



# Prompten, Agents, Innovators oder einfach KI

Wie Algorithmen Wirtschaft und  
Gesellschaft verändern.

06. März 2025 | Vortrag @ GCHH

---



**Jonas Wilinski**  
PhD Student @ TUHH  
[Jonas@Wilinski.me](mailto:Jonas@Wilinski.me)



Die Folien & Links werden an  
alle Teilnehmer verteilt.

# Wie ist Ihr aktueller Stand?



Haben Sie bereits Berührungspunkte mit KI gehabt?  
Wurden Sie bereits von KI-generierten Materialien konfrontiert?

# Aufbau des Vortrags

## **Grundlagen der Künstlichen Intelligenz**

Definition, Beispiele & Anwendung

## **Generative KI & Large Language Modelle**

Praktische Anwendung von ChatGPT, Diffusion, Reasoning Agenten und die zu Grunde liegenden Konzepte

## **Implikationen in Wirtschaft und Gesellschaft**

Auswirkungen von KI in verschiedenen Themenfeldern sowie mögliche Integration in bestehende Prozesse

# Grundlagen der Künstlichen Intelligenz

Definition, Beispiele & Anwendung

Picture by Shubham Dhage on [Unsplash](#)



# Definition KI



Künstliche Intelligenz (KI) ist die Fähigkeit einer Maschine, menschliche Fähigkeiten wie **logisches Denken, Lernen, Planen** und **Kreativität** zu imitieren.

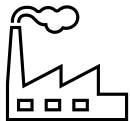
[EU: Was ist künstliche Intelligenz](#)



KI ermöglicht es technischen Systemen, **ihre Umwelt wahrzunehmen, mit dem Wahrgenommenen umzugehen und Probleme zu lösen**, um ein bestimmtes Ziel zu erreichen.

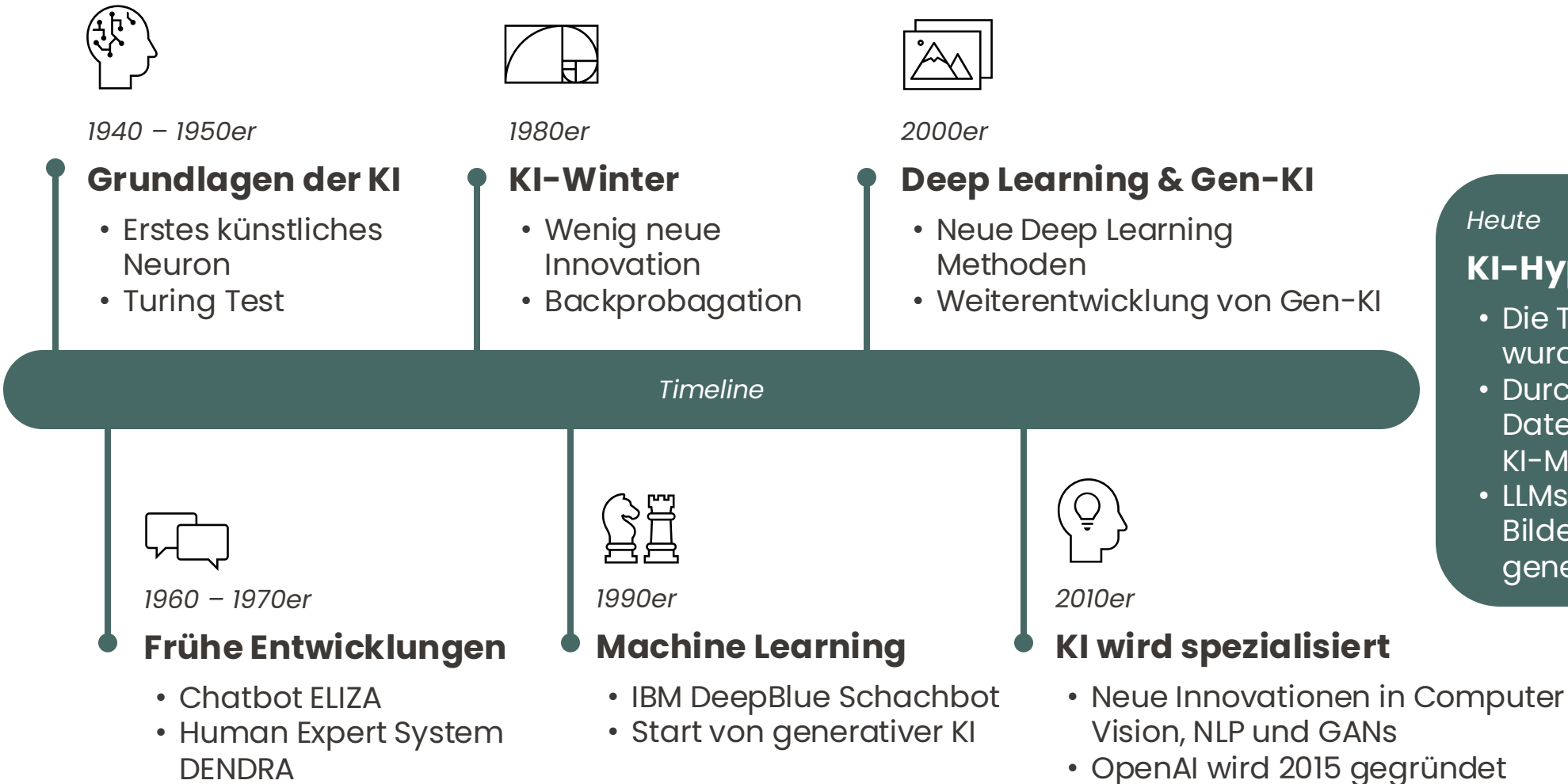


KI-Systeme können ihr **Handeln anpassen**, indem sie die Folgen früherer Aktionen analysieren und autonom arbeiten.



KI ist ein wesentlicher **Treiber für die digitale Transformation**. Sie wird in vielen Bereichen eingesetzt, um Effizienz zu steigern, Sicherheit zu verbessern und neue Erkenntnisse zu gewinnen.

# Historische Entwicklung von KI



*Heute*

**KI-Hype im vollen Gange**

- Die Transformer Architektur wurde 2017 veröffentlicht
- Durch immense Datenmengen können große KI-Modelle trainiert werden
- LLMs übernehmen Aufgaben, Bilder & Videos werden mit KI generiert

# KI-Terminologie: Ein inflationärer Begriff?



## Künstliche Intelligenz

Der Begriff “Künstliche Intelligenz” (KI) im weitesten Sinne bezieht sich auf die **Simulation menschlicher Intelligenzprozesse** durch Computersysteme



## Machine Learning

„Machine Learning“ (ML) ist ein Teilbereich der KI, welcher sich auf die **Entwicklung spezifischer Systeme** konzentriert, die auf der Grundlage von Daten lernen und Entscheidungen/Vorhersagen treffen können.



## Deep Learning

Deep Learning ist eine Unterkategorie des maschinellen Lernens, die eine bestimmte **Gruppe von Algorithmen** verwendet, die als **neuronale Netze** bekannt sind und oft viele Ebenen aufweisen.

# KI-Ethik: Verantwortung & Herausforderung



Die Nutzung von Künstlicher Intelligenz bringt **immense Chancen**, jedoch auch **ethische Fragestellungen**, die verantwortungsvoll adressiert werden müssen.



## Datenschutz und Privatsphäre

*Wie gehen wir mit den riesigen Datenmengen um, die KI-Systeme nutzen?*

- Risiken durch die Überwachung und Speicherung personenbezogener Daten (bspw. Gesichtserkennung in der Überwachung)



## Bias und Diskriminierung

*Wie vermeiden wir Vorurteile in KI-Modellen?*

- KI kann bestehende Vorurteile verstärken, wenn Trainingsdaten verzerrt sind (bspw. Diskriminierung bei Bewerberauswahl durch Algorithmen)



## Transparenz und Erklärbarkeit

*Wie machen wir KI-Entscheidungen nachvollziehbar?*

- „Black Box“-Modelle können schwer nachvollziehbar sein, was Vertrauen mindert (bspw. Entscheidungen autonomer Fahrzeuge)

# KI-Systeme im Alltag



## Google RankBrain

Seit 2015 setzt Google auf ein neues KI-basiertes Teilsystem für die rund 200 Suchalgorithmen, um die Websuche effizienter zu gestalten.



## DeepL

Das 2017 gegründete StartUp DeepL setzt neuronale Netze für die effiziente Übersetzung von Texten in allen Sprachen ein.



## Google Assistant, Siri & Alexa

Sprachassistenten nutzen KI für die Spracherkennung & -generierung, setzen jedoch auch Elemente des „Reinforcement Learning“ ein.



All diese Systeme sind in der Lage, von **Daten oder Interaktionen zu lernen**. Sie können nicht logisch denken, planen oder kreativ sein.

---

Exkurs:

# Generative KI & Large Language Models

ChatGPT, Stable Diffusion, Reasoning Modelle und die zu Grunde liegenden Konzepte

Picture by Shubham Dhage on [Unsplash](#)



# Kann ich Bildern trauen?



**Valderrama**  
Sotogrande, Spanien

on XT-4 Viltrox, landscape, Valderrama Golf Club: Lush Mediterranean landscape, cork oak trees, and a demanding test of accuracy.   
tlight, f/5, morning, very intricate, very detailed, realistic. sharp, bright, colorful, film grain, high contrast, <lora:Anti-Blur:1><lora:Insane  
x1\_lora\_flyway\_Epic-detail\_v2:0.3> <lora:fluxenhancer:0.25>

# Kann ich Audio trauen?



## Erneuerung der Rasenabschläge auf der Driving Range

Für die Erneuerung wurde die gesamte Fläche abgetragen, neu aufgefüllt, planiert und mit Rollrasen ausgelegt. Diese Maßnahme zielt darauf ab, die Qualität der Trainingsbedingungen zu verbessern und das Erscheinungsbild unserer Anlage zu optimieren.

### Wichtige Hinweise für die nächsten Wochen:

Die neu angelegte Fläche darf für die kommenden drei Wochen nicht betreten werden. Dies gilt auch für Bälle, die in der Nähe der Abschläge liegen. Bitte lassen Sie diese dort, um die neue Rasenfläche nicht zu beschädigen.

Sofern die Witterungsbedingungen es zulassen, planen wir, die Rasenabschläge in etwa sechs Wochen wieder freizugeben.



# Eleven | Labs



# Kann ich Videos trauen?

[OpenAI Sora](#)



# Intro: Was ist Generative KI

» Generative KI-Modelle können **neue Inhalte wie Text, Bilder, Audio oder Videos erzeugen**, basierend auf enormen Massen von Trainingsdaten.



Midjourney



synthesia



Anthropic



Microsoft  
Copilot



Adobe  
Firefly



GitHub  
Copilot

Eleven  
Labs

» Ein aktuelles Beispiel im Bereich der Audio-Generierung:  
[Google NotebookLM](#) – Fragen zu Dokumenten & Podcasts auf Basis von Text-Daten



Stets aktuelle Liste von KI-Tools nach Kategorie: [Awesome AI Tools](#)

# Fokus: Large Language Models

- **Kickoff durch ChatGPT**

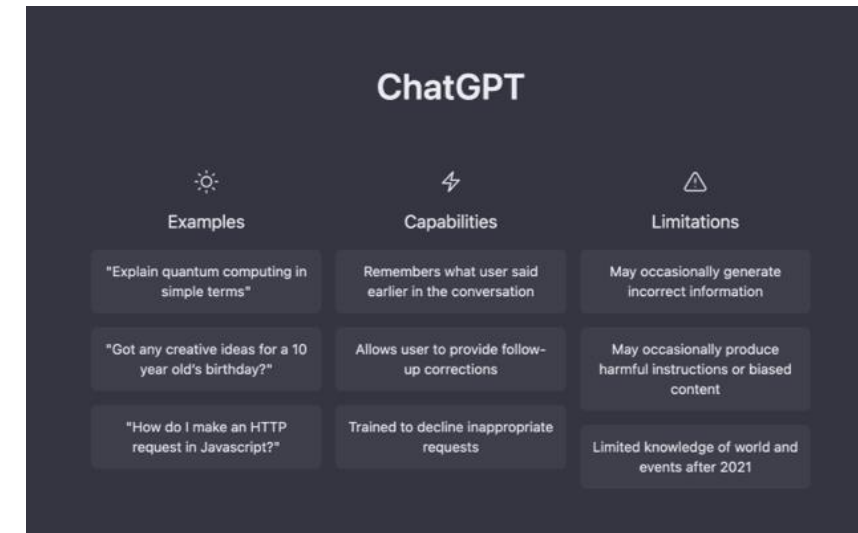
OpenAI zeigte durch ChatGPT 2022 das Potenzial von LLMs, menschenähnliche Dialoge und komplexe Textverarbeitung in Echtzeit zu ermöglichen.

- **Grundlagen**

LLMs wie GPT-4.5 sind auf **Milliarden von Textdaten** trainiert, um Sprache, Muster und Zusammenhänge zu verstehen und kontextbezogene Antworten zu generieren.

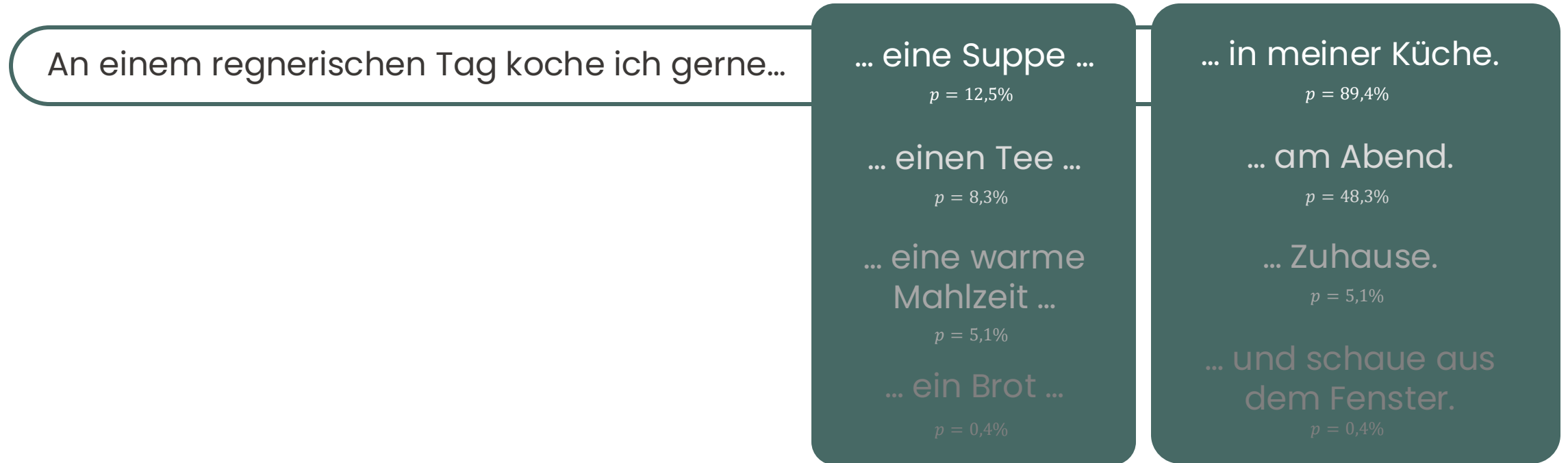
- **Anwendungen**

Von Textgenerierung und Übersetzungen bis hin zu Konversationsagenten, Codegenerierung und kreativen Prozessen.



Large Language Models wie ChatGPT **revolutionieren die Art und Weise, wie Maschinen menschliche Sprache verarbeiten und anwenden**, und eröffnen vielseitige Einsatzmöglichkeiten in Kommunikation, Kreativität und Automatisierung.

# Das Spiel mit den Wahrscheinlichkeiten

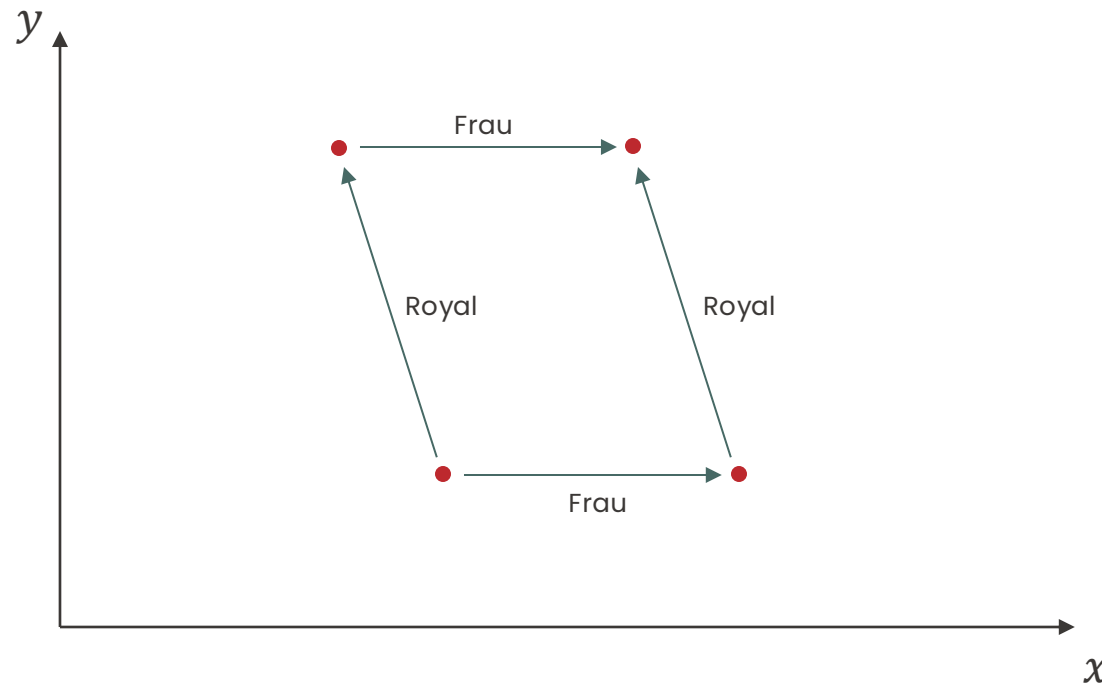


- Large Language Modelle **erlernen**, wie Wörter in welcher Konstellation in einem Satz vorkommen
- Sie sind in der Lage, auf Fragen (oder Prompts) **die wahrscheinlichste Antwort zu erzeugen**



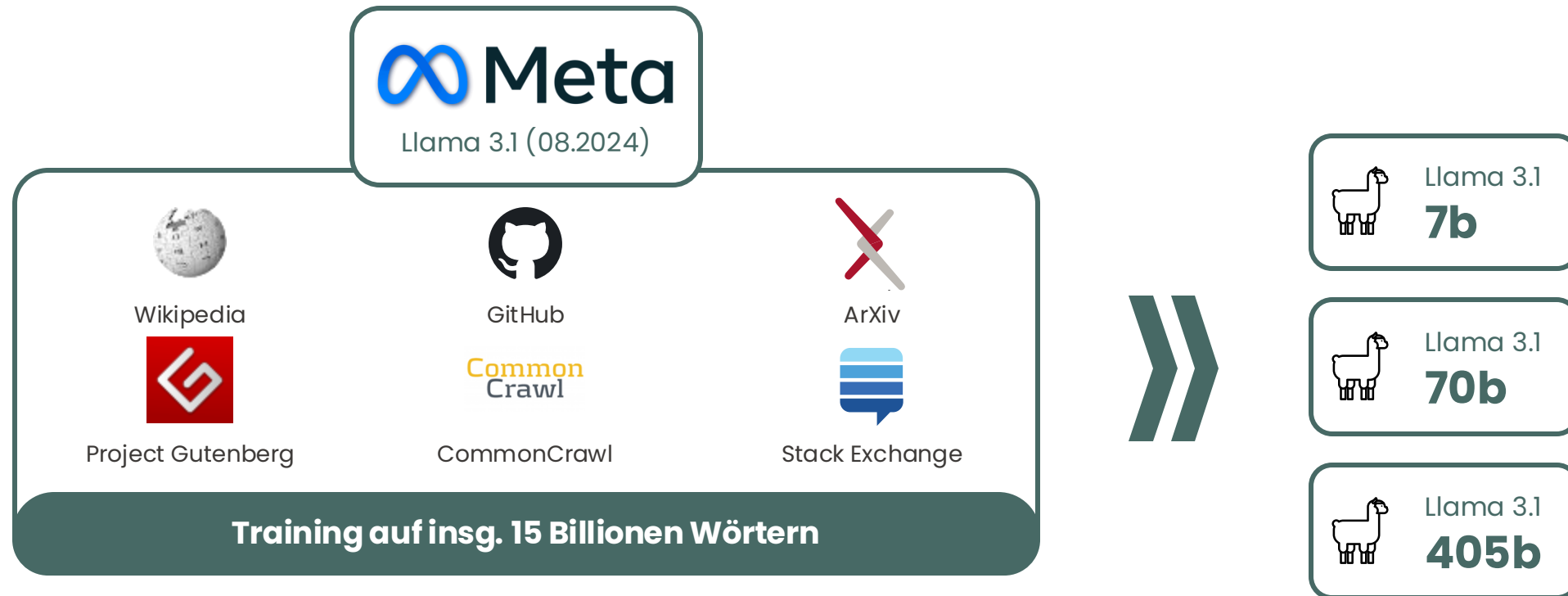
LLMs denken, argumentieren oder planen nicht – **sie folgen lediglich Wahrscheinlichkeiten**

König - Mann + Frau = Königin



# LLMs lernen und halluzinieren

- LLMs sind auf immense Massen von Textdaten angewiesen
- Es existieren einige freie Quellen, um Trainingsdaten zu akquirieren
- Die Herkunft der Trainingsdaten ist jedoch oft ungewiss und vereinzelt auch absichtlich nicht genannt



» LLMs **destillieren enorme Wissensmengen** auf eine Datei →

Sie erlauben es, mit diesem **Wissen zu interagieren** und **natürliche Fragen** zu stellen.



# Mehr Daten = Bessere Qualität?

- » Bisher galt, dass mit mehr Leistung und Trainingsdaten bessere Ergebnisse und weniger Halluzinationen erzielt werden.
- » Jüngste Ereignisse widerlegen diese Theorie und zeigen eine neue und innovative Herangehensweise.



## GPT 4.5 verfehlt Erwartungen

- Das neueste Modell von OpenAI kostete ca. 100mio \$ im Training
- Es scheint gut zu reagieren, ist aber in den meisten Tests unmerklich besser als 4o
- Die Kosten pro Anfrage sind ca. 30x so teuer als 4o



## Innovative Methoden übertreffen „Brute-Force“

- DeepSeek R1 (chinesisches KI-Modell) wurde mit einem Budget von 5,5mio \$ trainiert
- Es übertrifft einige Modelle und beherrscht sog. Chain-of-Thought Techniken
- OpenAI hat ebenfalls solche Modelle (o1/o3) – das Training kostete ca. 500mio \$.

# Reasoning & Deep Research

Warum Reasoning-Modelle die Zukunft sein können:

- Große Akteure in der KI-Welt ändern ihren Kurs – weg von immer größeren und komplexeren Modellen hin zu **Mixture-of-Experts** (MoE) und **Chain-of-Thought** (CoT) Implementierungen.
- Diese Modelle zerlegen komplexe Aufgaben in klare, sequenzielle Schritte, was die Genauigkeit verbessert.
- Ihre eingebaute Selbstkorrektur reduziert Halluzinationen und erhöht die Transparenz der Ausgaben.
- Verbesserte Anpassungsfähigkeit durch Fine-Tuning und Reinforcement Learning macht sie vielseitig einsetzbar in unterschiedlichsten Domänen.



# Reasoning & Deep Research



Kurzer Prompt / Einfache Frage:  
*„Ist Golf ein Anfängersport?“*

1

Sammeln von Informationen  
*Internet-Suche, Datenbank*

2

Lesen & Verstehen  
*Wichtige Dokumente identifizieren*

3

Selbstkorrektur  
*Iteratives Prompting (Brainstormen)*

4

Sammeln aller Informationen  
*Zusammenfassung & Struktur*



R1



o1 / o3



Ausgabe  
*Report oder Text an den User geben*

## Ist Golf ein Anfängersport?

### Argumentation mit R1

20 Quellen ^

#### Bewerten, ob Golf als Anfängersport gilt ^

Suche

Q Ist Golf ein Anfängersport Q Golf für Anfänger +1 more

Okay, ich muss die Frage beantworten, ob Golf ein Anfängersport ist. Zuerst schaue ich...

Die Platzreifekurse sind Pflicht, um auf den Platz zu spielen, und kosten etwa 179 Euro,...

Einige Quellen betonen, dass Golf nicht mehr elitär ist und für jeden Geldbeutel...

Golfclubs bieten verschiedene Trainingsmöglichkeiten, wie Driving Ranges und...

Zusammenfassend scheint Golf ein gut zugänglicher Sport für Anfänger zu sein, mit...

Abschluss der Analyse



Golf ist ein Sport, der sich hervorragend für Anfänger eignet, unabhängig von Alter oder Fitnesslevel. Die weit verbreitete Vorstellung von Golf als elitärer oder spießiger Sport entspricht nicht mehr der Realität – moderne Clubs bieten niedrighschwellige Einstiegsmöglichkeiten und haben sich zu inklusiven Sportstätten entwickelt.

### Altersunabhängiger Einstieg

Man kann Golf **ab 5 Jahren** beginnen, aber auch mit **75+** problemlos einsteigen <sup>1</sup>. Der Sport ermöglicht generationsübergreifendes Spielen durch ein ausgeklügeltes Handicap-System, das unterschiedliche Spielstärken ausgleicht <sup>1</sup> <sup>4</sup>. Selbst Senioren finden im Golf eine gelenkschonende Bewegungsform, die Ausdauer und Koordination trainiert <sup>3</sup> <sup>7</sup>.

### Strukturierte Lernwege

Der Einstieg erfolgt typischerweise über drei Phasen:

1. **Schnupperkurs** (ab 10-20€) mit Leihausrüstung <sup>1</sup> <sup>2</sup> <sup>4</sup>
2. **Platzreifekurs** (ca. 179€ für 10-16 Stunden) für Regelkunde und Grundtechniken <sup>1</sup> <sup>8</sup>
3. **Flexible Trainingsmöglichkeiten** auf Driving Ranges und Kurzplätzen <sup>6</sup> <sup>8</sup>

Viele Clubs wie der Golfclub Ansbach oder Gut Waldhof bieten spezielle Anfängerpakete, die bei späterer Mitgliedschaft die Kurskosten erstatten <sup>1</sup> <sup>3</sup>. Moderne Trainingsmethoden kombinieren Technikübungen mit spielerischen Elementen – etwa 60% der Ausbildungszeit entfallen auf praktische Platzübungen <sup>8</sup>.

### Kostengünstiger Einstieg

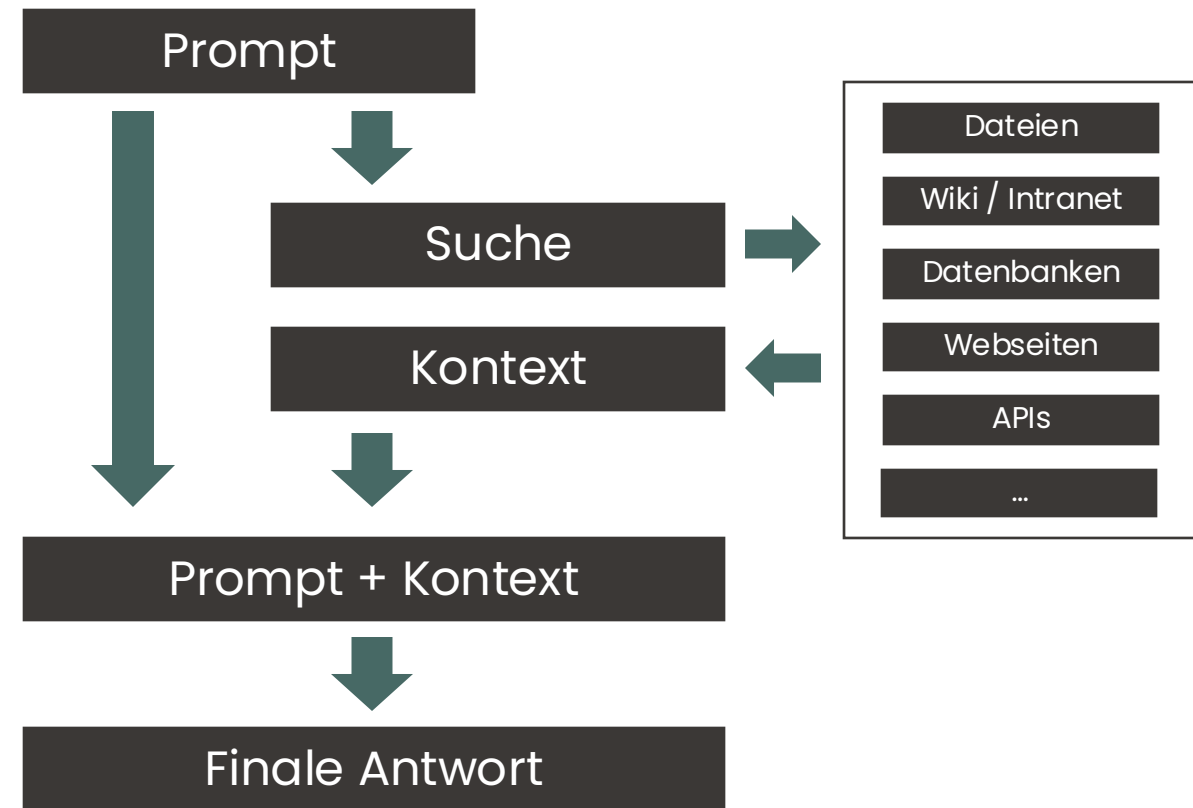
Entgegen landläufiger Meinung ist Golf kein reiner „Reichensport“ mehr:

- Erstausrüstung ab 200€ möglich
- Schlägerverleih an meisten Plätzen (3-10€/Runde)
- Jahresmitgliedschaften ab 500€ <sup>2</sup> <sup>4</sup>

<https://www.perplexity.ai/search/ist-golf-ein-anfangersport-4hLf5vzWQYOL5kAhVO8jzw>

# Beispiel **Retrieval Augmented Generation**

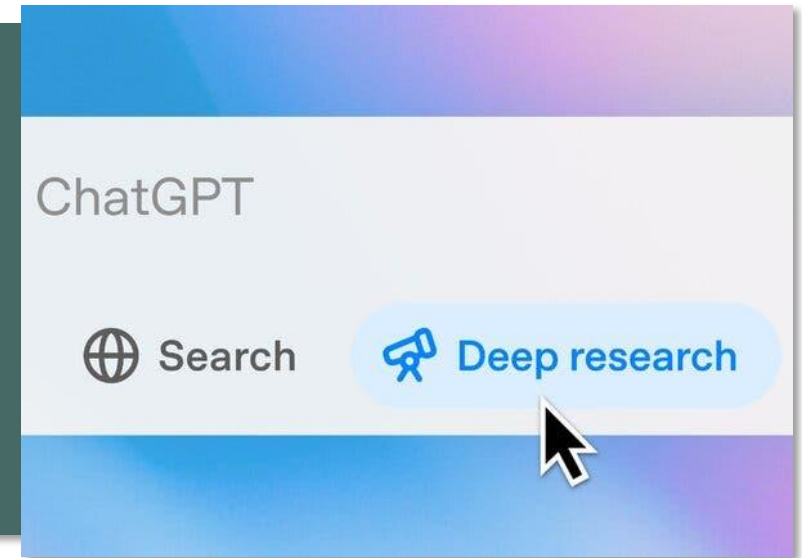
- **Externe Wissensquellen:**  
RAG kombiniert ein LLM mit der Fähigkeit, Informationen aus externen Datenquellen abzurufen.
- **Verbesserte Genauigkeit:**  
Das Modell liefert genauere und aktuellere Antworten, da es während der Generierung relevante Informationen bezieht.
- **Nachverfolgbarkeit:**  
Anders als herkömmliche LLMs kann eine RAG Pipeline Quellenangaben liefern und somit Halluzinationen vorbeugen
- **Zwei Schritte:**  
Erst werden passende Daten abgerufen, dann nutzt das Modell diese Informationen, um einen fundierten Text zu erstellen.



# Reasoning Modelle ermöglichen “Deep Research”

OpenAI und Perplexity bieten **erweiterte Recherche-Modi** an:

- Es werden zu einer klaren Aufgabenstellung simultan **viele Quellen** verarbeitet
- Der resultierende Report ist lang und detailliert – geht insbesondere auf die geforderten Punkte ein und kann als PDF exportiert werden



- ✓ Zeitersparnis: Ergebnisse in Minuten statt Stunden.
- ✓ Fundierte Entscheidungen durch detaillierte und strukturierte Berichte.
- ✓ Anpassbare Prompts für spezifische Anforderungen (z. B. Marktanalysen, Due Diligence)

# Forschungsberichte erstellen lassen (I/II)

Wie kann ich KI in mein Unternehmen integrieren? Ich führe ein mittelständisches Unternehmen für das verleihen von Baumaschinen. Ich habe ca. 15 Büromitarbeiter und 5 Mitarbeiter im Außendienst. Wir nutzen bereits Microsoft 365 für Outlook, Teams usw.

🔍 Tiefe Forschung

56 Quellen

KI in mittelständischen Unternehmen - KI...  
ki-trainingszentrum

Wie syniotec Künstliche Intelligenz für...  
syniotec

Microsoft entfesselt KI für Mittelstand - ORBIT...  
orbit

+53 quellen



🔍 Tiefe Forschung

56 Quellen

- I need to search the web to find information on how AI can be integrated into medium-...
- Die bisherigen Suchergebnisse haben allgemeine Strategien zur KI-Integration im...
- Ich habe zuvor branchenspezifische Informationen zur KI-Integration im...
- Ich habe genügend Informationen gesammelt, um Ihre Frage zu beantworten. Die...
- Okay, ich muss dem Benutzer helfen, KI in sein mittelständisches Unternehmen für...
- Aus Quelle 1 geht hervor, dass 20% der deutschen Unternehmen KI strategisch nutzen...
- Quelle 3 und 6 erwähnen Microsoft Copilot, der jetzt auch für KMUs verfügbar ist, ohn...
- Quelle 5 und 11 diskutieren KI in Baufahrzeugen und Erdbewegungsmaschinen,...
- Quelle 10 bietet KI-Beratung für den Mittelstand und nennt Einsatzmöglichkeiten wie...
- Quelle 18 geht auf prädiktive Wartung in der Werkstatt ein, was für die Wartungsplanun...
- Zusammenfassend sollte der Bericht die folgenden Bereiche abdecken: Einführung in KI...
- Recherchebericht schreiben

<https://www.perplexity.ai/search/wie-kann-ich-ki-in-mein-untern-jdUIRISjQIKS5bxcxqwFaQ>

# Forschungsberichte erstellen lassen (II/II)

- Forschungsberichte beinhalten viele Quellen (56 in diesem Beispiel)
- Sie benötigen längere Zeit, da mehr Rechenleistung zum Verarbeiten der Informationen benötigt wird
- Sie fassen auf mehreren Seiten das Thema zusammen und bieten **tiefgreifende Analysen & Handlungsempfehlungen**
- Texte sind mit Quellenangaben ausgestattet, um Fakten nachzulesen und weitere Schritte zu tätigen
- Ein Forschungsbericht kann dynamisch durch Prompts erweitert und verändert werden

## Strategische Schritte zur Implementierung

### Fallbeispiel:

Die Integration von KI in die Werkstattprozesse eines Baumaschinenunternehmens zeigt, dass



## Integration von Künstlicher Intelligenz in mittelständische Baumaschinenverleih-Unternehmen

Künstliche Intelligenz (KI) hat sich als transformative Kraft in der Industrie etabliert, insbesondere für mittelständische Unternehmen, die ihre Wettbewerbsfähigkeit durch Prozessoptimierung, Automatisierung und datengesteuerte Entscheidungen stärken möchten. Für ein Unternehmen im Baumaschinenverleih mit 15 Büro- und 5 Außendienstmitarbeitern, das bereits Microsoft 365 nutzt, bietet KI erhebliche Potenziale zur Effizienzsteigerung und Kundenzufriedenheit. Dieser Bericht analysiert konkrete Anwendungsfälle, technologische Lösungen und Implementierungsstrategien unter Berücksichtigung aktueller Entwicklungen und branchenspezifischer Herausforderungen.

### Grundlagen der KI-Integration im Baumaschinenverleih

#### Digitalisierung als Voraussetzung für KI

Die digitale Transformation bildet die Basis für den erfolgreichen Einsatz von KI. Laut Studien nutzen bereits 47 % der mittelständischen Unternehmen in Deutschland KI-Technologien, wobei der Schwerpunkt auf der Automatisierung von Geschäftsprozessen liegt. Ein zentraler Aspekt ist die Erfassung von Betriebsdaten, die für KI-Analysen erforderlich ist. Für Baumaschinenvermieter bedeutet dies die Integration von Sensordaten und die Nutzung von Plattformen wie Microsoft 365 um KI-Tools zu erweitern und IoT-Sensordaten zu erfassen [2] [3].

[https://www.perplexity.ai/page/integration-von-kunstlicher-in-RWVw3DEtqaqVeoGjO\\_A](https://www.perplexity.ai/page/integration-von-kunstlicher-in-RWVw3DEtqaqVeoGjO_A)

# Forschungsberichte via Perplexity oder OpenAI



- Benötigt das ChatGPT PRO Abonnement (Plus 5 p.M.)
- Kostet 200 USD / Monat
  - Plus kostet 20 USD / Monat
- Bietet „Deep Research“ & o1/o3 Pro Modelle
- Bindet Quellen aus dem Internet an



- Benötigt das Perplexity Abo
- Kostet 20 USD / Monat
- Bietet „Deep Research“, verschiedene Modelle (o3, Claude, Gemini, R1, ...)
- Bindet Quellen aus dem Internet an

# Large Language Models: Herausforderungen & Limitierungen



## Datenabhängigkeit und Bias

Trainingsdaten enthalten oft **Verzerrungen, Verunreinigungen und reichen nur bis zum Trainingsdatum**, was zu falschen oder alten Ergebnissen führen kann.

» Beispiel: Veraltete Informationen werden sehr elegant als Fakten dargestellt.



## Transparenz und Erklärbarkeit

LLMs agieren als **“Black Box”**, wodurch es schwierig ist, Entscheidungsprozesse nachzuvollziehen oder zu erklären.

» Fehlende Erklärbarkeit kann das Vertrauen in kritischen Bereichen wie Medizin oder Recht untergraben → RAG erforderlich.



## Rechenleistung und Ressourcen

Training und Betrieb von LLMs erfordern **extrem viel** Rechenleistung und Energie, was sie kostenintensiv und umweltbelastend macht.

» Die CO2-Emissionen von großen Modellen sind eine wachsende Besorgnis in der Technologiebranche.



## Halluzinationen / falsche Antworten

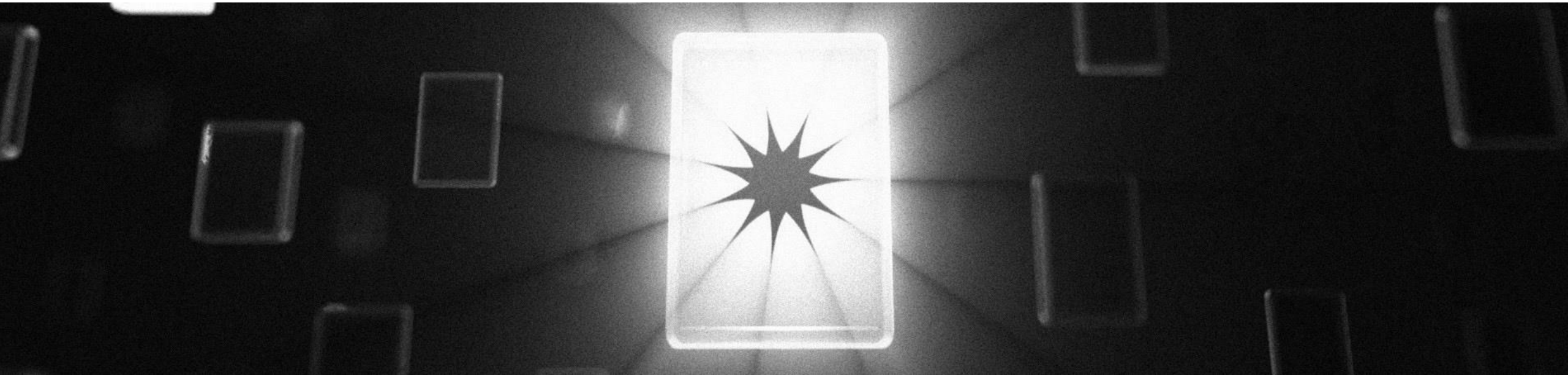
LLMs generieren vereinzelt falsche oder unlogische Informationen, die als Tatsachen dargestellt werden. Dies kann auf Training oder Betrieb zurückzuführen sein.

» Falschinformationen bei der Beantwortung von Wissensfragen → Reasoning kann hier helfen.

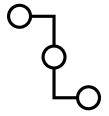
# Implikationen in Wirtschaft und Gesellschaft

Auswirkungen von KI in verschiedenen Themenfeldern sowie mögliche Integration in bestehende Prozesse

Picture by DeepMind on [Unsplash](#)



# Wirtschaftliche Chancen und Potenziale



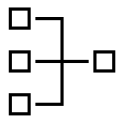
## Effizienzsteigerung & Automatisierung:

Optimierung von Prozessen, Reduktion von Fehlern, Kosteneinsparungen.



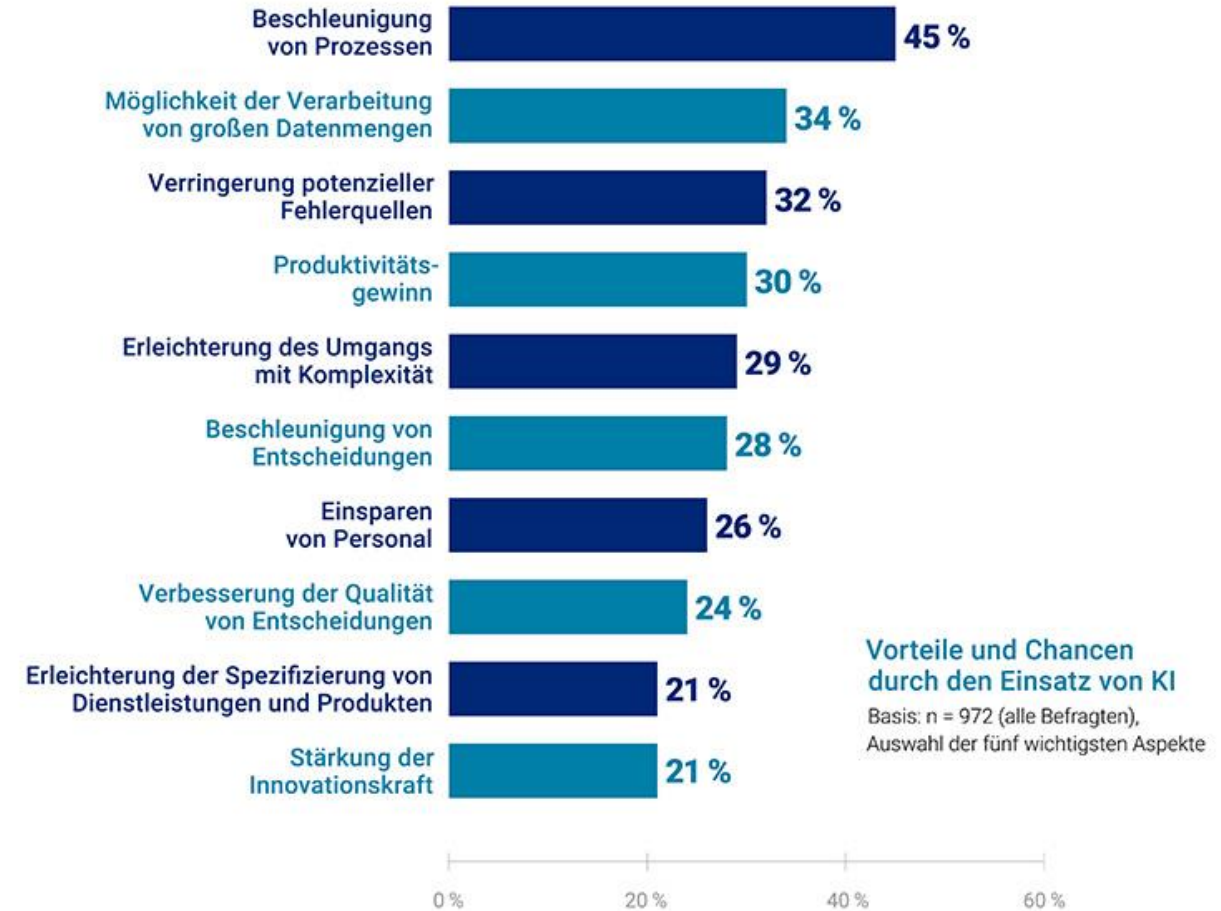
## Neue Geschäftsmodelle:

Personalisierte Produkte und Dienstleistungen, datengetriebene Innovation.



## Innovationskraft:

Einsatz von KI zur Produktentwicklung, Softwareentwicklung und Entscheidungsunterstützung.



HAYS: KI-Transformation: Effizienzsteigerung statt Innovation im Mittelpunkt (2024)

# Fallstudien und Praxisbeispiele aus der Wirtschaft



## **DB Smile**

Ein Chatbot, welcher mit KI und DB-Hintergrundwissen passende Stellen für Bewerber raussucht und bei Fragen assistiert.

## J.P.Morgan

### **JP Morgan – AI Fraud Detection**

JP Morgan nutzt KI-gestützte Systeme, um Transaktionen in Echtzeit zu analysieren und Betrugsmuster zu erkennen.



## **General Electric**

GE nutzt KI für Vorausschauende Wartung, um Geräteprobleme vorherzusagen und zu beheben, bevor sie die Produktion stören.



Viele Praxisbeispiele implementieren spezialisierte Formen von KI, ohne das volle Potential von generativer KI auszunutzen.

# Gesellschaftliche Auswirkungen und Herausforderungen



## **Arbeitsmarkt:**

Schaffung neuer Berufsbilder vs. Wegfall traditioneller Tätigkeiten.



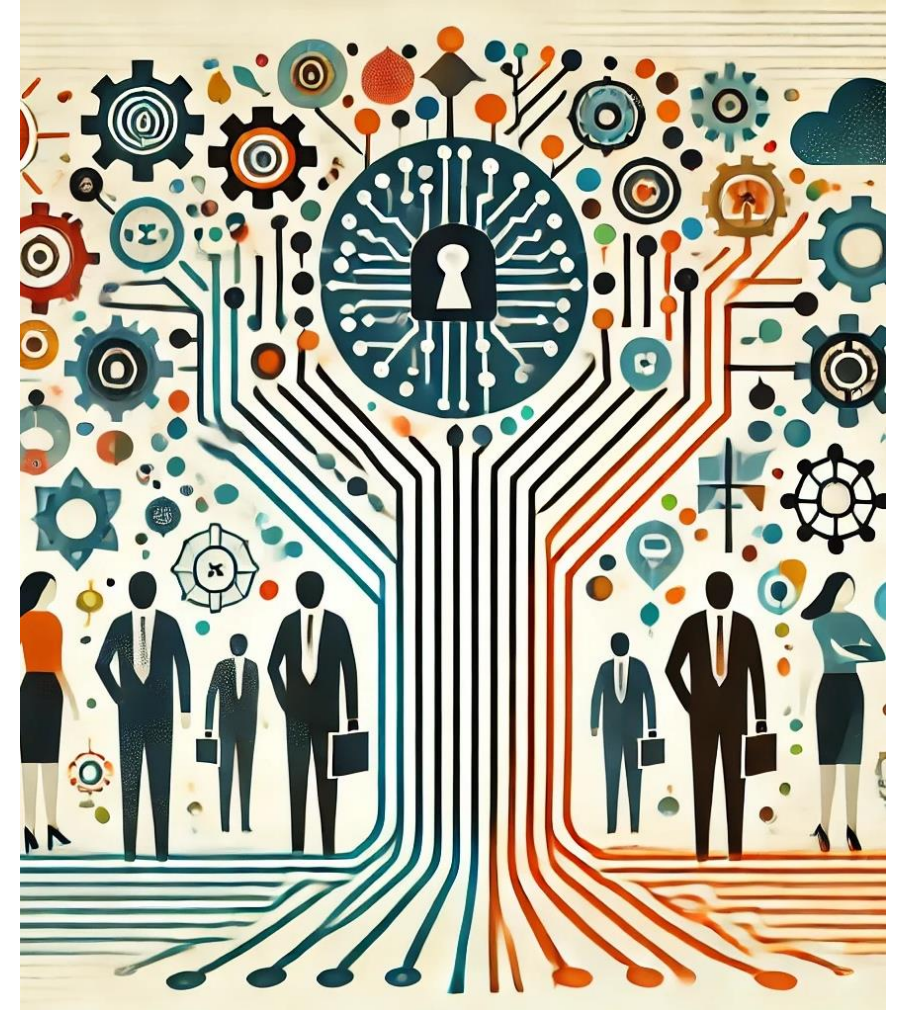
## **Soziale Teilhabe:**

Digitale Kluft: Unterschiedliche Zugänge zu KI-Technologien, Herausforderungen in Bildung und Weiterbildung.



## **Datenschutz & Sicherheit:**

Anforderungen der DSGVO, Risiken bei Datenmissbrauch.



# Zusammenfassung: Generative KI als Chance?

Effizienzsteigerung in vielen Branchen



**Automatisierung** von Text- und Kreativprozessen **spart Zeit und Ressourcen.**

Neue kreative Möglichkeiten



Generative KI ermöglicht **neue Formen von Kunst, Musik und Design**, die bisher unvorstellbar waren.

Domänenspezifische Innovationen



KI-Modelle werden gezielt auf **spezifische Branchen** (z.B. Gesundheitswesen, Finanzen) zugeschnitten.

Verantwortungsvoller Einsatz



Notwendigkeit von **ethischen Leitlinien** und **Regulierungen**, um Missbrauch und Risiken zu minimieren.

# Fragen? Use Cases?



Alle Folien, Links und den Podcast unter:  
[wilinski.me/ki-workshop](https://wilinski.me/ki-workshop)



**Jonas Wilinski**  
PhD Student @ TUHH  
[Jonas@Wilinski.me](mailto:Jonas@Wilinski.me)



[wilinski.me/ki-workshop](https://wilinski.me/ki-workshop)

# BACKUP

Picture by DeepMind on [Unsplash](#)

# Benötige ich einen Reasoning-Chat?

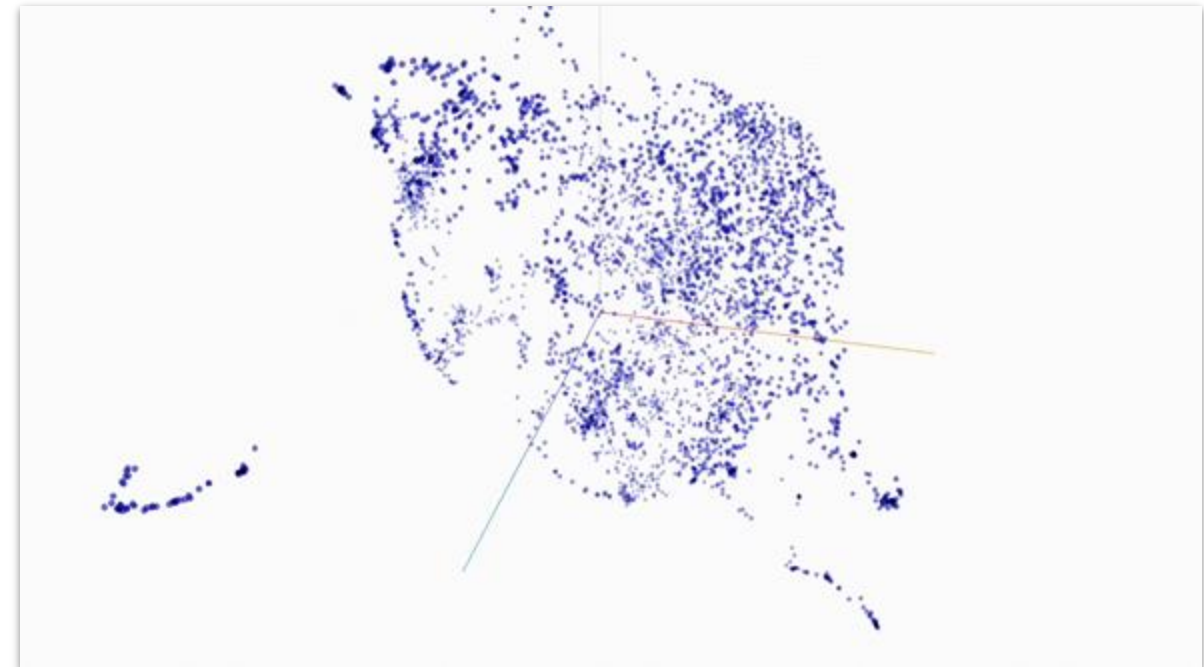
|                             | <b>Reasoning-Modelle</b><br><i>Der Planer</i>                                                                                                                                                                                                         | <b>GPT-Modelle</b><br><i>Der Arbeiter</i>                                                                                                                                                                |
|-----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Geeignete Aufgaben          | <ul style="list-style-type: none"><li>• Brainstorming / Diffuse Fragen</li><li>• Mehrschrittige Planungsaufgaben</li><li>• Code-Review / Debugging</li><li>• Logik &amp; Mathematik Aufgaben</li></ul>                                                | <ul style="list-style-type: none"><li>• Q&amp;A Chatbots</li><li>• Content-Erstellung mit genauem Plan</li><li>• Code-Vervollständigung</li><li>• Repetitive Aufgaben &amp; Analysen nach Plan</li></ul> |
| Aufgaben-<br>beschreibung   | <b>Knapp &amp; Direkt</b><br>Diese Modelle prompten sich selbst über mehrere Iterationen – sie planen und können mit wenigen Informationen auskommen.                                                                                                 | <b>Detailliert &amp; Beschreibend</b><br>Diese Modelle reagieren einmalig auf Ihren Prompt und halluzinieren schnell bei diffuser Aufgabenstellung.                                                      |
| Kontext                     | <b>Erforderlich</b><br>Sowohl Reasoning- als auch einfache GPT-Modelle benötigen ausreichend Kontext, um valide und hilfreiche Antworten zu generieren. Kontext kann durch erklärende Sätze im Prompt aber auch durch Dokumente (RAG) erzielt werden. |                                                                                                                                                                                                          |
| Formatierung<br>& Beispiele | <b>Optional</b><br>Spezielle Formatierung & Beispiele können bei komplexen Aufgaben helfen, werden i.d.R. aber vom Reasoning-Modell selber erstellt.                                                                                                  | <b>Erforderlich</b><br>Beispiele & Formatierungen verbessern die Präzision und Konsistenz der Ergebnisse enorm und beugen Halluzinationen vor.                                                           |

# Wörter & Texte als Vektoren

Einzelne Wörter und ganze Texte können als „Embeddings“ in eine hochdimensionale Vektorrepräsentation überführt werden:

$$f_{word2vec}: V \rightarrow \mathbb{R}^d \quad \gg \quad v_{play} = \begin{pmatrix} -0.224 \\ 0.130 \\ \dots \\ 0.276 \end{pmatrix} \quad \downarrow \quad \begin{array}{l} \text{Typischerweise} \\ 768\text{-dimensional} \end{array}$$

- Embedding-Modelle erlernen, dass **semantisch ähnliche Wörter ähnliche Vektoren** erhalten
- Die **Distanzen** zwischen diesen Vektoren sind aussagekräftig für **semantische Ähnlichkeit**
- Die **Qualität dieser Embeddings** ist von enormer Wichtigkeit, um im weiteren Verlauf gute Ergebnisse zu erzielen

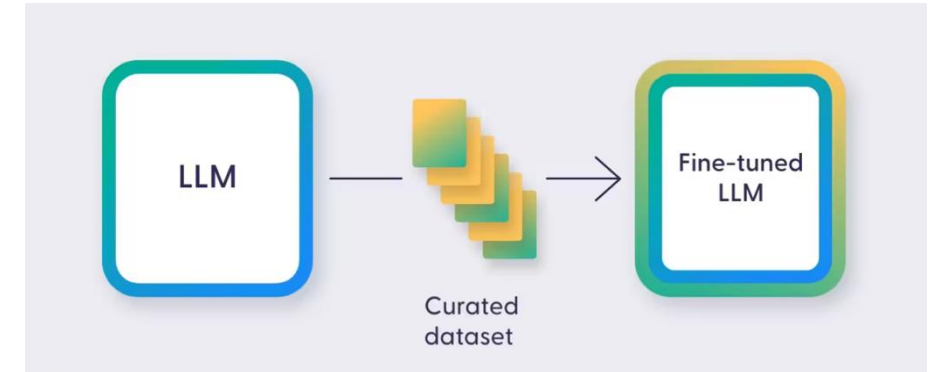


3D Darstellung von 768-dim Wort-Embeddings | Dimensionalitätsreduziert via TSNE

# Fine-Tuning vs. Retrieval Augmented Generation

## Fine-Tuning von LLMs

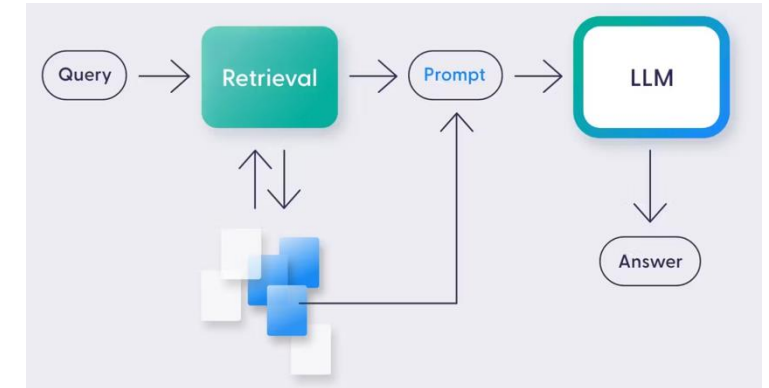
- Anstatt ein LLM von Grund auf zu trainieren, nehmen Sie ein vortrainiertes Modell und passen es an Ihren Anwendungsfall an.
  - Dadurch werden LLMs auf bestimmte Aufgaben oder Daten zugeschnitten, was ihre Leistung für bestimmte Szenarien verbessern kann
  - Kann mit Diensten wie [HF AutoTrain](#), [Ludwig](#) oder [OpenAI](#) durchgeführt werden
- Zwei Hauptziele bei der Feinabstimmung:
  - **Domänenspezifisch:**  
Abstimmung auf eine bestimmte Sprachdomäne oder einen bestimmten Ton, z. B. kann die interne Wissensbasis eines Pharmaunternehmens eine andere Terminologie verwenden als die Trainingsdaten des LLM
  - **Aufgabenspezifisch:**  
Verbesserung der Leistung bei einer bestimmten Aufgabe wie Klassifikation, NER oder Regression, z. B.  
Sentimentanalyse: Feinabstimmung eines Sprachmodells, um die Stimmung von Kundenrezensionen besser zu verstehen und zu klassifizieren, wobei zwischen positiven, negativen und neutralen Meinungen unterschieden wird:  
*Dies könnte die Gesamtfähigkeiten des LLM einschränken - "Ausrichtungssteuer"*



# Fine-Tuning vs. Retrieval Augmented Generation

## Retrieval Augmented Generation (RAG)

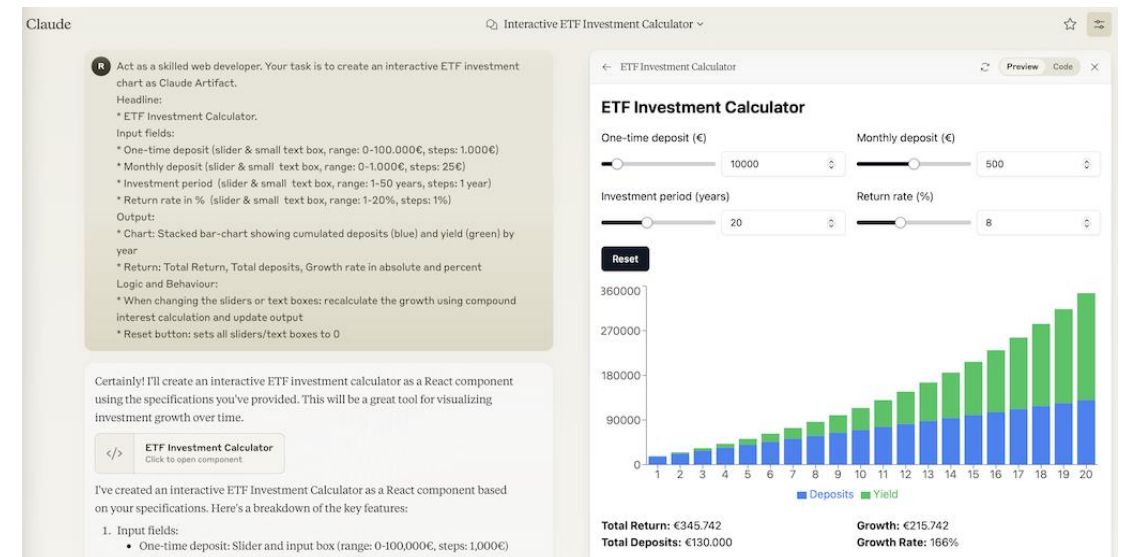
- RAG kombiniert ein Retrievalsystem mit einem Sprachgenerierungsmodell
  - Die Retrieval-Komponente ist meist eine Vektordatenbank
  - Sie sucht nach Textausschnitten, die für die Eingabeanfrage relevant sind, indem sie sowohl die Anfrage als auch die Dokumente in Vektordarstellungen umwandelt und diese vergleicht (Embeddings)
- Die abgerufenen Informationen werden in ein LLM eingespeist – es nutzt den zusätzlichen Kontext, um seine Antwort zu verbessern
  - Dies macht die Antwort genauer und kontextspezifischer.
- RAG ist nützlich bei Aufgaben, bei denen externes Wissen unerlässlich ist
  - Beantwortung komplexer, domänenspezifischer Fragen
  - Inhaltserstellung mit faktischer Genauigkeit (direkte Referenzen)
  - Szenarien, in denen der gewählte LLM möglicherweise nicht über ausreichenden Kontext oder aktuelle Informationen verfügt



# Code Generation mit LLMs

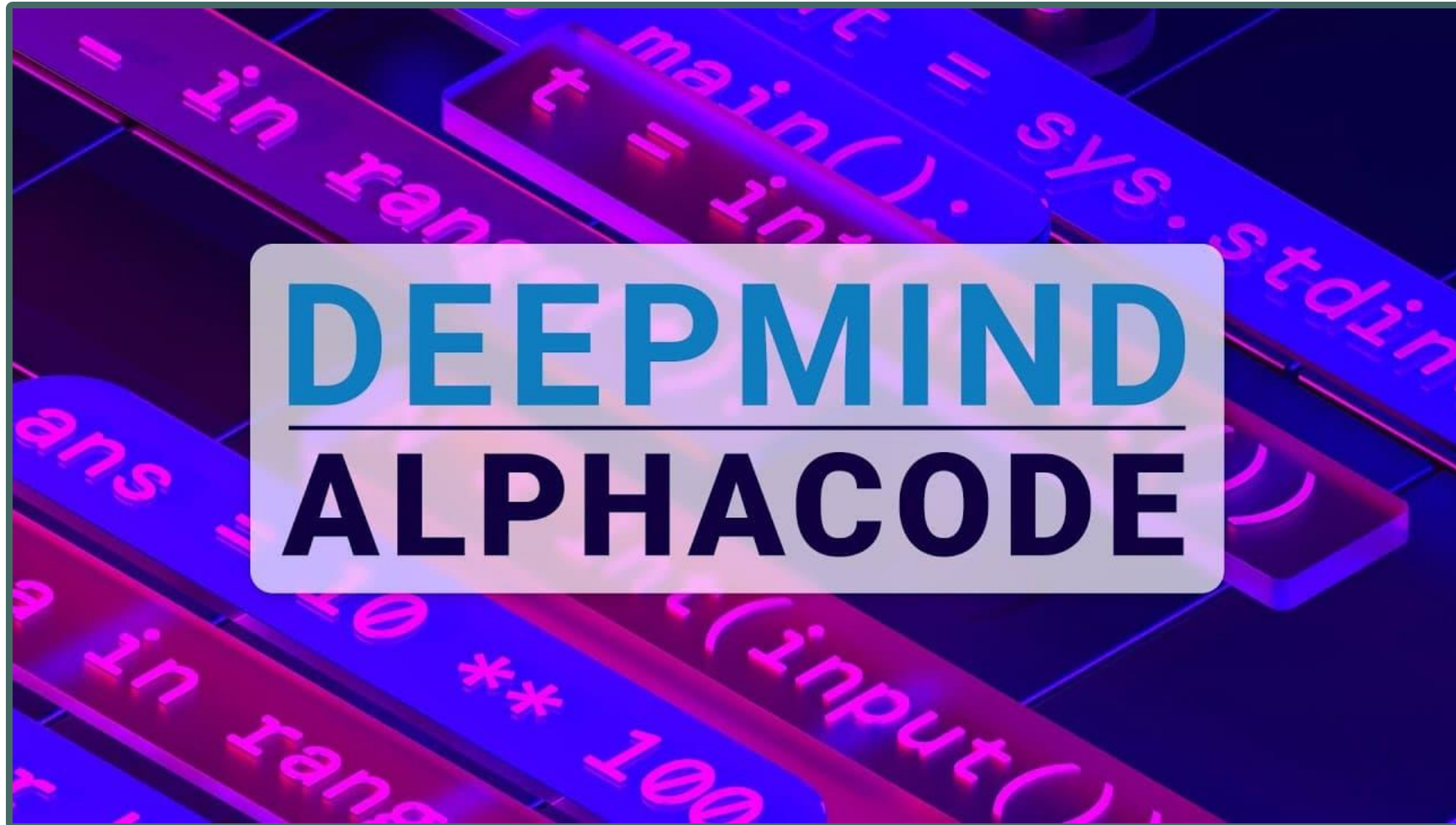
- **Chancen – Effizienz und Produktivität:**  
KI-gestützte Code-Generierung beschleunigt die Softwareentwicklung und automatisiert Routineaufgaben
- **Risiken – Halluzinationen:**  
KI-Modelle können syntaktisch korrekten aber faktisch falschen Code generieren, der fehlerhaft ist und Probleme verursacht, insbesondere bei komplexen Anforderungen.
- **Veränderte Lernkurve:**  
KI kann Entwicklern beim Lernen neuer Programmiersprachen oder Frameworks helfen, birgt jedoch das Risiko, dass Entwickler sich zu stark auf die KI verlassen und ihre eigenen Fähigkeiten vernachlässigen.

## KI-Tools zum Programmieren



Claude Artifacts

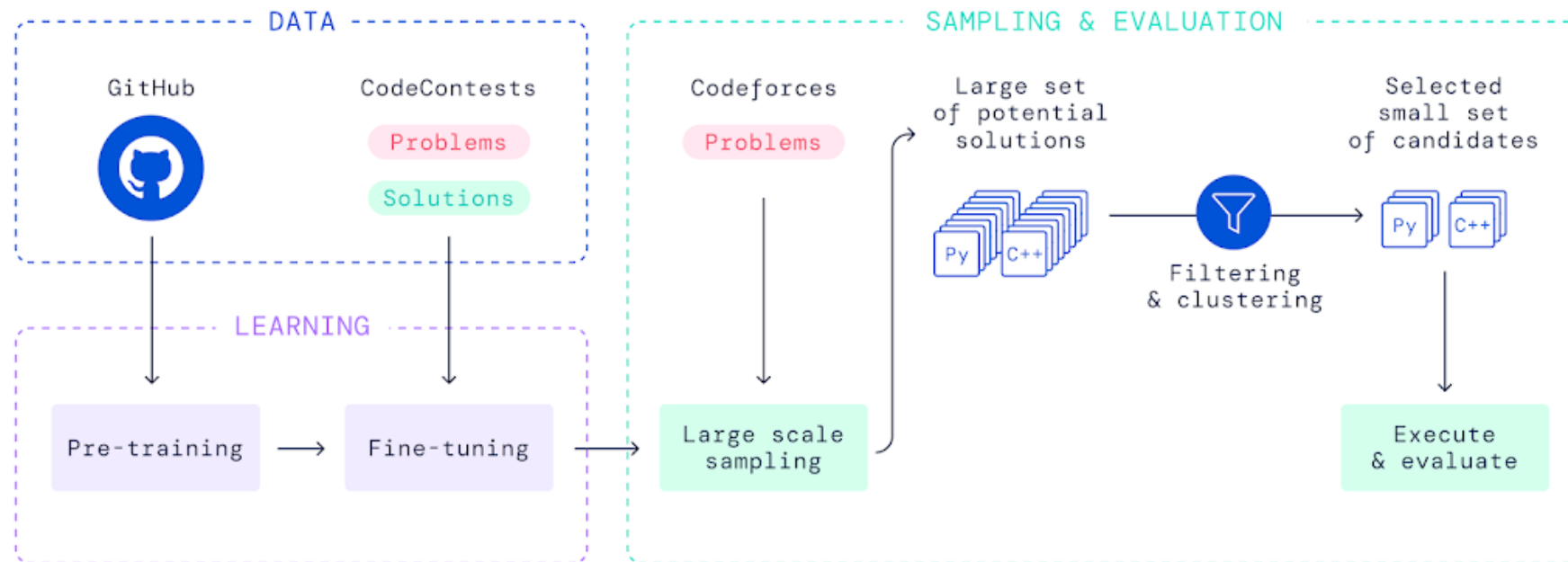
# AlphaCode – Wie LLMs programmieren:



<https://alphacode.deepmind.com/>

# Code Generation mit LLMs: Halluzinationen durch Training-Daten

- Spezialisierte LLMs werden wie allgemeine Modelle auch auf massiven Datenmengen trainiert
- Code LLMs nehmen hierzu gerne GitHub (als Basis) und StackOverflow (für Frage/Antwort Paare)
- Diese Daten können bereits Fehler enthalten, welche im LLM-Training als faktisch korrekt behandelt werden



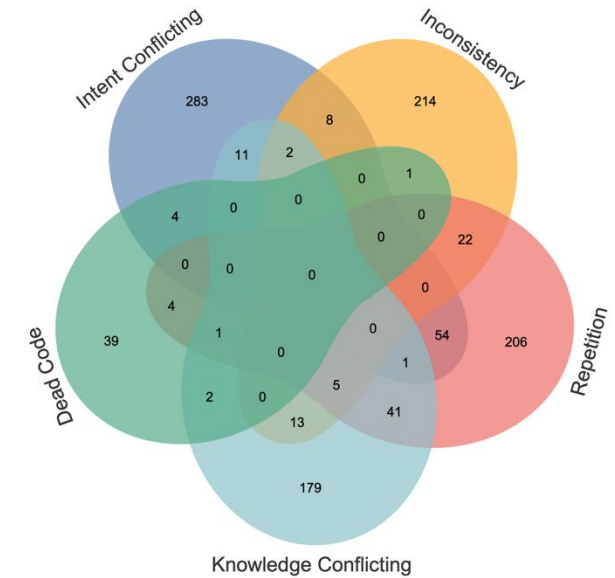
[DeepMind AlphaCode Blog Post](#)

# Code Generation mit LLMs: Welche Fehler passieren?

# Exploring and Evaluating Hallucinations in LLM-Powered Code Generation

(<https://arxiv.org/pdf/2404.00971v1>)

- Ein neues Paper behandelt die **Arten von Fehlern** bei LLM Code Generierung und entwickelt **Methoden zur Reduzierung** dieser Fehler
- Je nach Modell werden andere Arten von Fehlern erzeugt (bspw. ChatGPT: Knowledge Conflict & Inconsistency)
- Die Autoren betonen die Notwendigkeit von sog. **“Fine-Tuning”** auf bestimmte **Programmiersprachen** oder **Anwendungsdomänen** zur Reduzierung von Halluzinationen
- Das Paper schlägt darüber hinaus vor, **Wissensgraphen** in Sprachmodelle zu integrieren, um die Halluzinationen vorzubeugen



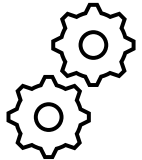
```
def filter_by_substring(strings: List[str], substring: str)
-> List[str]:
    """ Filter an input list of strings only for
        ones that contain given substring
    >>> filter_by_substring([], 'a')
    []
    >>> filter_by_substring(['abc', 'bacd', 'cde',
        'array'], 'a')
    ['abc', 'bacd', 'array']
    """
    # generated by CodeRL
    return [string for string in strings if
            string.startswith(substring)]

    # reference solution
    return [x for x in strings if substring in x]
```

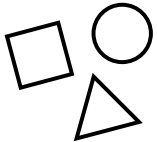
# Was wird Sie heute erwarten?



Praktisches Wissen **zum effizienten Prompten** verschiedener KI-Systeme und Maximierung der Ergebnisqualität



Grundlegendes Wissen um den **Aufbau von Large Language Modellen** und die Einordnung in den aktuellen Technologie-Kontext



Verständnis über die aktuelle Service-Landschaft rund um **Text-, Bild- und Videogenerierung mittels KI** & Nutzung von Reasoning-Systemen



Strategien zur Integration von KI-Assistenten und -Agenten in bestehende Prozesse sowie konkrete Methoden zur Entwicklung innovativer KI-basierter Geschäftsmodelle