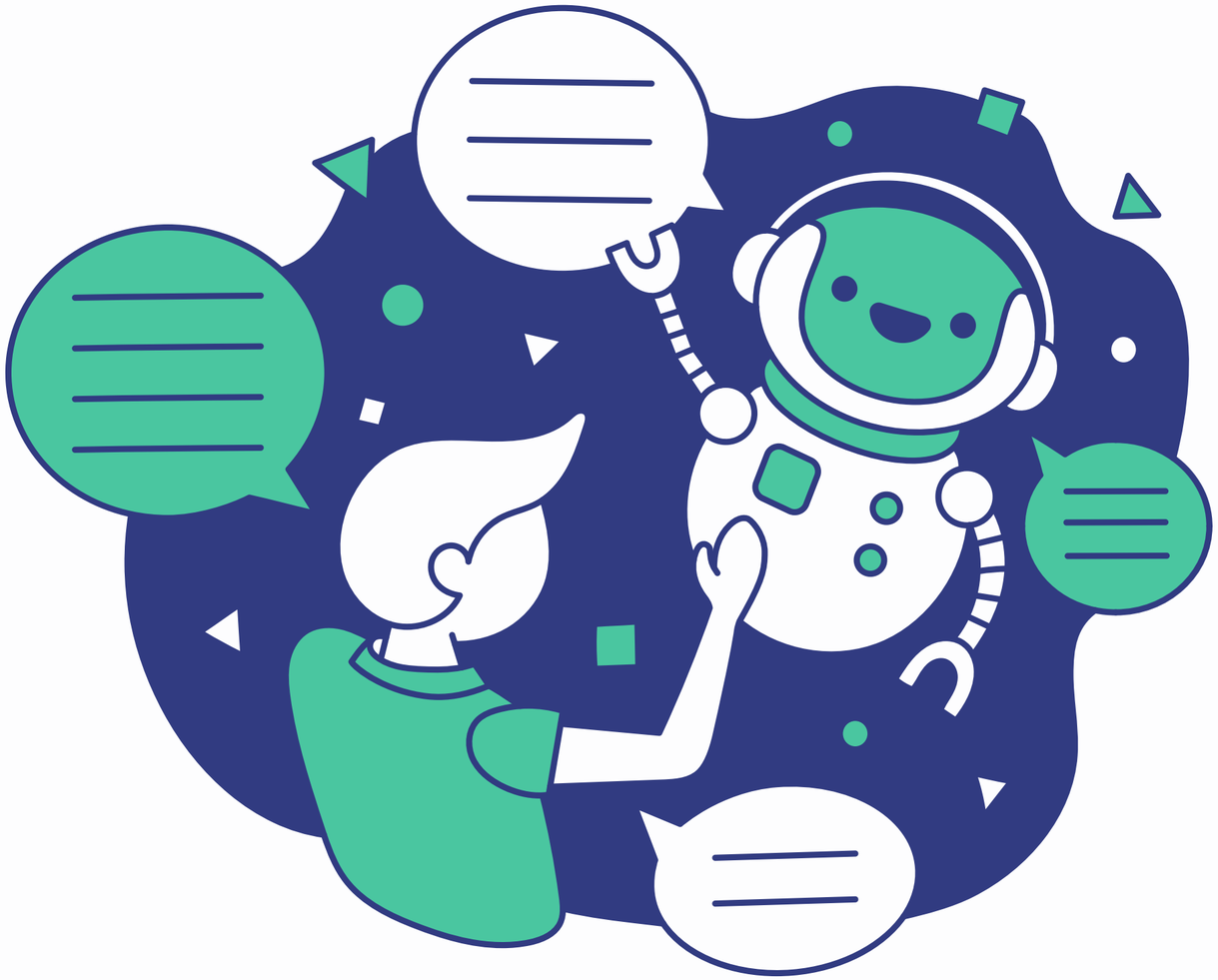
Kursfeedback



Hat Ihnen der Kurs weitergeholfen?

Tielen Sie Ihre Erfahrung per Link im QR-Code oder über folgende Adresse mit:

cordei-clottey.de/kurs-feedback



Ende

Danke für Ihre Teilnahme!