

# Wissenschaftliches Arbeiten mit KI

Lost in Dissertation

# Inhalt

- genKI-Grundlagen
  - Wie funktionieren Große Sprachmodelle
  - Prompten
- Übungen zum Prompten
- LLMs, Autorschaft und Verantwortung
- Urheberrechtliche Aspekte bei der Nutzung von LLMs
- Retrieval Augmented Generation und AI Research Assistants
- Anhang

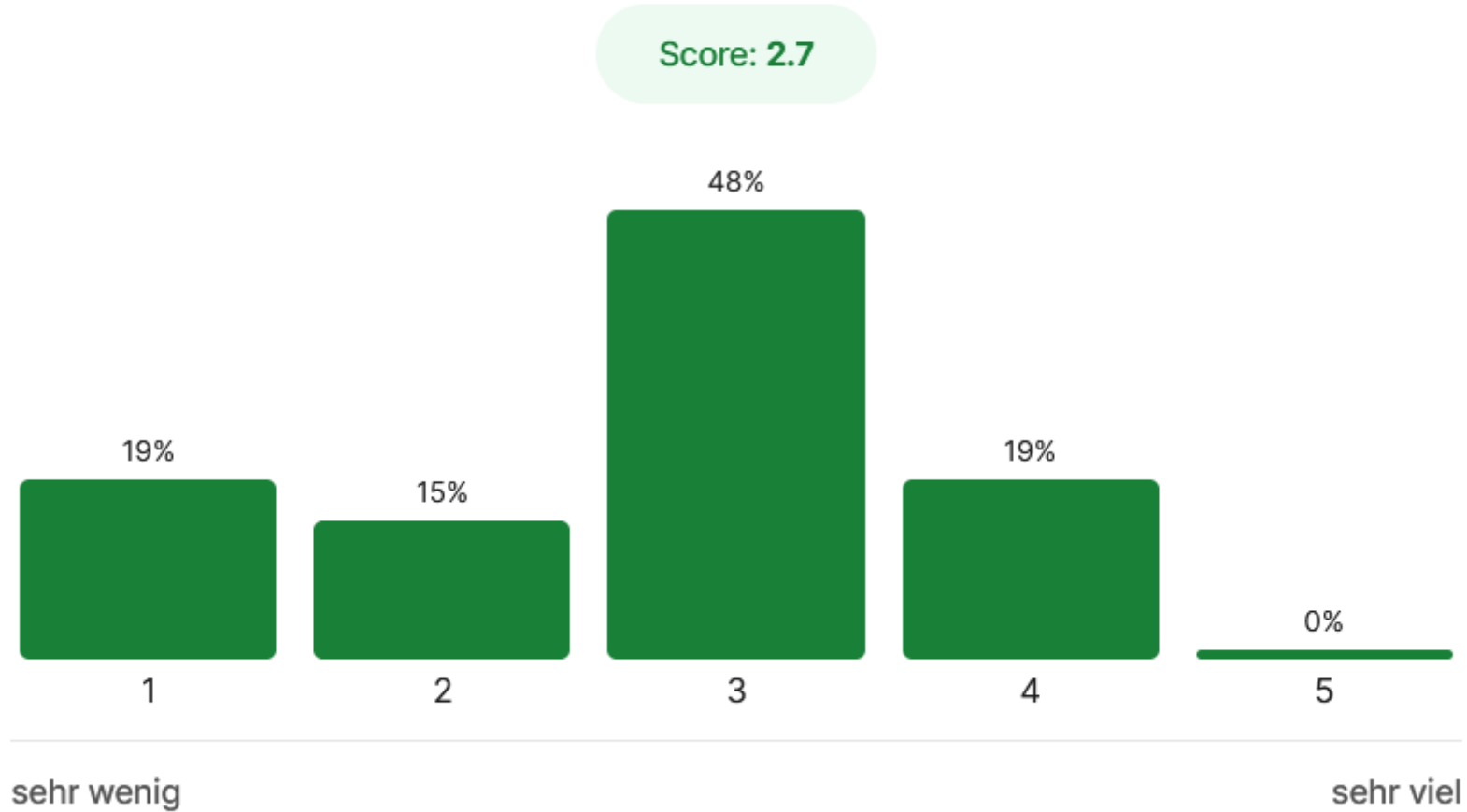
Ihre Erfahrungen mit generativer KI



## Wie viel wissen Sie über die Funktionsweise textgenerierender KI-Tools?

[Share](#) ▾

Rating Poll 27 votes 27 participants



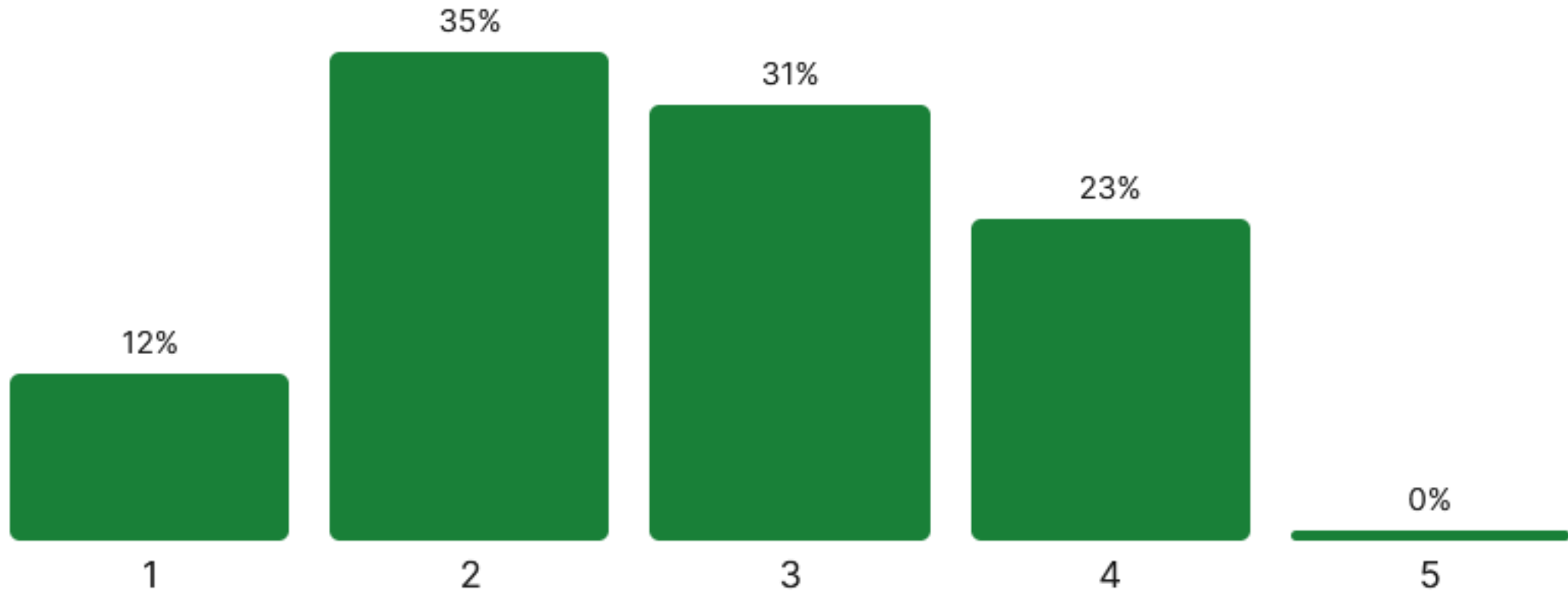


# Wie kompetent fühlen Sie sich in der Anwendung von KI-Tools?

Rating Poll  26 votes  26 participants

[Share](#) ▼

Score: 2.7





## Wofür haben sie KI-Tools bereits für Ihre wissenschaftliche Arbeit genutzt?

Share

Multiple Choice Poll 29 votes 29 participants

Brainstorming - 19 votes



Themenfindung/Formulierung einer Forschungsfrage - 9 votes



Gliederung - 10 votes



Textüberarbeitung/Redigieren - 19 votes



Zeitplanung - 3 votes



Übersetzen - 17 votes



Wissenschaftliche Recherche - 10 votes



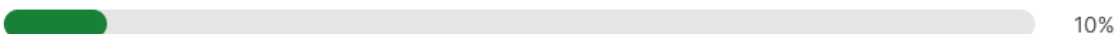
Zusammenfassung von Texten - 16 votes



Textproduktion - 9 votes



Sonstiges - 3 votes



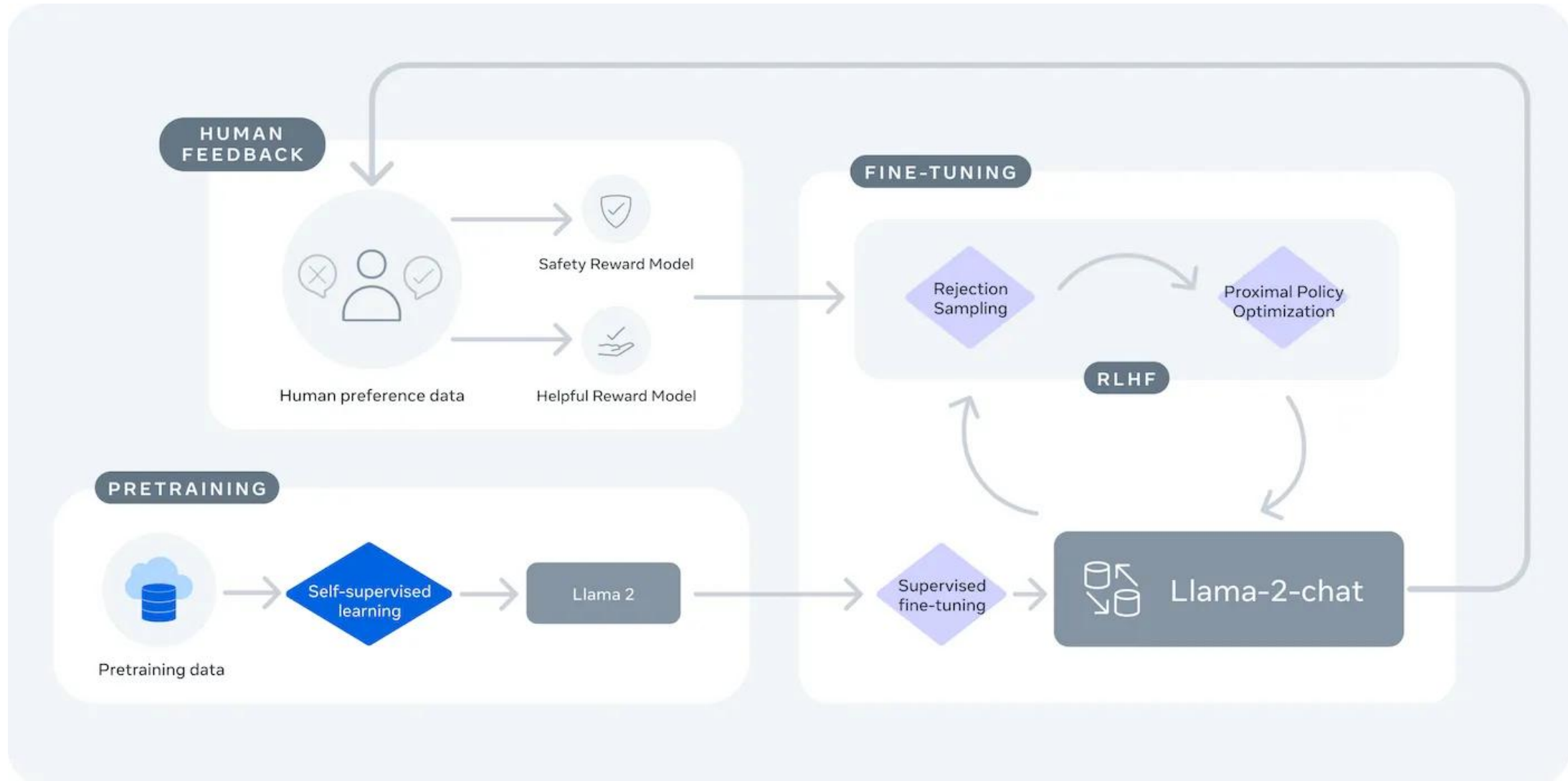
# Wie funktionieren Große Sprachmodelle (Large Language Models, LLMs)

# Input -> **Blackbox** -> Output

- Alle modernen Large Language Models (wie ChatGPT) basieren auf der Transformerarchitektur\* und führen von einer Texteingabe ausgehend Textoperationen durch

\* | Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser und Illia Polosukhin. „Attention Is All You Need“, 2017. <https://doi.org/10.48550/ARXIV.1706.03762>.

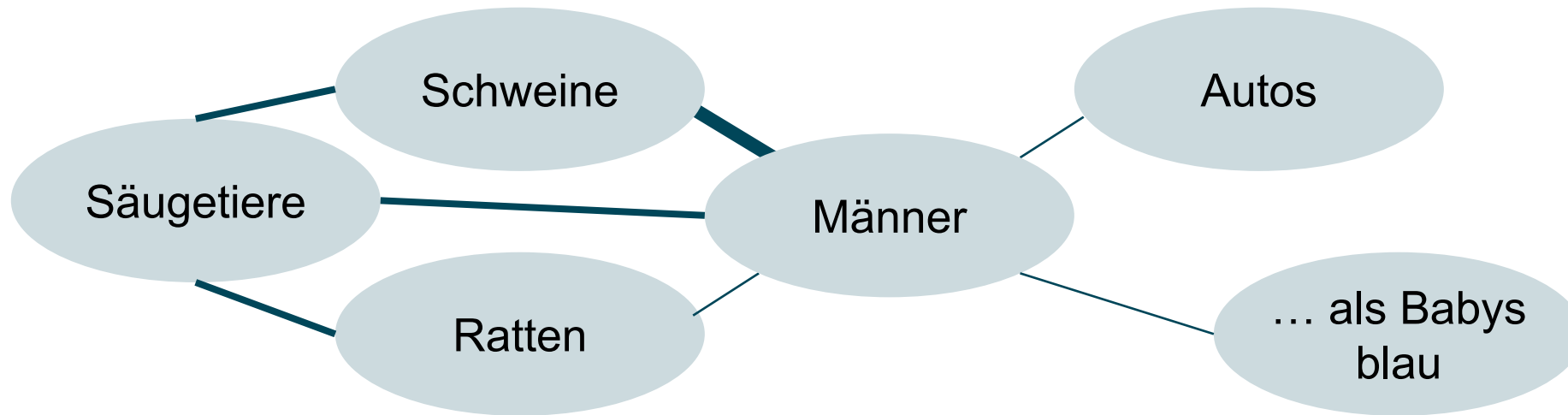
# Input -> **Blackbox** -> Output



\* | <https://llama.meta.com/llama2/>

# Input -> **Blackbox** -> Output

- Alle modernen Large Language Models (wie ChatGPT) basieren auf der Transformerarchitektur\* und führen von einer Texteingabe ausgehend Textoperationen durch
- Transformermodelle basieren auf einer neuronalen Netzwerkstruktur+



\* | Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser und Illia Polosukhin. „Attention Is All You Need“, 2017. <https://doi.org/10.48550/ARXIV.1706.03762>.

+ | Vgl. IBM. O.J. Was sind neuronale Netze? <https://www.ibm.com/de-de/topics/neural-networks>.

# Input -> **Blackbox** -> Output

- ChatGPT ist wie alle modernen Large Language Models ein Transformermodell\*, das von einer Texteingabe ausgehend Textoperationen durchführt
  - Transformermodelle basieren auf einer neuronalen Netzwerkstruktur\*
  - Die Texterzeugung folgt einer Wahrscheinlichkeitsheuristik
- Die Textproduktion ist i.d.R. nicht reproduzierbar (→ Ausnahme: Deterministische Modelle)
- Die Textproduktion beruht auf Wahrscheinlichkeit und wird durch die Trainingsdaten vordeterminiert (→ Stichwort: Halluzinieren | → Stichwort: Confirmation Bias)
- Das auf eine konkrete Anfrage (Prompt) erwartbare Output wird durch den Prompt begrenzt (→ Stichwort: Promptingstrategien | → Persönlichkeits-, Urheber- und Lizenzrechte)

# Glossar – Auswahl

- NLP = Natural Language Processing = Sprachverarbeitung natürlicher Sprache
- GPT = Generative Pretrained Transformer
- Generative = etwas Neues entsteht (Bild, Text, Ton ...)
- Pretrained = Modell lernt selbst Struktur von Sprache
- Transformer = Softwarearchitektur, auf der die großen Sprachmodelle aufgebaut sind
- Großes Sprachmodell = Software, die sprachliche Daten verarbeitet, um neuen Text zu generieren
- Trainingsdaten = frei verfügbare Texte aus dem Internet (Webseiten, E-Books, Onlinezeitschriften ...)
- Finetuning = Änderung des Outputs, damit verändern sich die Parameter/Wahrscheinlichkeiten zugunsten der Änderung
- Parameter = Wert, der beim Training gelernt wird
- RLHF = reinforcement learning from human feedback = Training durch menschliches Feedback auf die Ausgaben, damit “lernt” der Algorithmus, welche Ausgaben mehr “Erfolg” haben
- Prompts = Eingaben

KI-Glossar: <https://www.ki.nrw/ki-schluesselebegriffe/>

Sprachmodell  
= Large Language Model  
= LLM

Retrieval-Augmented  
Generation  
= RAG

Übersetzung und einsprachige  
Textbearbeitung

Beispiel

GPT's von ChatGPT, Gemini, Claude

Perplexity, CoPilot, NotebookLM, PrimoAI

DeepL (Write), Language Tool, Grammarly

Aufgabe

Ein Chatbot, der Antworten in menschlicher Art und Weise geben soll.

Sprachmodell kombiniert mit Web- oder Datenbanksuche

Übersetzung und einsprachige Textbearbeitung

Technische Grundlage

Ein Sprachmodell ist eine Software, die große Textmengen verarbeitet, diese in kleinere sprachliche Segmente (Token) aufteilt, in Zahlen umwandelt und berechnet, wie häufig und in welchem Zusammenhang diese Einheiten miteinander auftreten. Je häufiger Informationen im Trainingsmaterial vorhanden waren, desto wahrscheinlicher werden diese bei einer Anfrage korrekt wiedergegeben. Der spezielle Aufbau der Software (Transformer) sorgt dafür, dass das Token mit seinem Kontext rechts und links analysiert wird und nicht linear wie früher der Google Übersetzer.

Daten-  
grundlage

Webseiten, Bücher und Artikel, Inhalte aus Sozialen Medien, Blogs, Foren, Wikipedia usw

Sprachmodell plus Web- oder Datenbankergebnisse

Spezifische Sprachinhalte wie Übersetzungen usw.

Grenzen

Keine Datenbank!!  
Veröffentlichungszeitraum und Zusammensetzung des Trainingsmaterials beeinflusst Aktualität des Outputs

Kenntnisse in Quellenbewertung notwendig. Das Relevanzranking bei KI-Browsern wie Perplexity und Copilot ist nicht transparent und die Quellen oft höchstens zur ersten Orientierung geeignet. Deep Research hat oft intransparente Datenbasis. Erst selbst erstellte RAG-Systeme sind großer Mehrwert.

Output wird tlw. Als hölzerner als zB ChatGPTs Output empfunden, was am engeren Zuschnitt liegt

Prompten

# Prompten – Clear Framework von Lo 2023

- **C**oncise – Prägnant: Kurze und klare Eingabeaufforderungen
- **L**ogical – Logisch: Aufbau der Ausgabe vorgeben
- **E**xplicit – Explizit: Ausgabe wie Textlänge und Fokus des Inhalts klar spezifizieren
- **A**daptive – Adaptiv: Eingabeaufforderungen bei Bedarf anpassen
- **R**eflective – Reflektierend: Eingabe und Ausgabe bewerten und auswählen bzw. kombinieren für das geeignetste Ergebnis
- **Hilfreich**: dem ChatBot eine Rolle zuweisen und eine Zielgruppe ansprechen

Leo S. Lo (2023) The Art and Science of Prompt Engineering: A New Literacy in the Information Age, Internet Reference Services Quarterly, DOI: 10.1080/10875301.2023.2227621

# Promptmethoden

- Zero-Shot-Prompt → Einfache Anweisung  
„Erkläre mir künstliche Intelligenz aus der Sicht eines Informatikprofessors“
- One-Shot/Few-Shot-Prompt → Anweisung mit einem Beispiel, wie der Output aufgebaut sein soll, welche Sprache genutzt werden soll usw. Eignet sich insbesondere, wenn Sie Schulungsbeschreibungen o.Ä. erstellen (Beispiel in der Übung)
- Many-Shot-Prompt → Es werden sehr viele Beispiele mitgegeben, aus denen das KI-System Muster ableitet und entsprechend auf den Output anwendet
- Chain-of-Thought-Prompting → Anweisung, die konkrete Bearbeitungsschritte vorgibt  
„Ich erforsche den Einfluss von [Begriff] auf [Thema/Begriff]. Führe mich durch einen strukturierten Denkprozess, um relevante Theorien zu identifizieren. Beginne mit einer Analyse des Forschungsfeldes, identifiziere dann theoretische Ansätze aus verschiedenen Disziplinen, bewerte ihre Anwendbarkeit auf mein Thema und schließe mit einer Empfehlung der vielversprechendsten Theorien ab.“

# Übungen zum Prompten

# Brainstormen einer Forschungsfrage

Geben Sie folgenden Prompt beim Textgenerierungstool Ihrer Wahl ein und passen ggf. an:

Sie sind der Hochschullehrender an einer Hochschule und dafür verantwortlich, dass Promovierende aktuelle und in sich konsistente Themendispositionen für die Dissertation finden. Sie sind sehr streng und wollen von dem Promovierenden vor der Freigabe einige Fragen beantwortet haben. Sie stellen mir kritische Fragen zu den folgenden Bereichen: Themengebiet, Problemstellung, Forschungslücke, Ziele und Nicht-Ziele, wissenschaftliche Fragestellung, theoretischer Hintergrund, geplante Forschungsmethode, zu erwartende Ergebnisse. Ich bin der Promovierende und beantworte Ihre Fragen. Sie stellen jeweils eine Frage und stellen auf der Grundlage meiner Antwort die nächste Frage. Ich bekomme nur eine Zusage zur Themendisposition von Ihnen, wenn alle meine Antworten in sich konsistent sind und eine klare Vorgehensweise erkennbar ist.

# Erstellen einer Suchmatrix

Für eine Literaturrecherche zum Thema ["THEMA EINFÜGEN"] möchte ich die Recherchemethode Blockbildung/Building Blocks einsetzen. Bitte erkläre mir zunächst worum es bei der Recherchemethode geht. Gehe dann nach und nach folgende Schritte durch:

Schritt 1: Identifiziere die Schlüsselbegriffe meines Themas

Schritt 2: Zerlege das Thema bzw. die Forschungsfrage anhand der Schlüsselbegriffe in Blöcke

Schritt 3: Finde für jeden Schlüsselbegriff Synonyme, verwandte Begriffe, Oberbegriffe, Unterbegriffe und die englischen Übersetzungen und stelle es in einer Tabelle dar (eine Spalte für jeden Block)

Schritt 4: Trunkiere alle in Schritt 3 gefundenen Begriffe sinnvoll in einer Tabelle. Benutze als Trunkierungszeichen einen Asterisken.

Schritt 5: Verknüpfe die Wörter jeder Spalte mit einem OR, jeder Block hat runde Klammern und zwischen den Klammern wird der boolesche Operator AND eingesetzt. Gib mir dafür einen Suchstring aus, der alles enthält.

# Textkritik

Überprüfe die Argumentation kritisch des folgenden Textabschnitts hinsichtlich der Stringenz und erläutere die Stellen, die unklar oder schwach in der Argumentation sind:  
[IHR TEXT]

# Weitere Tipps

Sie könne auch immer das hier voranstellen, dann erhalten Sie einen modifizierten Prompt, aus dem Sie u.a. ableiten können, wie das Modell den Arbeitsauftrag versteht (wo Sie evtl. auch noch nachschärfen müssen):

Analysiere zunächst den folgenden Prompt und schlage Verbesserungen oder Ergänzungen vor, damit die dort beschriebene Aufgabe optimal gelöst werden kann:

[PROMPT]

Für längere, strukturierte Prompts können Sie mit Elementen/Anlehnung aus/an Markdown oder Meta-Tag-Analogen arbeiten

**Meta-Tags:** z.B. **Rollenzuschreibungen und Kontext**

`<role>Du bist ein Experte für Arbeitsrecht</role>`

`<context>Für eine rechtsvergleichende Analyse zwischen Deutschland und Großbritannien nach dem Brexit</context>`

**Markdown z.B. Rollenzuschreibungen und Kontext + Detaillierte Arbeitsanleitung**

Du bist ein Experte für akademische Integritätsprüfung. Analysiere die beigefügte Hausarbeit für einen Prüfungsausschuss systematisch in zwei Schritten:

---

### SCHRITT 1: Analyse der Eigenständigkeit

[...]

---

### SCHRITT 2: Plagiatsprüfung

[...]

# Promptotheken

- Promptkatalog von KI-Campus und Hochschulforum Digitalisierung: [https://coda.io/d/Prompt-Katalog\\_dCKiMW9kP-4/Offener-Prompt-Katalog\\_suubDjAW](https://coda.io/d/Prompt-Katalog_dCKiMW9kP-4/Offener-Prompt-Katalog_suubDjAW)
- Schritt-für-Schritt-Anleitung der Uni Hohenheim: [https://digital.uni-hohenheim.de/fileadmin/einrichtungen/digital/Von\\_Null\\_auf\\_ChatGPT\\_-\\_Anleitung.pdf](https://digital.uni-hohenheim.de/fileadmin/einrichtungen/digital/Von_Null_auf_ChatGPT_-_Anleitung.pdf)
- Beispielprompts und Basics zum Prompten für Studierende der TH Augsburg: <https://www.tha.de/ChatGPT-fuer-Studierende.html>
- Promptsammlung für Studierende und Lehrkräfte: <https://www.aiforeducation.io/prompt-library> (Englisch)
- Promptsammlung für Studierende und Lehrkräfte: <https://github.com/microsoft/prompts-for-edu> (Englisch)

# LLMs, Autorschaft und Verantwortung

# DFG-Leitlinie 14: Autorschaft

„Autorin oder Autor ist, wer einen genuinen, nachvollziehbaren Beitrag zu dem Inhalt einer wissenschaftlichen Text-, Daten- oder Softwarepublikation geleistet hat. [...]. Sie tragen für die Publikation die gemeinsame Verantwortung, es sei denn, es wird explizit anders ausgewiesen.“\*

- Für LLM-generierte Texte kann keine Autorschaft des LLMs angenommen werden.  
→ Daher auch nicht plagiatfähig

\* | Deutsche Forschungsgemeinschaft. 2019. *Leitlinien zur Sicherung guter wissenschaftlicher Praxis: Kodex*. Bonn: DFG.  
<https://doi.org/10.5281/zenodo.3923601>

# DFG-Leitlinie 14: Autorschaft

„Autorin oder Autor ist, wer einen genuinen, nachvollziehbaren Beitrag zu dem Inhalt einer wissenschaftlichen Text-, Daten- oder Softwarepublikation geleistet hat. [...]. Sie tragen für die Publikation die gemeinsame Verantwortung, es sei denn, es wird explizit anders ausgewiesen.“\*

- Für LLM-generierte Texte kann keine Autorschaft des LLMs angenommen werden.  
→ Daher auch nicht plagiatfähig
- Generieren LLMs Fehlinformationen, Falschangaben oder (in seltenen Fällen) wörtliche Textplagiate liegt die Verantwortung bei der Person, die diese Texte verwendet (und allen Mitautor\*innen → author's contributions section)
- Urheberrechtlich geschützte Texte dürfen nicht ohne weiteres (per Prompting) an ein LLM übergeben werden

\* | Deutsche Forschungsgemeinschaft. 2019. *Leitlinien zur Sicherung guter wissenschaftlicher Praxis: Kodex*. Bonn: DFG.  
<https://doi.org/10.5281/zenodo.3923601>

# Stellungnahme des Präsidiums der DFG zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG

- „Transparenz und Nachvollziehbarkeit des Forschungsprozesses und der gewonnenen Erkenntnisse für Dritte sind wesentliche Grundprinzipien wissenschaftlicher Integrität. Dieses Wertesystem bietet im Hinblick auf den Umgang mit generativen Modellen weiterhin wertvolle Leitlinien.“
- „Es entspricht dem Berufsethos von Wissenschaftlerinnen und Wissenschaftlern, dass sie selbst für die Einhaltung der Grundprinzipien wissenschaftlicher Integrität einstehen. Der Einsatz generativer Modelle kann Wissenschaftlerinnen und Wissenschaftler von dieser inhaltlichen und formalen Verantwortung nicht entbinden.“

\* | <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf>

# Stellungnahme des Präsidiums der DFG zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG

- „Wissenschaftlerinnen und Wissenschaftler sollten bei der öffentlichen Zugänglichmachung ihrer Ergebnisse im Sinne wissenschaftlicher Integrität offenlegen, **ob** und **welche generativen Modelle** sie **zu welchem Zweck** und **in welchem Umfang** eingesetzt haben.“
- „In wissenschaftlichen Publikationen können nur die verantwortlich handelnden natürlichen Personen als Autorinnen und Autoren in Erscheinung treten. Sie müssen **sicherstellen, dass durch die Verwendung generativer Modelle kein fremdes geistiges Eigentum verletzt wird** und kein wissenschaftliches Fehlverhalten etwa in Form von **Plagiaten** entsteht.“

\* | <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf>

# Stellungnahme des Präsidiums der DFG zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG

- „Daraus folgt nach aktueller Einschätzung, dass der Einsatz von generativen Modellen bei der Antragstellung bei der DFG im Prozess der Begutachtung, Bewertung und Entscheidung als solcher grundsätzlich weder positiv noch negativ zu bewerten ist.“
- „Bei der Erstellung von Gutachten ist der Einsatz von generativen Modellen mit Blick auf die Vertraulichkeit des Begutachtungsverfahrens unzulässig. Zur Begutachtung bereitgestellte Unterlagen sind vertraulich und dürfen insbesondere nicht als Eingabe für generative Modelle genutzt werden.“

\* | <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf>

# Grundregeln

- Besprechen Sie mit dem/der Betreuer\*in:
  - Welche Tools wollen Sie verwenden?
  - Wozu wollen Sie diese verwenden?
  - Wie wird die Verwendung der Tools dokumentiert?
- Ist vollständige Transparenz gewünscht, sollten dokumentiert werden:
  - Prompt
  - Output
  - Verwendung des Outputs
  - Hersteller des LLMs
  - Name des LLMs
  - Version des LLMs (wenn bekannt)

# Zitation von KI Tools

- APA
  - <https://apastyle.apa.org/blog/how-to-cite-chatgpt>
- Chicago
  - <https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation.html>
- MLA
  - <https://style.mla.org/citing-generative-ai/>
- DFG Pressemitteilung (Leitlinien sind unten verlinkt)
  - <https://www.dfg.de/de/service/presse/pressemitteilungen/2023/pressemitteilung-nr-39>
- Berlin Universities Publishing
  - <https://www.berlin-universities-publishing.de/ueber-uns/policies/ki-leitlinie/ki-handreichung/index.html>

# Urheberrechtlich Aspekte bei der Nutzung von LLMs

# Eigentum an den generierten Inhalten

- Autorschaft ist ein personenbezogenes Konzept, daher kommen LLMs oder LMMs (Large Multimodal Models) nicht als Autoren infrage
  - Generierte Inhalte sind somit urheberrechtsfrei (-> keine aus dem Urheberrecht begründete Kennzeichnungspflicht)
- In der Regel (-> AGB, Terms of Use) können Nutzende die generierten Inhalte frei – selbst kommerziell – verwenden
- Fraglich ist, ob für generierte Inhalte die Urheberschaft übernommen werden kann:
  - USA: LLM-generierte Texte sind weder die Texte einer dritten Person noch als eigenes Werk urheberrechtlich geschützt (so die US Copyright Authority 2023 im Fall Zarya of the Dawn; diskutierbar)

# Eigentum an den generierten Inhalten

The Office has completed its review of the Work's original registration application and deposit copy, as well as the relevant correspondence in the administrative record.<sup>1</sup> We conclude that Ms. Kashtanova is the author of the Work's text as well as the selection, coordination, and arrangement of the Work's written and visual elements. That authorship is protected by copyright. However, as discussed below, the images in the Work that were generated by the Midjourney technology are not the product of human authorship. Because the current registration for the Work does not disclaim its Midjourney-generated content, we intend to cancel the original certificate issued to Ms. Kashtanova and issue a new one covering only the expressive material that she created.

<https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>

# Prompting: Übertragung von Nutzungsrechten

## OpenAI (ChatGPT)

### Content

**Your content.** You may provide input to the Services (“**Input**”), and receive output from the Services based on the Input (“**Output**”). Input and Output are collectively “**Content**”. You are responsible for Content, including ensuring that it does not violate any applicable law or these Terms. You represent and warrant that you have all rights, licenses, and permissions needed to provide Input to our Services.

**Ownership of content.** As between you and OpenAI, and to the extent permitted by applicable law, you (a) retain your ownership rights in Input and (b) own the Output. We hereby assign to you all our right, title, and interest, if any, in and to Output.

**Similarity of content.** Due to the nature of our Services and artificial intelligence generally, Output may not be unique and other users may receive similar output from our Services. Our assignment above does not extend to other users’ output or any Third Party Output.

**Our use of content.** We can use your Content worldwide to provide, maintain, develop, and improve our Services, comply with applicable law, enforce our terms and policies and keep our Services safe.

**Opt out.** If you do not want us to use your Content to train our models, you have the option to opt out by updating your account settings. Further information can be found in [this Help Center article](#). Please note that in some cases this may limit the ability of our Services to better address your specific use case.

## Anthropic (Claude)

**Rights and Responsibilities.** You are responsible for all Inputs you submit to our Services and all Actions. You must ensure that you have all rights, licenses, and permissions that are necessary for us to process such Inputs in accordance with our Terms and to provide the Services to you, including for example, to integrate with third-party services, to share Materials with others at your direction, and to take Actions. You must ensure that your submission of Inputs to us, your sharing them with others, or directing Claude to take Actions will not violate our Terms, our [Acceptable Use Policy](#), or any laws or regulations applicable to those Inputs. As between you and Anthropic, and to the extent permitted by applicable law, you retain any right, title, and interest that you have in such Inputs. Subject to your compliance with our Terms, we assign to you all our right, title, and interest (if any) in Outputs.

**Our use of Materials.** We may use Materials to provide, maintain, and improve the Services and to develop other products and services, including training our models, unless you opt out of training through your account settings. Even if you opt out, we will use Materials for model training when: (1) you provide Feedback to us regarding any Materials, or (2) your Materials are flagged for safety review to improve our ability to detect harmful content, enforce our policies, or advance our safety research.

# Datenschutzrechtliche Aspekte bei der Nutzung von LLMs

# Eigene und personenbezogene Daten Dritter

- Die meisten LLMs werden auf Servern außerhalb der EU gehostet und sind daher nicht DSGVO-konform
- Datensparsamkeit mit Blick auf die eigenen Daten
- Personenbezogene Daten Dritter dürfen beim Prompten nicht verwendet werden, wenn die Nutzung dieser Daten durch den Anbieter des LLMs nicht zweifelsfrei ausgeschlossen werden kann oder die ausdrückliche Zustimmung der betroffenen Person(en) eingeholt wurde

# Retrieval Augmented Generation und AI Research Assistants

# Recherche mit einem Sprachmodell?



## GPT-4o-mini

Hier ist eine Beispiel-Bibliografie im Chicago Footnote-Stil zum Thema Künstliche Intelligenz (KI) im wissenschaftlichen Schreiben:

1. Bhatia, Rajiv. "Artificial Intelligence in Academic Writing: Opportunities and Challenges." *Journal of Writing Research* 12, no. 3 (2020): 345-367. <https://doi.org/10.17239/jwr.2020.12.03.01>
2. Elgammal, A., Liu, B., Elhoseiny, M., und M. M. D. A. "Can: Creative Adversarial Networks, Generating" Art by Learning About Styles and Deviating from Style Norms." *arXiv preprint arXiv:1706.07068* (2017). <https://arxiv.org/abs/1706.07068>.
3. Kuo, Chih-Hsiung, und Chih-Hsiu Chen. "The Impact of Artificial Intelligence on Academic Writing: A Review." *International Journal of Artificial Intelligence in Education* 30, no. 4 (2020): 607-623. <https://doi.org/10.1007/s40593-020-00213-0>.
4. Liao, Qiang, und Xiaofei Wang. "The Role of AI in Enhancing Research Writing: A Systematic Review." *Computers and Education* 149 (2020): 103804. <https://doi.org/10.1016/j.compedu.2020.103804>.
5. Smith, John. "AI-Assisted Writing Tools: A New Era in Academic Publishing." *Research in Higher Education* 61, no. 5 (2020): 745-762. <https://doi.org/10.1007/s11162-019-09563-0>.
6. Zhang, Wei, und Jie Liu. "Ethical Considerations of Using AI in Academic Writing." *AI & Society* 35, no. 4 (2020): 877-887. <https://doi.org/10.1007/s00146-020-00997-7>.

Diese Quellen bieten einen Überblick über verschiedene Aspekte der Künstlichen Intelligenz im Kontext des wissenschaftlichen Schreibens. Achten Sie darauf, die Links und Informationen entsprechend Ihrer tatsächlichen Quellen zu überprüfen und anzupassen.

# Recherche – Retrieval Augmented Generation

- Bei einer Anfrage wird zunächst ein Informationssystem wie eine Suchmaschine (z.B. bei Perplexity oder einem Bibliothekskatalog) oder einer Kundendatenbank nach passenden Informationen gesucht und die Anfrage mit Kontext aus dieser Datenbank angereichert. Selten wird ein kompletter Volltext oder eine Webseite dabei berücksichtigt, sondern nur die Metadaten, Keywords, Abstracts, Suchmaschinenoptimierungsangaben usw., da die Volltextanalyse derzeit noch zu teuer ist.
- Mithilfe der angereicherten Anfrage generiert das Large Language Model (LLM) einen Text und setzt im Text Fußnoten zu den entsprechenden Quellen
- Ziel: Korrektere Textgenerierung
- Vorsicht: Kann das Ergebnis verbessern, aber ein inhaltlicher Abgleich von Text und Quellen ist zwingend notwendig. Oft sind die Quellen nicht vertrauenswürdig.

Quelle / Link

# Retrieval Augmented Generation (RAG)

Wie verändert KI Bibliotheksarbeit?



Sprachmodell wandelt die natürlichsprachige Anfrage um in Suchstring

((KI OR kuenstliche intelligenz OR artificial intelligence) AND (bibliothek OR bibliotheksarbeit OR library) AND (transformation OR veraenderung OR innovation))

Anfrage mit dem Suchstring wird je nach Tool an eine Datenbank, das World Wide Web oder eine Dokumentensammlung gesendet



## Datenbank-RAG

Künstliche Intelligenz (KI) verändert die Bibliotheksarbeit grundlegend – sowohl in der Nutzung durch Besucher:innen als auch in internen Prozessen und der Rolle von Bibliotheken als Wissensvermittler. Hier die zentralen Entwicklungen:

### 1. Revolution der Recherche & Wissensvermittlung



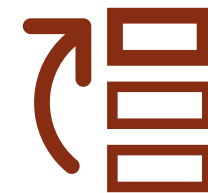
KI fungiert als **intelligente Schnittstelle** zwischen Nutzer:innen und Beständen:

- **Präzisere, kontextbewusste Suche:** Während klassische Recherche (z. B. über Google) mit stichwortbasierten Treffern arbeitet, analysiert KI die **Fragestellung und das Vorwissen der Nutzer:innen**, filtert relevante Quellen und liefert **strukturierte, verständliche Zusammenfassungen** – ähnlich einer „KI-Bibliothekarin“, die nicht nur Bücher findet, sondern sie auch interpretiert <sup>1</sup>.
- **Integration in Bibliothekssysteme:** Projekte wie *LibraryAI* der Staats- und Universitätsbibliothek Göttingen verknüpfen KI direkt mit **Katalogdaten, Volltexten und Fachressourcen**. Die KI greift auf aktuelle Bestände zu, vermeidet so veraltete oder fehlerhafte Inhalte generativer Modelle und wird zum „**verlässlichen Begleiter im Forschungsalltag**“ – speziell für wissenschaftliche Nutzer:innen <sup>2</sup>.
- **Schulungen zur KI-Nutzung:** Bibliotheken passen ihre **Medienkompetenz-Angebote** an, um Nutzer:innen – besonders Schüler:innen und Studierende – im **kritischen Umgang mit KI** zu schulen. Dazu gehören:
  - Quellenprüfung bei KI-generierten Inhalten.
  - Effiziente Nutzung von KI-Tools für Recherche und Präsentationen <sup>3, 4</sup>.
  - Bewertung von **Barrierefreiheit** (z. B. automatisierte Braille-Übersetzungen durch KI) <sup>5</sup>.

Abstracts, Metadaten (Schlagworte, Suchmaschinenoptimierungsangaben usw.) werden genutzt, um die ursprüngliche Anfrage semantisch anzureichern. Auf Basis dessen wird ein Text mit der Verlinkung zu Quellen im Web bzw. der entsprechenden Datenbank erstellt.



Davon werden noch mal je nach Tool 5, 8, 10 oder mehr Treffer als am passendsten eingestuft.



## Websuchen-RAG

KI verändert Bibliotheksarbeit vor allem durch Automatisierung routinierter Aufgaben und neue, stärker beratende Rollen für Bibliothekar\*innen. <sup>ao-itc +2</sup>

### Zentrale Veränderungen



- **Automatisierung:** Erschließung, Katalogisierung, Schrift- und Layouterkennung, Bestandsverwaltung und Ausleihprozesse werden zunehmend von KI-Systemen unterstützt oder teilweise übernommen. Das entlastet Fachangestellte deutlich, gerade bei immer gleichen, volumenstarken Tätigkeiten. <sup>zbw-mediatalk +2</sup>
- **Digitale Services:** Chatbots und virtuelle Assistenten beantworten Standardfragen, führen durch Kataloge, geben Öffnungszeiten an und schlagen passende Medien oder Veranstaltungen vor – auch außerhalb der Öffnungszeiten. <sup>zbw-mediatalk +2</sup>
- **Empfehlungs- und Suchsysteme:** KI-gestützte Recommendation-Engines und semantische Suche helfen Nutzer\*innen, relevante Literatur schneller zu finden und große digitale Bestände besser zu nutzen. <sup>bibliotheksportal +2</sup>

Die am besten zur Suchanfrage passenden Ergebnisse werden sortiert. Je nach Tool können das 30, 50, 100 oder mehr Ergebnisse sein. Das ist oft nicht transparent.



# Potenzial und Grenzen

## **Potenzial, z.B.**

Reduzierung von Halluzination

Erschließung von Resultaten, die ggf. mit klassischer Suchanfragenformulierung so nicht zu finden gewesen wären

Schneller Überblick über die Hauptaspekte der Resultate und über ein Thema

## **Grenzen**

Recherche ist nicht reproduzierbar

Quellen sind ggf. nicht zitierwürdig oder irrelevant

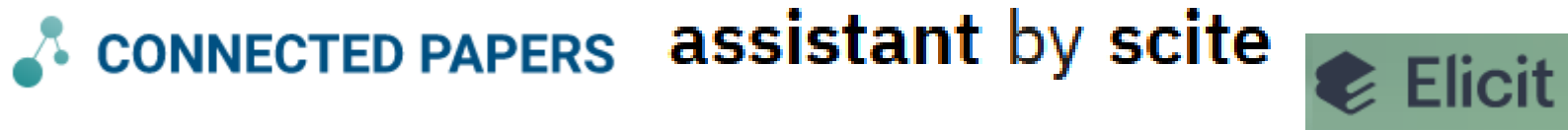
Generierter Text hat ggf. keinen Bezug zu den Quellen

Teilweise werde Textpassagen aus den Quellen übernommen → Plagiatsgefahr

Oft werden nur Metadaten und Abstracts als Basis für die Textgenerierung verwendet

# Recherche – AI Research Assistants

KI-gestützte Anwendungen zur Unterstützung von Literatursuche, -auswahl und Auswertung.  
Funktionsumfang sehr unterschiedlich.



Siehe auch:

- Toolbeschreibung der Uni Tübingen: <https://uni-tuebingen.de/einrichtungen/universitaetsbibliothek/lernen-arbeiten/schulungen-beratungen-fuehrungen/literaturrecherche-mit-ki/#c2070900>
- Auswahl von Rechercheassistenten der Hochschule Rhein Main: [https://www.hs-rm.de/fileadmin/user\\_upload/KI-Tools\\_Literaturrecherche\\_aktuelle\\_Version\\_Mai\\_24.pdf](https://www.hs-rm.de/fileadmin/user_upload/KI-Tools_Literaturrecherche_aktuelle_Version_Mai_24.pdf)
- Toolssammlung nach Funktionen des VK:KIWA: <https://www.vkkiwa.de/ki-ressourcen/>

# AI Research Assistants



Kullmann, Hiebl (2023): [https://www.linkedin.com/posts/sylvia-kullmann-499904267\\_ki-transformation-activity-7108740735847276545-w3PY?utm\\_source=share&utm\\_medium=member\\_desktop](https://www.linkedin.com/posts/sylvia-kullmann-499904267_ki-transformation-activity-7108740735847276545-w3PY?utm_source=share&utm_medium=member_desktop)

# Recherche – AI Research Assistants

## Mögliche Funktionen

- Natürlichsprachige Suchanfragen
- Newsfeeds/Alerts auf Basis vorheriger Literaturlauswahl
- Literatursuche auf Basis einer Bibliothek im Literaturverwaltungsprogramm
- Visuelle Darstellung von Zitations- oder Inhaltsbeziehungen
- Zusammenfassung mehrerer Abstracts
- Chatten mit dem Text
- Outline z.B. der Methoden und Ergebnisse
- Datenextraktion

# Semantic Scholar – <https://www.semanticscholar.org/>

- Indexierung von offenen und lizenzierten Journals
- Auswertung von Referenzen und Zitationen (wie bei klassischen Datenbanken auch)
- Vorschläge aufgrund der Abstracts ausgewählter Artikel
- Erstellung von Alerts auf Basis von Kollektionen
- Zusammenfassung von Abstracts → TLDR /too long, didn't read
- Semantic Reader
- Datenquellen tlw. einsehbar unter <https://www.semanticscholar.org/about/publishers>
- Kostenlos, Anmeldung nur zur Speicherung von Favoriten und Research Feed notwendig

## About 1,950 results for "ai academic writng education"

Fields of Study ▾

Date Range ▾

Has PDF

Author ▾

Journals &amp; Conferences ▾

### The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT

D. Dalalah

Osama M.A. Dalalah

Education, Computer Science ·

[The International Journal of Management Education](#) · 1 July 2023

36

Publisher

Save

Cite



Related Papers



Chat with paper

### Detection of GPT-4 Generated Text in Higher Education: Combining Academic Judgement and Software to Identify Generative AI Tool Misuse

Mike Perkins

Jasper Roe

+5 authors

Singapore

Computer Science, Education ·

[Journal of Academic Ethics](#) · 29 May 2023

**TLDR** Recommendations include adjusting the overall strategies for assessing university students in light of the availability of new Generative AI tools, and reducing the overall reliance on assessments where AI tools may be used to mimic human writing, or by using AI-inclusive assessments. [Expand](#)

22

[PDF]

Cites a paper in your library

Springer

Save

Cite



Related Papers



Chat with paper

Semantic  
Reader

Armin Glatzmeier & Johanna Gröpler (FU Berlin)

Semantic Reader

# UX Research on Conversational Human-AI Interaction: A Literature Review of the ACM Digital Library

Qingxiao Zheng, Yiliu Tang, +2 authors Yun Huang • Published in International Conference on

Computer Science

**TLDR** These findings show that designing with social boundaries, such as privacy, disclosure, and identification, is crucial for ethical polyadic CAs and should also advance usability testing methods and trust-building guidelines for conversational AI. [Expand](#)

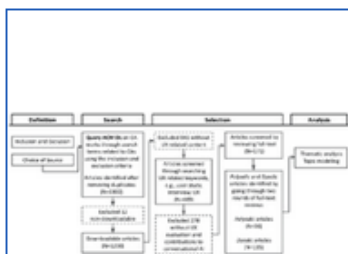
[PDF] Semantic Reader Save to Library Create Alert Cite

Figures and Tables

Topics

Ask This

## Figures and Tables from this paper



Armin Glatzmeier & Johanna Gröpler (FU Berlin)

26 Citations

Highly Influential Citations

Background Citations

Methods Citations

7

View All



View all papers from this venue

## Ask This Paper

BETA

AI-Powered

What is the goal of this paper?

What are the key results of this paper?

What methods are used?

Ask a question that can be answered by this paper.

Submit

0/800 Characters

Our system tries to constrain to information found in this paper. Results quality may vary.

[Learn more](#) about how we generate these answers.

Zheng, Q., Tang, Y., Liu, Y., Liu, W., & Huang, Y. (2022). UX Research on Conversational Human-AI Interaction: A Literature Review of the ACM Digital Library. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. <https://www.semanticscholar.org/reader/db2c80017284d8492026868aeba957efd8aa7cf3>

are designed to mediate human-human interactions. CAs for polyadic interactions are unique because they encompass hybrid social interactions, i.e., human-CA, human-to-human, and human-to-group behaviors. However, research on polyadic CAs is scattered across different fields, making it challenging to identify, compare, and accumulate existing knowledge. To promote the future design of CA systems, we conducted a literature review of ACM publications and identified a set of works that conducted UX (user experience) research. We qualitatively synthesized the effects of polyadic CAs into four aspects of human-human interactions, i.e., communication, engagement, connection, and relationship maintenance. Through a mixed-method analysis of the selected polyadic and dyadic CA studies, we developed a suite of evaluation measurements on the effects. Our findings show that designing with social boundaries, such as privacy, disclosure, and identification, is crucial for ethical polyadic CAs. Future research should also advance usability testing methods and trust-building guidelines for conversational AI.

**Keywords** Conversational Agent · Chatbot · Conversational AI · UX Research · Literature Review

### 1 Introduction

There is a rapidly growing body of literature on conversational agents or chatbots [Adamopoulou and Moussiades, 2020a]. As promising Artificial intelligence (AI) technologies, conversational agents are defined as "software that accepts natural language as input and generates natural language as output, engaging in a conversation with the user" [Griol et al., 2013]; chatbots, meanwhile, are computer programs designed to simulate conversation with human users via text [Adamopoulou and Moussiades, 2020a,b]. As these two terms are often perceived as interchangeable [Rapp

Page Loading

Method

et al., 2021, McTear, 2020], in the remainder of this paper, we refer to both conversational agents and chatbots as CAs. Scholars have shown that these machines are able to compensate for human shortcomings or exceed human capacities [Fox and Gambino, 2021, Guzman and Lewis, 2020, Whittaker et al., 2018]. However, prior works focus on designing and evaluating dyadic human-AI interaction, which involve only one-to-one interactions between humans and their CAs [Bickmore et al., 2005, Schulman and Bickmore, 2009, Xu et al., 2017, Kopp et al., 2005, Anabuki et al., 2000]; whereas more recent works start tapping into polyadic human-AI interactions that also support human-human interactions [Kim et al., 2021, Wang et al., 2021, Kim et al., 2020, Toxtli et al., 2018, Benke et al., 2020].

Method

Even though there are extensive literature reviews on CAs, e.g., [Seering et al., 2019, Chaves and Gerosa, 2021, de Barcelos Silva et al., 2020, Montenegro et al., 2019, Laranjo et al., 2018], they do not address how polyadic CAs are designed and evaluated, nor present the effects of using polyadic CAs on handling the challenges of human-human interaction [Hohenstein and Jung, 2020]. In this paper, we overview UX (user experience) research on polyadic CAs that 1) interact with more than one user in the same conversation and 2) engage in bidirectional conversations between all parties (human-AI and human-human). These polyadic CAs encompass a wide variety of complexities that dyadic Human-AI may not encounter, including multi-party interactions, social roles taking, group hierarchy, or social tension [Van Dijk, 1997]. To evaluate the effects of CAs' support in human-human interactions, we need to examine real human-to-human behaviors and human-to-human behaviors, as well as human-to-group behaviors potentially. Given the unique challenges of the design space, little is known about how polyadic CAs should be designed to address these

Method

## Skimming Highlights ✕

AI-generated highlighting to support skimming Back

**Enable Skimming**  
Show automatically-generated highlights

**Goals** (23 highlights)  
Paper objectives are highlighted in blue.  
Example: **In this work, we propose...**

**Methods** (56 highlights)  
Methods used are highlighted in orange.  
Example: **The methods used...**

**Results** (17 highlights)  
Paper results are highlighted in pink.  
Example: **Our results show...**

**Show Margin Flags**  
Label highlight types in the margin.

Number of Highlights

Feedback? [Contact Us](#)



# Consensus

<https://consensus.app/>

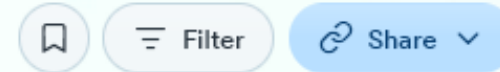
- Natürlichsprachige Anfragen
- Zusammenfassungen der Abstracts auf verschiedenen Sprachen möglich
- Bei Ja-/Nein-Fragen: „Stimmungsbarometer“
- Datenbasis: Semantic Scholar
- Freeversion mit 20 AI Credits pro Monat

The screenshot displays the Consensus app interface. At the top, a search bar contains the query "wie verändert generative KI die Recherche?". Below the search bar, there are toggle switches for "Synthesize" and "Copilot", along with "Filter" and "Share" buttons. The main content area shows three search results, each with a summary, citation information, and a "Study Snapshot" button. The first result is titled "Ethics of generative AI" and includes a summary about generative AI's impact on research integrity. The second result is "Leveraging the Potential of Generative AI to Accelerate Systematic Literature Reviews: An Example in the Area of Educational Technology" with a summary about AI's role in systematic reviews. The third result is "Assessing the impact of generative AI on medicinal chemistry" with a summary about AI's potential in drug discovery. Each result also features a "Study Snapshot" button and a "Share" button. A dropdown menu is visible on the right side of the first result, showing options to download in ".RIS" and ".CSV" formats.

verändert generative KI bisherige Recherchepraktiken?



Synthesize  Copilot



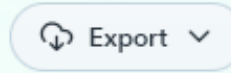
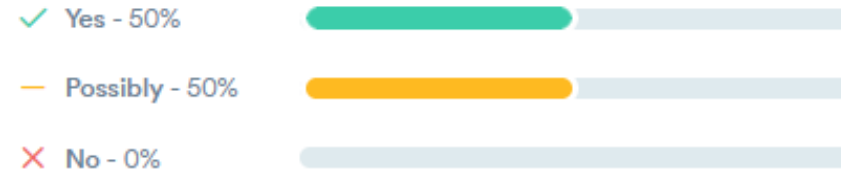
### Summary Beta

9 papers analyzed

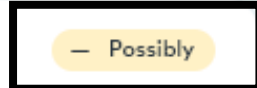
Diese Studien legen nahe, dass generative KI bisherige Recherchepraktiken verändert, indem sie die Komplexität versteht, Daten zusammenfasst, relevante Artikel identifiziert, Forschung automatisiert, wissenschaftliche Entdeckungen beschleunigt und die Produktivität steigert, während einige Studien auch auf potenzielle und tatsächliche Schäden hinweisen.

### Consensus Meter Beta

8 papers analyzed



### Generative artificial intelligence community of practice for research



Generative KI kann Forschern helfen, Komplexität zu verstehen, genaue Zusammenfassungen von Daten und Ergebnissen bereitzustellen und möglicherweise aktuelle Forschungsprozesse zu automatisieren.

International Wound Journal | Steven H. Cohen et al. | 0 citations | 2023

Study Snapshot ▼



## Einführung

Generative Künstliche Intelligenz (KI) verändert die Art und Weise, wie Forschung betrieben wird, erheblich. Diese Technologien, insbesondere große Sprachmodelle (LLMs), bieten neue Möglichkeiten zur Automatisierung und Verbesserung von Forschungsprozessen in verschiedenen Disziplinen.

## Wichtige Erkenntnisse

- **Automatisierung und Beschleunigung von Forschungsprozessen:**
  - Generative KI kann die Identifikation relevanter Artikel in systematischen Literaturübersichten (SLRs) erleichtern und so den Forschungsprozess beschleunigen **2** **4**.
  - KI-Modelle können Mikrotätigkeiten wie Ideengenerierung, Feedback, Schreiben, Hintergrundrecherche, Datenanalyse und Codierung automatisieren, was zu erheblichen Produktivitätsgewinnen führt **4**.
- **Personalisierung und Verbesserung der Forschungsergebnisse:**
  - Generative KI kann personalisierte Forschungsgemeinschaften schaffen, die auf die Bedürfnisse und Profile der Mitglieder zugeschnitten sind, was die Suche nach relevanten Inhalten erleichtert und die Benutzerbindung erhöht **1**.
  - In phygitalen Umgebungen kann generative KI qualitative Forschung validieren und Einblicke über menschliche Intuition hinaus generieren **5**.
- **Erweiterung der Forschungskapazitäten:**
  - Generative KI kann virtuelle Modelle des menschlichen Körpers erstellen, um Verletzungen zu simulieren und die Wirksamkeit von Behandlungen zu testen, sowie neue Medikamente und Behandlungsprotokolle entwickeln **6**.
  - In der Umweltgesundheitsforschung kann KI Forschungsergebnisse für nicht-akademische Leser zusammenfassen und so die Forschungstranslation verbessern **8**.
- **Herausforderungen und Bedenken:**
  - Es gibt Bedenken hinsichtlich der potenziellen und tatsächlichen Schäden durch die Nutzung von generativer KI, einschließlich der Notwendigkeit, die organisatorischen, technologischen und interaktionellen Praktiken zu untersuchen, durch die KI und ihre Ergebnisse produziert werden **7**.
  - Die schnelle Entwicklung und die Einschränkungen der Technologie, wie die Anzahl der Anfragen pro Minute und geografische Verfügbarkeit, stellen Herausforderungen dar **2**.

## Zusammenfassung

Generative KI revolutioniert die Forschungspraxis, indem sie Prozesse automatisiert, personalisierte und validierte Forschungsergebnisse liefert und die Forschungskapazitäten erweitert. Obwohl es viele Vorteile gibt, es auch Herausforderungen und Bedenken, die berücksichtigt werden müssen, um die verantwortungsvolle Nutzung dieser Technologien sicherzustellen.

Armin Glatzmeier & Johanna Gröpler (FU Berlin)

**Methods**

**Study types** ⓘ

- Meta Analysis
- Systematic Review
- Case Report**
- Animal Trial
- In Vitro Trial

**Case Report**

A case report is a paper that describes an individual case or subject ( $n = 1$ ). While case reports can be useful anecdotes to justify further research, findings should not be generalized.

[Learn more](#) ↗

# Recherche – AI Research Assistants

## Begrenzungen

- Nutzung erfordert oft Rechercheerfahrung
- Datenbasis oft nicht transparent
- Hauptsächlich Open-Access-Journals (bei den freien Versionen)
- Fächerschwerpunkt lebens-, technik- und naturwissenschaftlich
- Datenbasis häufig Englisch
- Anfragen auf Englisch bringen tlw. bessere Ergebnisse
- Zusammenfassungen meistens basierend auf den Abstracts, nicht auf Volltext
- Funktionsumfang abhängig von Kostenmodell

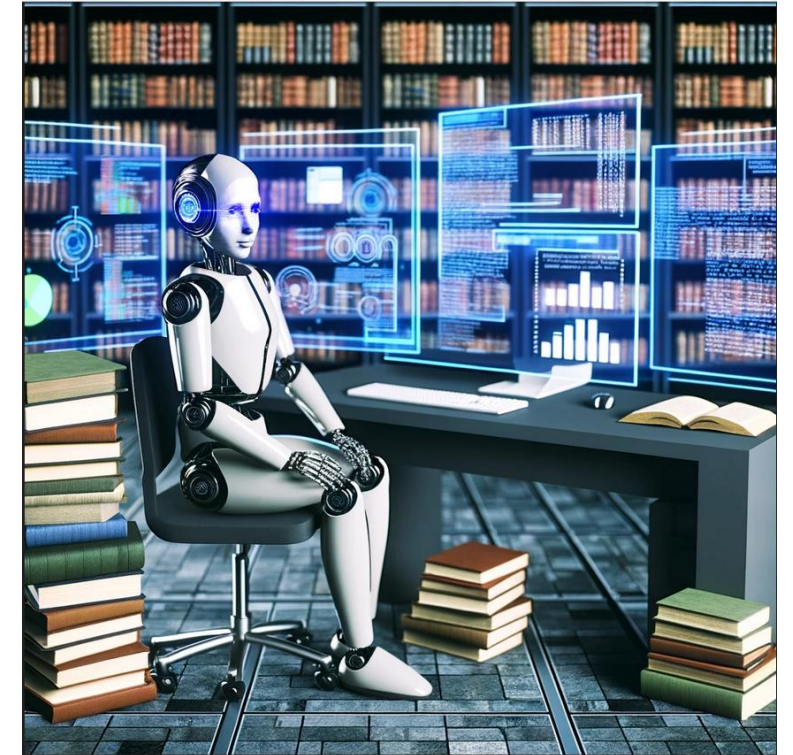
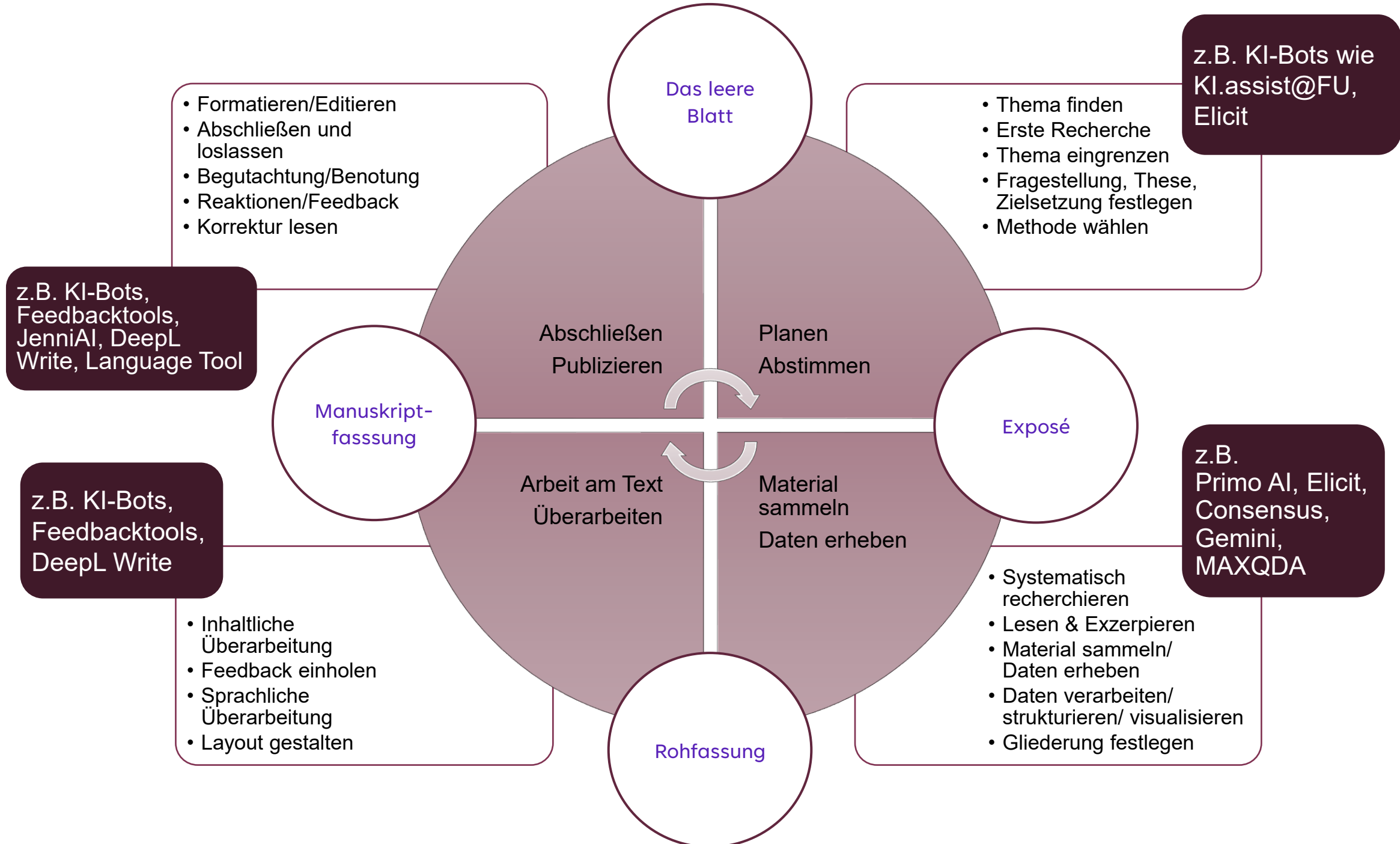


Bild erstellt mit DALL-E

Anhang

# Mögliche Einsatzgebiete von KI-Tools im Schreibprozess



Schreibprozess leicht abgeändert nach Otto Kruse (2007). Keine Angst vor dem leeren Blatt: Ohne Schreibblockaden durchs Studium (12., völlig neu bearbeitete Aufl.). Campus-Verl. S. 112

# Welches Tool für welchen Arbeitsschritt?

(da die Bedürfnisse so unterschiedlich sind, nur eine kurze Liste)

- DeepL Write zur Textbearbeitung wie Synonyme, Grammatik, Satzzeichen, Formulierungen, Ton: <https://www.deepl.com/de/write>
- Überprüfung der Argumentation usw.: Sprachmodelle möglich
- Recherche:
  - Suchbegriffe generieren: Große Sprachmodelle möglich, zusätzlich aber vor allem auch nicht KI-basierte Nachschlagewerke – Synonymwörterbücher wie OpenThesaurus, Duden-Werke, Duden Online, Oxford Learners Dictionaries
  - Zur Orientierung und Ergänzung der Suche in klassischen Datenbanken AI Research Assistants (Tools unter Literaturrecherche): <https://www.vkkiwa.de/ki-ressourcen/ki-tools/> (ACHTUNG: Reichen aufgrund der Datengrundlage nicht für eine umfängliche Recherche. Niemals Quellen in ChatGPT recherchieren bzw immer sehr gründlich nachrecherchieren und durch klassische Datenbankrecherche komplettieren)
- Academic Cloud bietet datenschutzkonformen Zugang zu Sprachmodellen. Anleitung am Ende der Präsentation. Anmeldung mit Ihrem Hochschulaccount.

**Sprechen Sie IMMER ab, ob Sie KI-Tools für einen bestimmten Zweck nutzen dürfen.**

# Worst Practices



Was ‚wissen‘ Große Sprachmodelle

# Trainingsdaten – Was wissen wir darüber?

## 2 Scope and Limitations of this Technical Report

This report focuses on the capabilities, limitations, and safety properties of GPT-4. GPT-4 is a Transformer-style model [39] pre-trained to predict the next token in a document, using both publicly available data (such as internet data) and data licensed from third-party providers. The model was then fine-tuned using Reinforcement Learning from Human Feedback (RLHF) [40]. Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar.

We are committed to independent auditing of our technologies, and shared some initial steps and ideas in this area in the system card accompanying this release.<sup>2</sup> We plan to make further technical details available to additional third parties who can advise us on how to weigh the competitive and safety considerations above against the scientific value of further transparency.

\* | OpenAI et al. 2023. GPT-4 Technical Report. arXiv:2303.08774, <https://doi.org/10.48550/arXiv.2303.08774>

# Trainingsdaten – Was „weiß“ GPT-3?

- Das Modell GPT-3 wurde mit folgenden Sammlungen trainiert\*

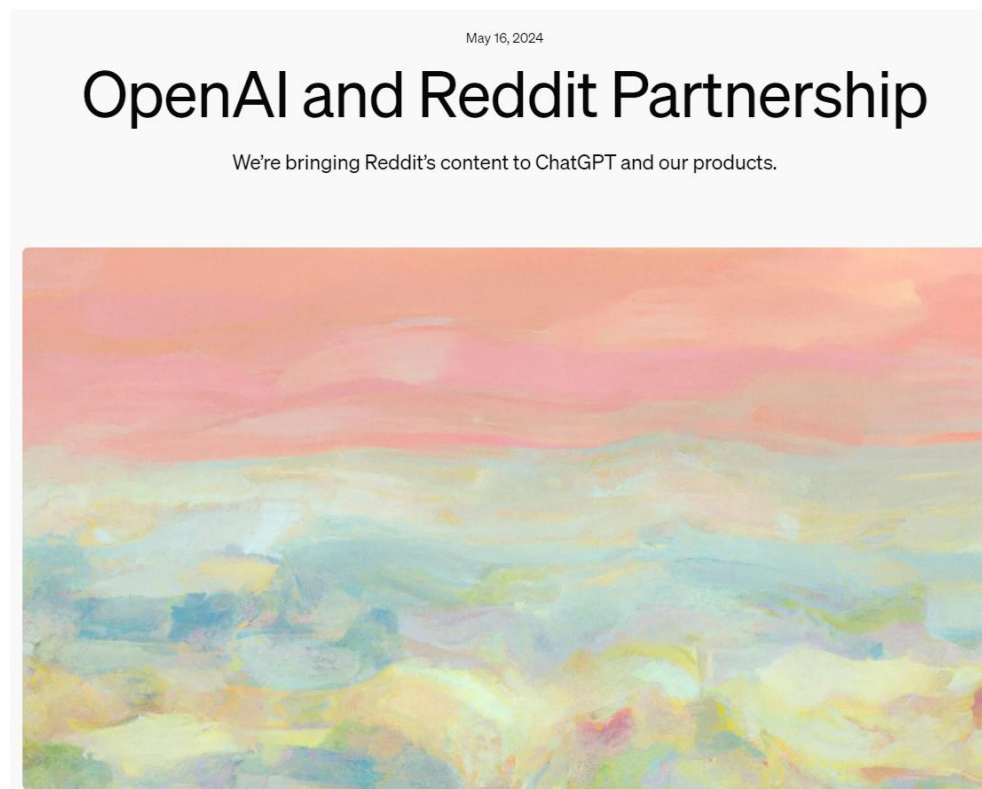
Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

- Diese Datensätze enthalten<sup>+</sup>
  - Webseiten
  - Bücher und Artikel
  - Inhalte aus Sozialen Medien, Blogs, Foren, Wikipedia usw.

\* | Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah et al. 2020. "Language Models are Few-Shot Learners". *Arxiv* 2005.14165: 9; <https://doi.org/10.48550/arXiv.2005.14165>

+ | Rudolph, Jürgen, Samson Tan, and Shannon Tan. 2023. "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of Applied Learning & Teaching* 6(1): 3; <https://doi.org/10.37074/jalt.2023.6.1.9>

# Trainingsdaten – Was „weiß“ GPT?



# Trainingsdaten – Was „weiß“ GPT?

[Home](#) > [SDNY Blog](#) > Copyright Infringement Lawsuits Against OpenAI And Microsoft Are Mounting

Newcomer, Maghan. 2024. Copyright Infringement Lawsuits Against OpenAI and Microsoft Are Mounting, <https://www.sdneyblog.com/copyright-infringement-lawsuits-against-openai-and-microsoft-are-mounting/>

## Copyright Infringement Lawsuits Against OpenAI and Microsoft Are Mounting

By [Meghan Newcomer](#) on March 5, 2024

In [two complaints](#) filed last week, The Intercept Media and AlterNet Media, Inc. became the latest companies to sue for copyright infringement in violation of the Digital Millennium Copyright Act. The complaints included Microsoft as a defendant.

Both complaints were filed by self-identified “news organizations” whose copyrighted works were used to train AI systems, ChatGPT, on how to mimic human speech. The news organizations, when deciding what information to feed to ChatGPT:

“Defendants had a choice: they could train ChatGPT using works of journalism with the copyright management information protected by the [Digital Millennium Copyright Act] intact, or they could strip it away. Defendants chose the latter, and in the process, trained ChatGPT not to acknowledge or respect copyright, not to notify ChatGPT users when the responses they received were protected by journalists’ copyrights, and not to provide attribution when using the works of human journalists.”

# Trainingsdaten vs. Urheberrecht

[Home](#) > [SDNY Blog](#) > Copyright Infringement Lawsuits Against OpenAI And Microsoft Are Mounting

## Copyright Infringement Lawsuits Against OpenAI and Microsoft Are Mounting

By [Meghan](#)

In [two](#) [con](#)  
and AlterN  
infringeme  
included N

Both comp

those organizations' copyrighted works were used to train systems, ChatGPT, on how to mimic human speech. News organizations, when deciding what information materials fed to ChatGPT:

First, the News Plaintiffs continue to bear significant burden and expense in searching for their copyrighted works in OpenAI's training datasets within a tightly controlled environment that this Court and the parties have previously referred to as "the sandbox." OpenAI has provided the News Plaintiffs with two dedicated virtual machines with improved computing resources for performing their searches, and News Plaintiffs have spent an additional 150 person-hours (and even more computing hours) since November 1 searching OpenAI's training data. On November 14, all of News Plaintiffs' programs and search result data stored on one of the dedicated virtual machines was erased by OpenAI engineers. Maisel Decl. ¶ 3; Ex. A at 5. While OpenAI was able to recover much of the data that it erased, the folder structure and file names of the News Plaintiffs'

when the responses they received were protected by journalists' copyrights, and not to provide attribution when using the works of human journalists.

Newcomer, Maghan. 2024. Copyright Infringement Lawsuits Against OpenAI and Microsoft Are Mounting, <https://www.sdneyblog.com/copyright-infringement-lawsuits-against-openai-and-microsoft-are-mounting/>  
Das Schreiben der Anwälte hier: <https://storage.courtlistener.com/recap/gov.uscourts.nysd.612697/gov.uscourts.nysd.612697.328.0.pdf>

# Trainingsdaten – Was „wissen“ LLMs?

- Vortrainierte LLMs haben idR keine Internetanbindung (aber: Retrieval augmented generation)
- Die Trainingsdaten sind idR bereinigt, um problematische Inhalte wie Gewalt, Vorurteile, Hate Speech etc. auszuschließen\*
  - Die Trainingsdaten enthalten ein umfangreiches Spektrum unterschiedlicher menschlicher Sprache
  - Die Trainingsdaten allgemeiner LLMs haben *keinen spezifischen wissenschaftlichen Zuschnitt*
  - Die Trainingsdaten können *Fehler, Verzerrungen, Biases und Mißrepräsentationen* enthalten (und tun dies auch)
  - Die Auswahl der *Trainingsdaten* und die Kriterien ihrer Bereinigung liegen *in der ausschließlichen Hoheit der jeweiligen Anbieter*



## Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

\* | Perrigo, Billy. 2023. "The \$2 Per Hour Workers Who Made ChatGPT Safer". *Time*, 18.01.2023; <https://time.com/6247678/openai-chatgpt-kenya-workers/>

# Lizenzrecht und Nachhaltigkeit



# LIZENZRECHTLICHE BEGRENZUNGEN – Z.B. WILEY

- 5.5. Wiley, the DEAL Operating Entity, and the Member Institution all acknowledge the challenges and complexities created by rapidly emerging technologies such as AI technologies, and commit to mutual support in developing policies and protocols for all parties' benefit. Member Institution and its Authorized Users may not, without Wiley's express written permission, (a) use or copy the Licensed Electronic Products or any substantial part thereof as input for artificial Intelligence (AI) tools or technologies; (b) use or enable AI technologies (machine learning models) to ingest, train, test, analyze, process, copy, distribute, make publicly accessible, and/or generate Output based on the Licensed Electronic Products or any part thereof; or (c) upload or copy the Licensed Electronic Products or any part thereof to or for use in AI technologies directly or indirectly (e.g., through the use of third-party plugins associated with those AI technologies). The foregoing restrictions do not apply to Licensed Electronic Products made available under a CClicense, for which the terms of the respective CClicense apply (including but not limited to the requirement of attribution). AI Tools and Technologies means Software that is developed with one or more

# Nachhaltigkeit

## ChatGPT-3 vs. ChatGPT-4:

### Training Power Consumption and Costs in U.S. Dollars

Model	 <b>GPT-3</b>	 <b>GPT-4</b>
Days of training	34	100
Training energy consumption	1,287,000 kWh	62,318,800 kWh
Cost of training*	\$168,597	\$8,163,763

Data sources: U.S. Energy Information Administration, Electric Power Research Institute (EPRI)

\*Calculated based on the average U.S. commercial electricity rate of \$0.131 per kWh as of June 2024



<https://www.bestbrokers.com/forex-brokers/ais-power-demand-calculating-chatgpts-electricity-consumption-for-handling-over-78-billion-user-queries-every-year/>

# Nachhaltigkeit



<https://www.appypie.com/blog/hardware-requirements-for-llm-training>

# Nachhaltigkeit

## ChatGPT's Energy Consumption for Responding to Prompts and Its Cost in the U.S.

Time Period	Number of prompts	Energy Consumption	Cost
Day	214,285,714	621,429 kWh	\$81,407
Week	1,500,000,000	4,350,000 kWh	\$569,850
Month	6,517,860,000	18,901,794 kWh	\$2,476,135
Year	78,214,350,000	226,821,615 kWh	\$29,713,632

“Apparently, each time you ask ChatGPT a question, it uses about 0.0029 kilowatt-hours of electricity. This is nearly ten times more than the energy needed for a typical Google search, which consumes about 0.0003 kilowatt-hours per query, according to The Electric Power Research Institute (EPRI).”

Data sources: U.S. Energy Information Administration, Electric Power Research Institute (EPRI)  
Calculations based on: 100 million weekly users, 15 weekly prompts per user, 0.0029 kWh of energy consumption per prompt, average U.S. commercial electricity rate of \$0.131/kWh as of June 2024

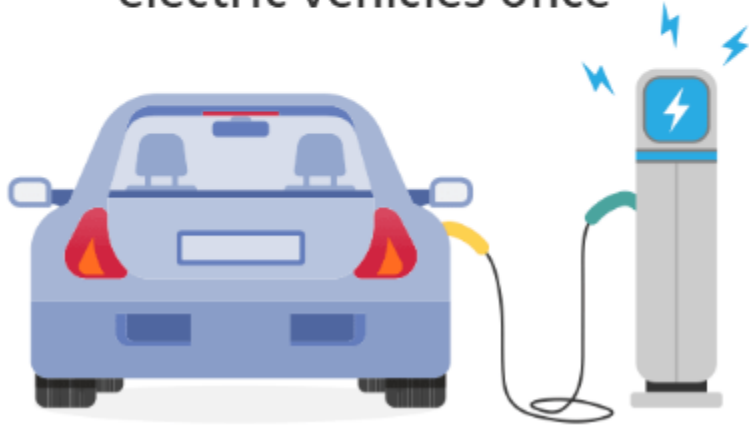


<https://www.bestbrokers.com/forex-brokers/ais-power-demand-calculating-chatgpts-electricity-consumption-for-handling-over-78-billion-user-queries-every-year/>

# Nachhaltigkeit

**The annual amount of electricity ChatGPT uses  
to respond to prompts (226.82 million kWh)  
would be enough to:**

**Fully charge 3.13M  
electric vehicles once**



**Power 21,602  
U.S. households for a year**



<https://www.bestbrokers.com/forex-brokers/ais-power-demand-calculating-chatgpts-electricity-consumption-for-handling-over-78-billion-user-queries-every-year/>

Tools, Toolsammlungen und Kurse

# KI-Textgenerierungstools (Auswahl, tlw. Anmelde-/kostenpflichtig)

- ChatGPT von Open AI (basiert je nach Kostenmodell auf GPT-3.5, GPT-4 oder GPT-4o (aktuellstes), <https://chatgpt.com/>)
- BingChat/Copilot von Microsoft (basiert auf den GPTs von Open AI), Microsoft-Account notwendig, Bildgenerierung möglich, Live-Suche im Internet <https://copilot.microsoft.com/>)
- Claude (Sprachmodell von Anthropic, Canada), Anmeldung notwendig, <https://claude.ai/>)
- Gemini (Sprachmodell von Google), Anmeldung notwendig, <https://gemini.google.com/>)
- Mistral LeChat (Sprachmodell vom französischen Unternehmen Mixtral), <https://chat.mistral.ai/chat>)
- Perplexity (nach Anmeldung Zugriff auf verschiedenen Sprachmodelle, bei Fokus „all“ Live-Suche im Internet, <https://www.perplexity.ai/>)
- You.com (nach Anmeldung Zugriff auf verschiedenen Sprachmodelle), <https://you.com/>)

# Academic cloud der GWDG I

Chat mit verschiedenen Sprachmodellen möglich, Zugriff über „föderierte Anmeldung“ und dann FU-Account ODER kostenfrei Registrieren

The screenshot shows the Chat-AI interface. At the top left is the 'Chat-AI' logo. At the top right are logos for 'KISSKI' (KI-Servicezentrum für sensible und kritische Infrastrukturen) and 'GWDG' (Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen). A warning message in a light blue box states: 'Hinweis: Der Chatbot ist anfällig für halluzination und ihre Antworten sollten nicht als korrekt angesehen werden. Datenschutz und Impressum.' Below this is a large empty text area for the chat. On the right side, there is a text input field with the placeholder 'Fragen Sie mich' and icons for adding attachments and voice recording. Below the input field is a list of AI models: 'Meta LLaMA 3 8B Instruct', 'Mixtral 8x7B Instruct', 'Meta LLaMA 3 70B Instruct', 'LLaMA 3 SauerkrautLM 70B Instruct', and 'Qwen 2 72B Instruct'. Below the list is a dropdown menu labeled 'Modell' with a question mark icon, currently showing 'Meta LLaMA 3 8B Instruct'. At the bottom right, there is a link for 'Erweiterte Optionen' with a question mark icon. A small blue icon with an upward arrow is located at the bottom center of the chat area.

<https://chat-ai.academiccloud.de/chat>

# Academic cloud der GWDG II

Zugriff über „föderierte Anmeldung“ und dann FU-Account ODER kostenfrei Registrieren






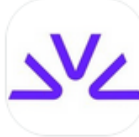


- ChatAI mit verschiedenen Open Source-Modellen: <https://chat-ai.academiccloud.de/chat>  
Dokumentation zu den eingebundenen Sprachmodellen: <https://docs.hpc.gwdg.de/services/chat-ai/models/index.html>
- ImageAI (Text zu Bild): <https://image-ai.academiccloud.de/>
- VoiceAI (Sprache zu Text): <https://voice-ai.academiccloud.de/>
- Retrieval Augmented Generation mithilfe der Arcanas (d.h. Dokumentensammlung anlegen, um diese dann in ChatAI befragen zu können):  
Arcanas-Manager: <https://chat-ai.academiccloud.de/arcanas/arcana>  
Anleitung: <https://docs.hpc.gwdg.de/services/arcana/getting-started/index.html>

# Poe.com

Plattform, um Sprach- und Bildmodelle zu testen, mit Anmeldung, ohne Bezahlung






## Offizielle Bots

[Alles anzeigen](#)

 <p><b>Assistant</b> Allzweck-Assistenten-Bot.</p>	 <p><b>Llama-3.1-405B-T</b> Llama 3.1 405B Instruct from Meta. Supports 4k tokens ...</p>	 <p><b>Claude-3.5-Son...</b> Anthropic's Claude 3.5 Sonnet eignet sich besonders für...</p>	 <p><b>GPT-4o</b> Das leistungsstärkste Modell von OpenAI. Übertrifft GPT-3.5 be</p>
 <p><b>Web-Search</b> Allzweck-Assistenten- Bot, der bei Bedarf eine Websuche...</p>	 <p><b>Llama-3.1-405B-...</b> The Meta Llama 3.1 collection of multilingual large...</p>	 <p><b>GPT-4o-Mini</b> Das neueste Modell von OpenAI. Dieses intelligente kleine...</p>	 <p><b>Llama-3.1-70B-F...</b> The Meta Llama 3.1 collection of multilingual large...</p>

## Deutsch-Bots

[Alles anzeigen](#)

 <p><b>Motivationschr...</b> Ich helfe dir dabei, ein gutes und... 438 monatliche Nutzer</p>	 <p><b>Bicki-GPT</b> Ich schreibe und erkläre für Sie Code... 91 monatliche Nutzer</p>	 <p><b>Buch-Generator</b> Generiere spannende Geschichten mit etw... 159 monatliche Nutzer</p>	 <p><b>AnkiMeisterBeta</b> Feeling like the Russian bear trappe... 2 monatliche Nutzer</p>
 <p><b>AnalyseErstellung</b> Anweisungen: Analysiere den... 87 monatliche Nutzer</p>	 <p><b>WINGERv0_4e5</b> It's me, Gennaio! V0_1 1 monatlicher Nutzer</p>	 <p><b>Herr_Fruehblue...</b> Herr Frühblüher beantwortet nur... 19 monatliche Nutzer</p>	 <p><b>QUANTEN-ANA...</b> Q-BOT ist der fortschrittlichste... 3 monatliche Nutzer</p>

## Beliebte Bots

[Alles anzeigen](#)

Armin Glatzmeier & Johanna Gröpler (FU Berlin)

 <p><b>Image_Creators</b> Describe the image</p>	 <p><b>Photo_CreateE</b> This bot generates</p>	 <p><b>RealVisXL</b> Generates photos</p>	 <p><b>LOGO-CREATOR...</b> ★★★★★</p>
---	--	--	---

# Lmstudio – Sprachmodelle lokal auf eurem Rechner

The image shows the LM Studio website on the left and a screenshot of the LM Studio application interface on the right. The website features the LM Studio logo, a navigation menu, and a main heading "Discover, download, and run local LLMs". Below this, there are buttons for various models: Llama 3, Phi 3, Falcon, Mistral, StarCoder, and Gemma, along with a "gguf" icon and a note "models from Hugging Face". A "Technology Preview: LM Studio 0.2.28 with AMD ROCm" link is also present. Three download buttons are shown for M1/M2/M3, Windows (x86), and Linux (Beta), all version 0.2.28. At the bottom, it says "LM Studio is provided under the terms of use." The application interface on the right shows a chat window with a model selected as "TheBloke · codellama instruct 7B q3\_k\_l gguf". The interface includes a chat area with a "USER" prompt "Please write a snake", a settings panel on the right, and a top status bar showing "RAM Usage: 3.71 GB" and "CPU: 0 %". A purple banner at the top of the application window reads: "New in v0.2.28: Meta's Llama 3.1 is here! It's the most capable local LLM ever. Read more about it here."

<https://lmstudio.ai/>

Armin Glatzmeier & Johanna Gröpler (FU Berlin)

# Recherchertools

Auswahl des VK:KIWA mit Fokus auf wissenschaftliches Schreiben:

<https://www.vkkiwa.de/ki-ressourcen/>

Übersicht zu Recherche mit KI der Hochschule RheinMain: <https://www.hs-rm.de/de/service/hochschul-und-landesbibliothek/suchen-finden/ki-tools#recherche-plattformen-mit-ki-unterstuetzung-132409>

Toolsteckbriefe der ZHAW:

<https://www.taskcards.de/#/board/19a19b6d-f2a5-42eb-bb92-da42cbe2863e/view?token=a1e19e20-e8bd-4b34-a53a-367a354b3a6f>

# Links und Literatur

# Ressourcen

## **Artikel zu KI-basierten Forschungsassistenten:**

Kullmann, S., & Hiebl, J. (2024). Artificial Intelligence (AI) Research Assistants in der Praxis. *Information – Wissenschaft & Praxis*, 75(1), 32–33. <https://doi.org/10.1515/iwp-2023-2045>

## **Sammlung von Papern mit mutmaßlichem Einsatz von KI**

<https://retractionwatch.com/papers-and-peer-reviews-with-evidence-of-chatgpt-writing/>

## **Artikel zum Vertrauen in KI-Output von Forschenden**

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002), 49–58. <https://doi.org/10.1038/s41586-024-07146-0>

# Weitere Ressourcen z.B. zu rechtlichen Rahmenbedingungen

## **Stellungnahme des DFG-Präsidiums**

Deutsche Forschungsgemeinschaft (Hrsg.). (2023, 21. September). Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG.

[https://www.dfg.de/download/pdf/dfg\\_im\\_profil/geschaeftsstelle/publikationen/stellungnahmen\\_papiere/2023/230921\\_stellungnahme\\_praesidium\\_ki\\_ai.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2023/230921_stellungnahme_praesidium_ki_ai.pdf)

## **Rechtsgutachten des Projektes KI:edu.nrw im Auftrag des Ministeriums für Kultur und Wissenschaft des Landes Nordrhein-Westfalen → didaktische und rechtliche Einordnung**

Salden, P., & Leschke, J. (2023). Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung. <https://doi.org/10.13154/294-9734>

## **Hintergrundpapier des Büros für Technikfolgenabschätzung des Deutschen Bundestags**

Albrecht, S. (2023). ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen (Hintergrundpapier Nr. 26). Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag. <https://doi.org/10.5445/IR/1000158070>

## **EU AI Act/KI Verordnung**

[https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=OJ:L_202401689)

## **Datenschutzvereinbarung Europa-USA:**

[https://ec.europa.eu/commission/presscorner/detail/de/ip\\_23\\_3721](https://ec.europa.eu/commission/presscorner/detail/de/ip_23_3721)

# Weitere Ressourcen zu Energieverbrauch

Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and Policy Considerations for Deep Learning in NLP* (arXiv:1906.02243). arXiv.

<https://doi.org/10.48550/arXiv.1906.02243>

Forumseintrag “How much energy consumption is involved in Chat GPT responses being generated?”

<https://ai.stackexchange.com/questions/38970/how-much-energy-consumption-is-involved-in-chat-gpt-responses-being-generated>

Weitere Links:

- <https://arxiv.org/abs/2309.14393>
- <https://tinymml.substack.com/p/the-carbon-impact-of-large-language>
- <https://www.piie.com/blogs/realtime-economics/2024/ais-carbon-footprint-appears-likely-be-alarming>