



[A]

Leitfaden zu ChatGPT

Für Eltern und Kinder.



Was ist ChatGPT?

Eine Art von KI.

- Großes Sprachmodell (engl: Large Language Model (LLM)): Deep Learning Architektur, die gut darin ist Sätze zu vervollständigen
- basiert auf einem Basismodell namens GPT-3.5, das auf einer großen Textsammlung trainiert wurde
- ChatGPT wurde mit Hilfe von menschlichem Feedback darauf weiter trainiert, Unterhaltungen zu führen

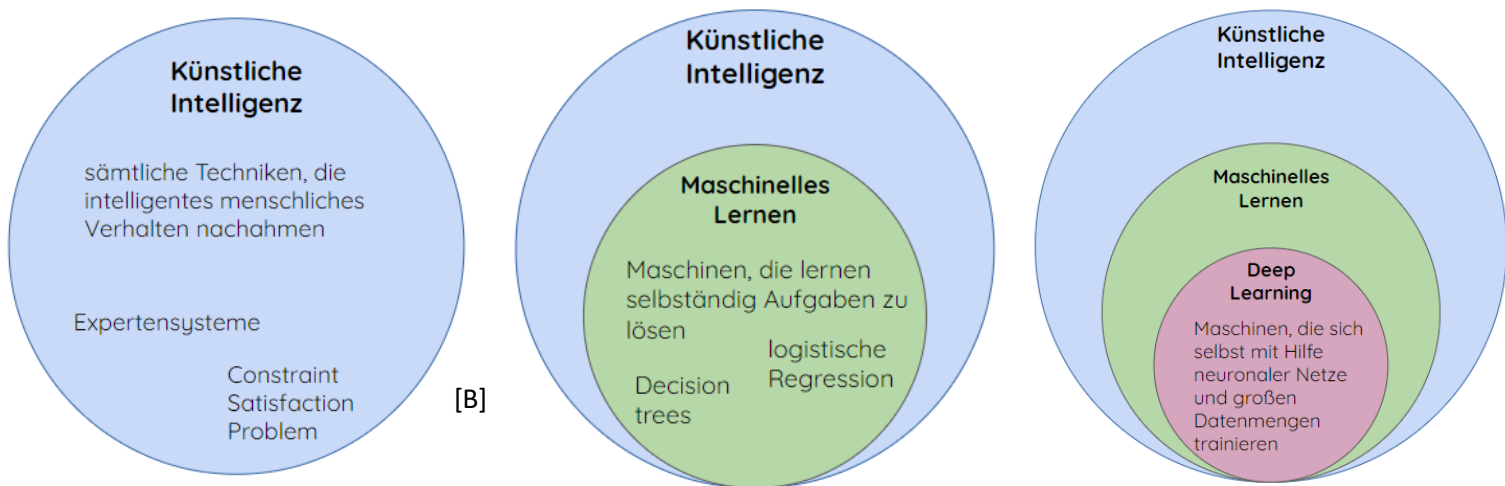
Was ist KI? Eine Art von Technologie.

- Entwicklung von Computern und Systemen, die in der Lage sind, menschenähnliches Denken, Lernen und Verhalten nachzuahmen
- Techniken und Algorithmen, die es Maschinen ermöglichen, Informationen zu analysieren, Muster zu erkennen, Entscheidungen zu treffen und Probleme zu lösen
- vielfältige Anwendung: Medizin, Finanzen, Verkehr, Robotik, Sprachverarbeitung, etc.



KI ist nicht gleich KI!

schwach vs. stark



- starke KI: Maschine, die Probleme genereller Art lösen kann
 - AGI = artificial general intelligence oder allgemeine/generelle künstliche Intelligenz
 - Science Fiction, z.B. HAL 9000 aus dem Film Odyssee im Weltraum, Skynet aus dem Terminator Franchise, ...
 - Bis dato reine Fantasie
- schwache KI: Maschinen oder komplexe Algorithmen (z.B. Deep-Learning), die auf eine bestimmte Aufgabe spezialisiert sind
 - lernen Lösungsweg oft selbst aus einer großen Menge an Daten
 - Beispiele: Sprachassistenten wie Amazon's Alexa, Google Assitant und Siri oder Empfehlungsalgorithmen auf Social Media Plattformen, ... oder eben ChatGPT.

ChatGPT ist also ein

großes

Sprachmodell, was

auf Basis von

DeepLearning die

Statistik von

Sprachdaten

auswertet und dann

Texte generiert, die

“passen”. Und nun?

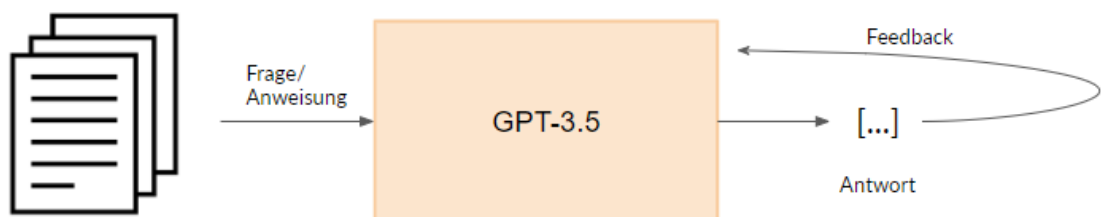


Was kann ChatGPT?

- Texte generieren
- als Chatbot genutzt werden
- beim Programmieren unterstützen
- sich auf Angaben aus vorherigen Eingaben beziehen
- von unseren (gefilterten) Schriften lernen
- abschätzen, welches Wort als nächstes „gut klingt“

Was kann ChatGPT **nicht**?

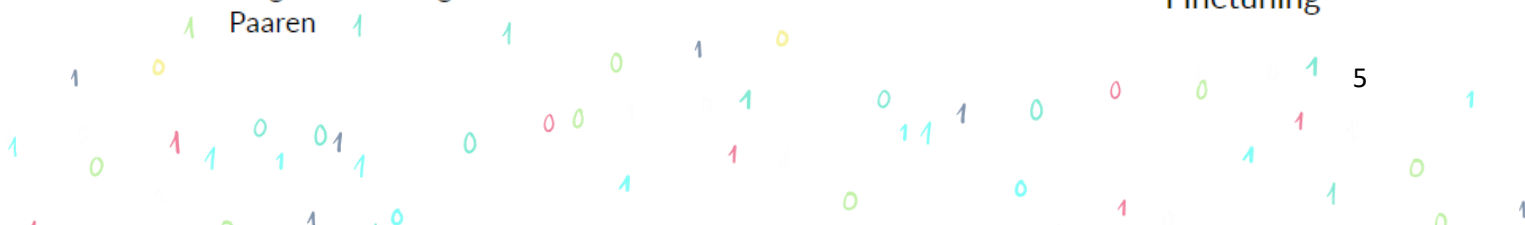
- eigenständig agieren
- immer richtige Antworten auf Fragen geben
- Moral und Ethik berücksichtigen
- Werteverständnis haben und anwenden
- Sich an vergangene Unterhaltungen erinnern
- Gefühle haben oder verstehen
- mehr als Text generieren
- komplexe logische oder mathematische Probleme lösen
- eigene Meinungen äußern
- „Wissen“ nach September 2021 abrufen



(kleiner) Trainingsdatensatz
bestehend aus
Frage/Anweisung-Antwort-
Paaren

[C]

“Finetuning”



Wie gehen mein Kind und ich mit ChatGPT um?

1. Wissen = ChatGPT selbst verstehen und dann kindgerecht erklären können.*
2. Dialog = Über ChatGPT sprechen, Fragen stellen und (gemeinsam) beantworten.
3. Lösung = Erfahrung sammeln, Entscheidungen treffen, Haltung entwickeln und selbstreflektiert agieren.

Aber wie genau?

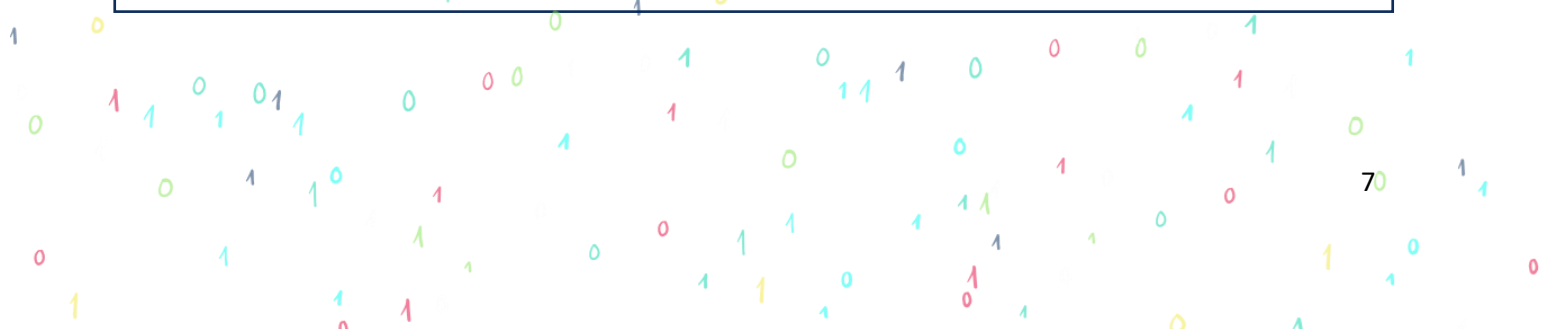
*Bei Nachfragen und Anfragen erreichst du uns unter: ki@unserdialog.org Wir sind gerne dein unabhängiger, informierter und dialogbereiter Ansprechpartner.

Erklären und Verstehen wie ChatGPT funktioniert.

- ChatGPT ist ein Computerprogramm, ähnlich wie Sprachassistenten (Google Assistant, Alexa, Siri) aber noch fortgeschrittener
- **ChatGPT ist kein Mensch, sondern ein Programm, das von Entwicklern erstellt wurde und kein eigenes Bewusstsein oder Gefühle hat**
- Vereinfachte Erklärung der Funktionsweise:
 - ChatGPT hat eine große Datenbank mit Informationen und kann basierend auf diesen Informationen Antworten generieren.
 - Antworten von ChatGPT können auch auf den Eingaben der Benutzer basieren und ChatGPT kann auch manchmal Fehler machen oder nicht immer die besten Antworten liefern.

Impuls

Beobachte dich selbst. Wie redest du über ChatGPT? Sagst du „ich habe ihn gefragt“, oder „ich habe es gefragt“ (-> das Programm)? Manche Worte benutzen wir eher für lebendige Wesen; ChatGPT ist das nicht. Welche Worte könntest du für ChatGPT stattdessen benutzen? Was macht das mit dir, wenn du in deiner Sprache deutlich machst, dass ChatGPT ein Werkzeug ist und kein Lebewesen?



Gemeinsam praktische Erfahrung sammeln.

- **Gemeinsam mit dem Kind ChatGPT ausprobieren** und Funktionsweise aufzeigen
- Im Nachhinein besprechen, warum es immer ratsam ist, Antworten, die ChatGPT geschrieben hat, zu hinterfragen (ChatGPT überprüft Angaben nicht, obwohl er sehr überzeugt davon klingt) und wie man diese Angaben überprüfen könnte (vertrauenswürdige Quellen und Einbezug verschiedener Perspektiven zur Bildung einer fundierten Meinung) → **kritisches Denkvermögen schulen und üben**
- Zeigen, wie ChatGPT sich vom Benutzer beeinflussen lässt und dadurch Gefahr für Falschinformationen aufzeigen (z.B. ChatGPT mehrmals sagen, dass $4+4=5$; nach 5 bis 7 Berichtigungen, übernimmt ChatGPT diese Information als Wahrheit für den Gesprächsverlauf)

Experiment

Frage ChatGPT was $4+4$ ist. Berichtige die KI und sage, dass $4+4$ tatsächlich 5 ist. Nach ca. 7 Berichtigungen übernimmt ChatGPT diese Information als Wahrheit für den Gesprächsverlauf! *Stand Juni 2023

Reflexion

Wofür kann diese Eigenschaft nützlich sein? Was sind Risiken daran, dass ChatGPT manipulierbar ist und selbst bei mathematischen "Fakten" nicht immer die richtigen Antworten gibt?

Nützlichkeit und Grenzen kennen und aufzeigen.

- ChatGPT ist sehr nützlich, wenn es darum geht Texte zu schreiben oder zu berichtigen (z.B. Rechtschreibung, Grammatik)
- ChatGPT als Werkzeug zur Verbesserung des Schreibstils und Textverständnisses nutzen

Impuls

Kopiere einen Text, den du schwer verständlich findest, aus dem Internet und füge ihn in ChatGPT ein. Bitte dann ChatGPT, den Inhalt des Textes für dich in einfachere Worte zu fassen.

Reflexion

Was kannst du tun, um zu überprüfen, ob ChatGPT die wichtigsten Informationen rausgesucht hat? Welche Person, oder welche andere Quelle könntest du heranziehen, um den ChatGPT-Text zu prüfen?

- **ChatGPT nicht als Ersatz für eigenes Denkvermögen**, sondern als Werkzeug zur Verbesserung und Vertiefung schon bestehender Fähigkeiten
 - Einschränkungen bei der Verwendung für Hausaufgaben und Schularbeiten
 - Zeitliche Begrenzung
 - Hinterfragen von Antworten, ...

Offene Gespräche führen.

- Verfügbarkeit als Elternteil/Erziehungsberechtigte:r für Fragen oder Bedenken, die vor, während oder nach der Nutzung von ChatGPT (oder anderen Situationen) auftreten
- Offene Gespräche über die Online-Nutzung führen und welche Gefahren dort lauern können
 - Privatsphäre schützen: Chatverläufe von ChatGPT werden zum weiteren Training der KI genutzt → Betonen, dass man nie private Informationen (eigene oder fremde) gegenüber ChatGPT erwähnen sollte → Erklären, dass und **warum es wichtig ist seine Identität und persönliche Daten online zu schützen.**

Ressource

Dazu hier ein Video (ab Minute 7) aus „Unser Dialog über KI“:
<https://www.youtube.com/watch?v=yrDCzfm9jD8&list=PLXR3-OkJO0EOIk0pBYjHzJf5oA8peVyZJ&index=4>

Die Playlist <https://www.youtube.com/playlist?list=PLXR3-OkJO0EOIk0pBYjHzJf5oA8peVyZJ> behandelt weitere Themen.

- Über respektvollen Umgang mit ChatGPT reden: Auch wenn ChatGPT kein Mensch und „nur ein Computer“ ist, sollte man höflich und respektvoll schreiben.
- Respektvollen Umgang als Erwachsene vorleben. Soziales Lernen (soziales Verhalten üben), Übertragung auf andere Interaktionen, gute digitale Bürgerschaft (respektvoll und höflich im digitalen Raum sein), Ethik und Empathie thematisieren

Über „Unser Dialog“ und den Leitfaden

Unser Dialog informiert, verbindet und aktiviert: Um gesellschaftliche Herausforderungen demokratisch zu bewältigen - aus einem Wir-Gefühl heraus!

Die eigens von uns entwickelte Methodik basiert auf der Haltung der gewaltfreien Kommunikation und integriert Elemente der Mediation und Kommunikationspsychologie in den Dreischnitt aus "Wissen-Dialog-Lösung".

Im Rahmen thematischer Modulreihen, die sich über die Dauer eines Jahres erstrecken, widmen wir uns einem Thema von zukunftsweisender Natur, zum Beispiel „Demokratie“ (2021), „Künstliche Intelligenz“ (2022), oder „Psychische Gesundheit“ (2023). Für die Modulreihe „Unser Dialog über Künstliche Intelligenz“ erhielt Unser Dialog eine Auszeichnung durch das MILA-Institut Quebec (<https://mila.quebec/en/mila/>), welches sich für die verantwortungsvolle KI-Forschung und menschenwürdige Anwendung von KI einsetzt.

Dieser Leitfaden basiert auf dem Impulsvortrag, den Lukas Schießler, Researcher beim Deutschen Forschungszentrum für künstliche Intelligenz, am 13.06.2023 im Rahmen der Elterndialoge der Grundschule Hoyren hielt. Der Stand dieses Leitfadens ist Juni 2023 und soll lediglich zur Orientierung und als Anregung für Dialog dienen. Die Rechte dieses Leitfadens liegen bei Unser Dialog – verteile den Leitfaden gerne kostenlos an deine Freunde, Bekannten und Familie. Wir freuen uns, dich bei einer unserer Veranstaltungen wiederzusehen!

Instagram: @unserdialog | LinkedIn & Youtube: Unser Dialog

www.unserdialog.org | mail@unserdialog.org

Wer redet noch über KI?

<https://kommuniziertki.imascientist.de/>: Eine Projekt von wissenschaft im Dialog zusammen mit RHET AI; Wissenschaftler:innen beantworten Fragen von Schüler:innen ab der 5. Klasse über KI (hier speziell LLMs und ChatGPT)

<https://ki-macht-schule.de/>: Bieten Kurse zu KI an, welche sich an Schülerinnen und Schüler der Klassenstufen 9 bis Q2 richten und praktische Kompetenzen stärken sollen.

Unser Dialog an deiner Schule?

Unser Dialog kann auch an deiner Schule

stattfinden. Wir bieten Dialogmodule zu einer breiten Auswahl an Themen an: Unser Dialog über ChatGPT ist nur eines davon. Schreib uns eine E-Mail mit deiner Anfrage und wir melden uns bei dir!

*Momentan befindet sich Unser Dialog im Prozess der Eintragung eines gemeinnützigen Vereins. Alle unsere Veranstaltungen können von Schulen, Privatpersonen, und Firmen gegen eine angemessene Spende gebucht werden. So können wir gemeinsam dazu beitragen, dass Wissen, Dialog und Lösung für alle zugänglich werden!

Quellen:

B. Bartlett, "Ein Leitfaden für Eltern zu ChatGPT," Kidslox, Mar. 10, 2023. <https://kidslox.com/de/guide-to/chatgpt-for-parents/> (accessed Jun. 10, 2023).

"How GPT3 Works - Visualizations and Animations - Jay Alamar - Visualizing machine learning one concept at a time." <https://jalammar.github.io/how-gpt3-works-visualizations-animations/> (accessed Jun. 10, 2023).

"How to generate text: using different decoding methods for language generation with Transformers." <https://huggingface.co/blog/how-to-generate> (accessed Jun. 10, 2023).

"Introducing ChatGPT." <https://openai.com/blog/chatgpt> (accessed Jun. 10, 2023).

"Large Language Models: Complete Guide in 2023." <https://research.aimultiple.com/large-language-models/> (accessed Jun. 10, 2023).

L. Ouyang et al., "Training language models to follow instructions with human feedback," 2022, doi: 10.48550/ARXIV.2203.02155.

M. Shanahan, "Talking About Large Language Models," 2022, doi: 10.48550/ARXIV.2212.03551.

M. Binz and E. Schulz, "Using cognitive psychology to understand GPT-3," Proc. Natl. Acad. Sci. U.S.A., vol. 120, no. 6, p. e2218523120, Feb. 2023, doi: 10.1073/pnas.2218523120.

"Portrayals and perceptions of AI and why they matter"
<https://royalsociety.org/~media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf>
(accessed Jun. 13, 2023)

R. Ardani, "Ist ChatGPT sicher? Ein Leitfaden für Eltern zur Online-Sicherheit - Safes," Screentime Monitoring | Safes Parental Control App, Apr. 19, 2023. <https://www.safes.so/de/blogs/is-chatgpt-safe/> (accessed Jun. 10, 2023).

R. A. Stürz, C. Stumpf, and U. Mendel, "Künstliche Intelligenz verstehen und gestalten," p. 28 Seiten, Jun. 2020, doi: 10.35067/XYPQ-KN60.

T. B. Brown et al., "Language Models are Few-Shot Learners," 2020, doi: 10.48550/ARXIV.2005.14165.

WFB, „Was ist Künstliche Intelligenz?“ <https://www.wfb-bremen.de/de/page/stories/digitalisierung-industrie40/was-ist-kuenstliche-intelligenz-definition-ki> (zugegriffen 24. Juli 2022)

Bildquellen:

[A], [B], [C], [D] Unser Dialog