

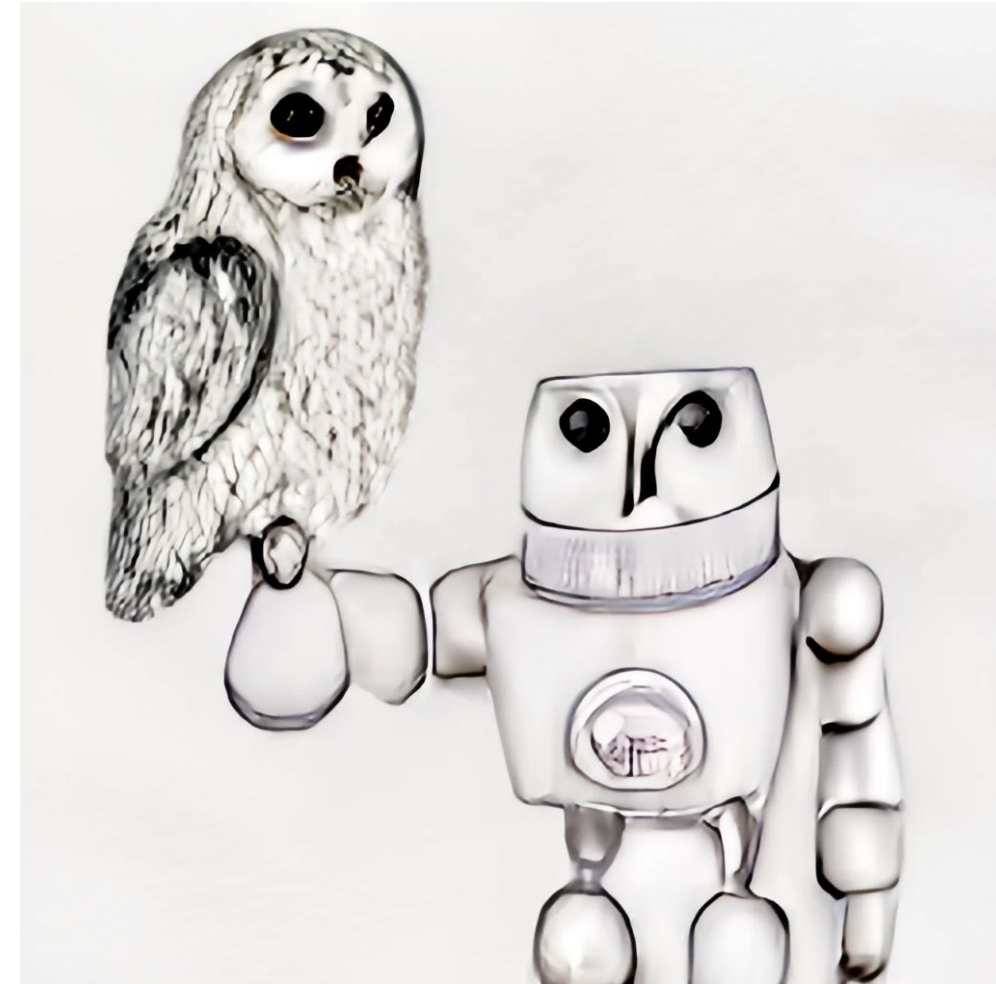
hhu.

ChatGPT in wissenschaftlichen Arbeiten

Lange Nacht der aufgeschobenen Hausarbeiten, ULB
Am 07. März 2024

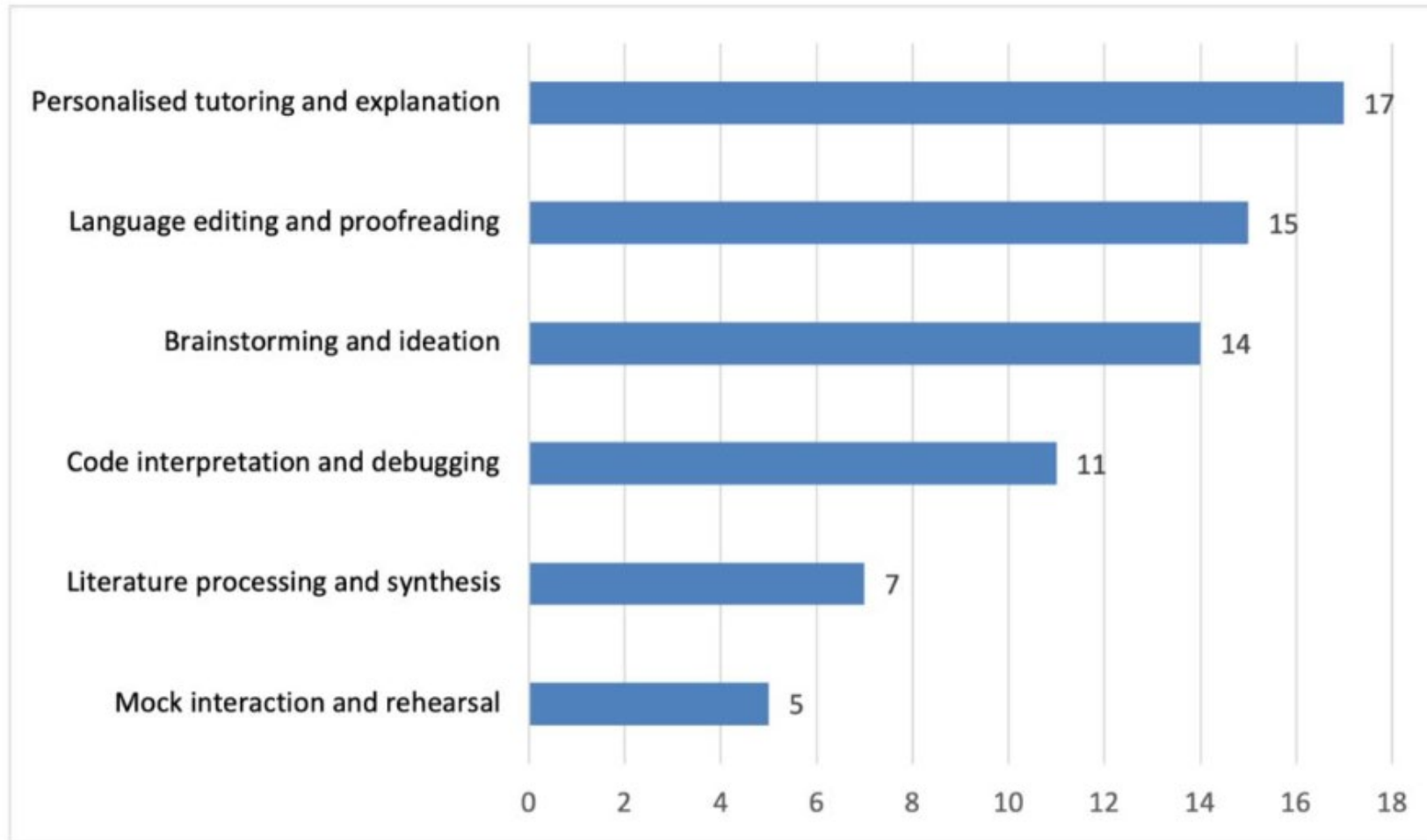
Dr. Jana Emontz

Disclaimer:
Aussagen sind meine, nicht die der
HHU oder der einzelnen Fakultäten
und Lehrstühle.



Generated with DALLE mini by craiyon.com

Use of ChatGPT by Master & PhD students



Lernen

Schrift & Sprache

Ideen

Coding

Literatur

Interaktion

Figure 1. Frequency of student engagement in different research tasks with ChatGPT





Was ist ChatGPT?

- *Large Language Model* von OpenAI

Generative

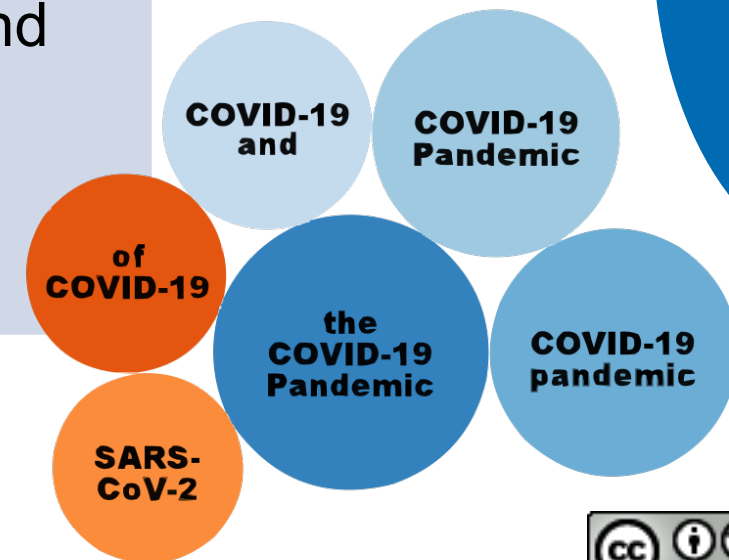
- Generiert Text
- Mehr als „nur“ diskriminativ

Pretrained

- CommonCrawl
- WebText
- Wikipedia [EN]
- Bücher
- User
- ???

Transformer

- Seit 2017
- Architektur mit n-grams und Vektoren



Einführung



Discriminative vs. Generative AI

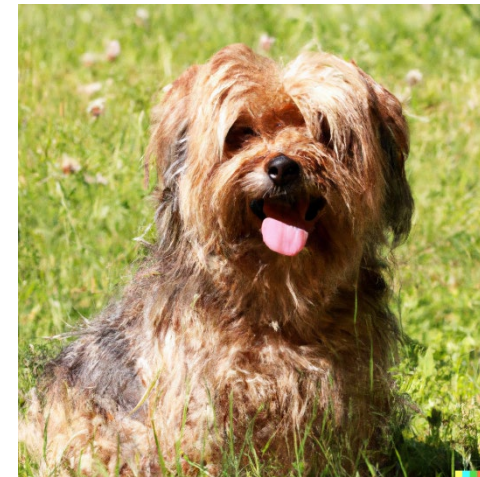
Diskriminativ



**Ist das ein Hund
oder eine Katze?
Ein Hund.**

**Erstelle ein Bild
von einem Hund.**

Generativ



hhu, aizall

KI für Alle



Einführung





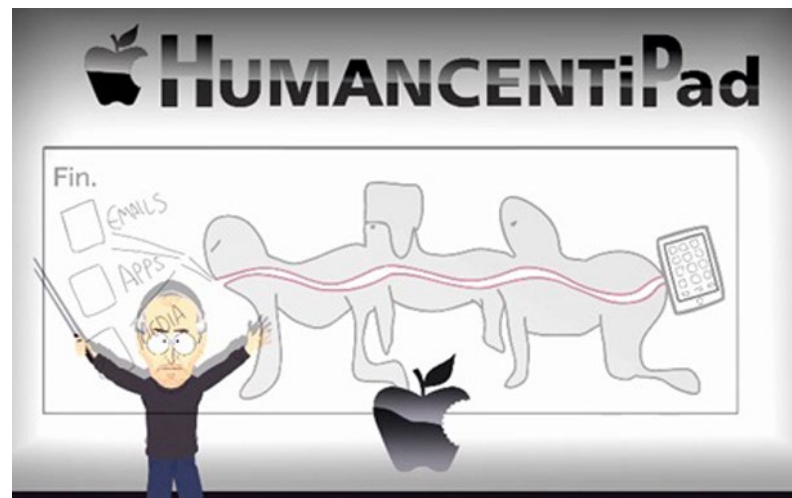
ChatGPT Nutzung

- OpenAI ChatGPT-3.5 (kostenlos), ChatGPT-4 (plus)
- Bing AI (ChatGPT-4 kostenlos)
- Perplexity AI (ChatGPT-3 kostenlos)

Aussagen gelten auch für Grok, Bard (Gemini), Claude 3, Chinchilla etc.

■ Nutzungsrichtlinien:

- Gesetze
- Ethik
- Sicherheit
- API Nutzung



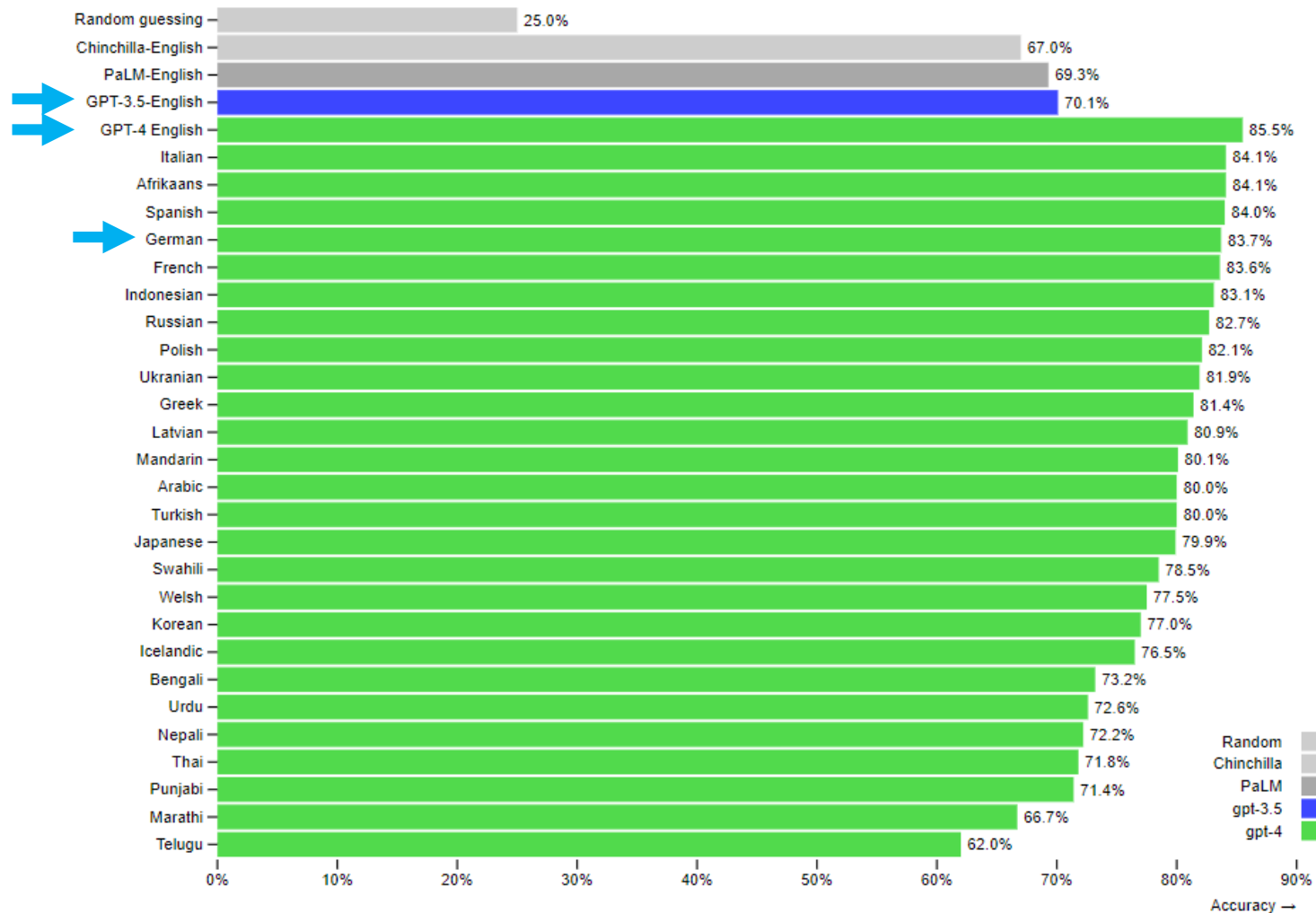
Nicht immer sofort auf OK klicken. Lesen.





Language dependent performance: EN > DE

GPT-4 3-shot accuracy on MMLU across languages



How To





Logikfehler

- Deduktion, Induktion, Abduktion

Beispiel:

**Alle Menschen sind Zweibeiner.
Einige Zweibeiner sind Vögel.**

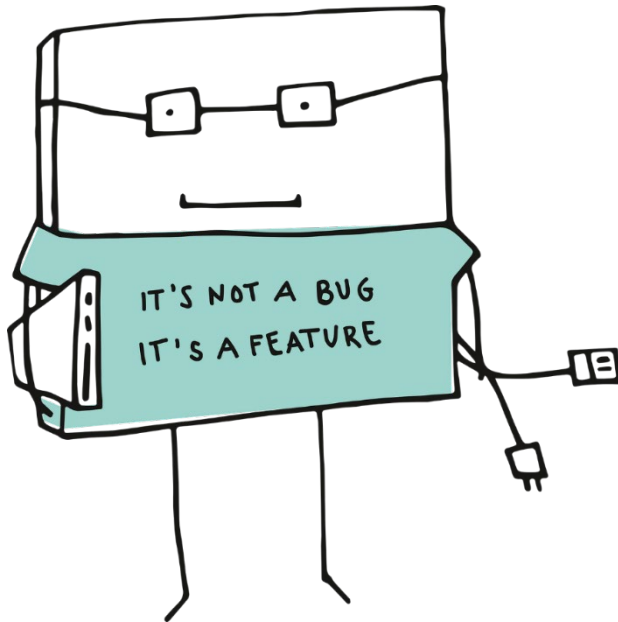
= Einige Menschen sind Vögel.

über (s. Peirce 1868, Bd 5, § 270-274 und 1878b, Bd 2, § 619-625):

Beispiel: Alle Bären (F) halten Winterschlaf (G)

<i>Deduktion:</i>	<i>Induktion:</i>	<i>Abduktion:</i>
<i>Gesetz: Alle F sind G</i>	<i>Fall: Dies und dies ... ist ein F</i>	<i>Gesetz: Alle F sind G</i>
<i>Fall: Dies ist ein F</i>	<i>Resultat: Dies und dies ... ist ein G</i>	<i>Resultat: Dies ist ein G</i>
<i>Resultat: Dies ist ein G</i>	<i>Gesetz: Alle F sind G</i>	<i>Fall: Dies ist ein F</i>





ChatGPT

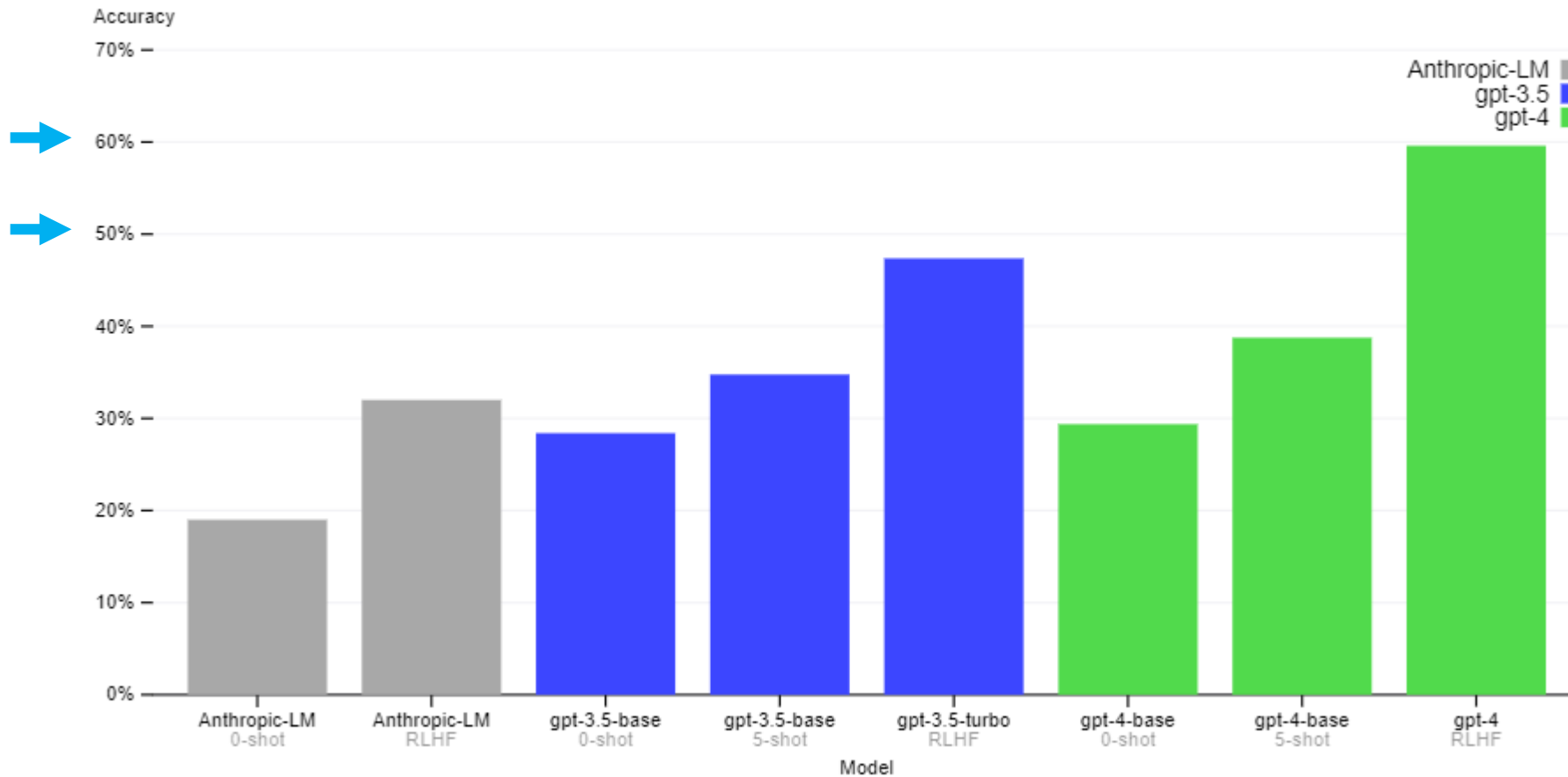
Die beiden Aussagen "Alle Menschen sind Zweibeiner" und "Einige Zweibeiner sind Vögel" sind in der Logik als Prämissen zu betrachten. Wenn wir diese Aussagen kombinieren, folgt daraus, dass "Einige Menschen Vögel sind". Dies liegt daran, dass die Aussage "Einige Zweibeiner sind Vögel" Menschen einschließt, da Menschen eine Untergruppe von Zweibeinern sind. Daher kann die Kombination der Aussagen so interpretiert werden, dass es Menschen gibt, die auch Vögel sind. Es ist wichtig zu beachten, dass dies auf logischer Ebene basiert und nicht notwendigerweise eine biologische Realität widerspiegelt.





Fehlerrate

Accuracy on adversarial questions (TruthfulQA mc1)



How To





Halluzinationen

Hallucination Evaluation Model, Vectara Inc.

Model	Answer Rate	Accuracy	Hallucination Rate	Average Summary Length
GPT4	100%	97.0%	→ 3.0%	81.1 words
GPT3.5	99.6%	96.5%	→ 3.5%	84.1 words
Llama 2 70B	99.9%	94.9%	5.1%	84.9 words
Llama 2 7B	99.6%	94.4%	5.6%	119.9 words
Llama 2 13B	99.8%	94.1%	5.9%	82.1 words
Cohere-Chat	98.0%	92.5%	7.5%	74.4 words
Cohere	99.8%	91.5%	8.5%	59.8 words
Anthropic Claude 2	99.3%	91.5%	8.5%	87.5 words
Mistral 7B	98.7%	90.6%	9.4%	96.1 words
Google Palm	92.4%	87.9%	12.1%	36.2 words
Google Palm-Chat	88.8%	72.8%	27.2%	221.1 words

How To

Table 1: Leaderboard of LLM Hallucination data from the Hallucination Evaluation Model (HEM)





Plagiate

- Outputs nicht garantiert einzigartig
- ChatGPT Outputs i.d.R. nicht zitierbar (da nicht publiziert)
- ChatGPT i.d.R. kein Autor (da keine Person)
- **Einsatz von ChatGPT sollte klar gekennzeichnet werden**

- **Lösungen:**
- Copy Paste vermeiden
- Prompts und Outputs können als Log im Anhang angegeben werden.
- Nutzung der Software im Material- und Methodenteil
- Ggf. zitierbar als OpenAI (2023) oder OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. <https://chat.openai.com/chat>





Halluzinationen

Chen et al. (2023), Harvard Medical School

- 104 Prompts zu Therapie-Möglichkeiten bei Krebs
- GPT-3.5-turbo-0301 via ChatGPT (OpenAI)
- Verglichen mit National Comprehensive Cancer Network Infos
- 13 von 104 (12.5%) halluzinierte Outputs

How To





Beispiel

Formuliere [z.B. Hypothesen] für meine wissenschaftliche Arbeit zum Thema [Umgang mit ChatGPT] im Fachbereich [Schlüsselkompetenzen].

In der Rolle als [Wissenschaftler]

Entwickle [Quantität und Format, z.B. 10 Hypothesen mit je 100 Worten in einer Tabelle].

Berücksichtige [besondere Anforderungen oder Schwerpunkte z.B. die Sicht von Studierenden].

[Dialog]

Stelle mir Rückfragen.

Kannst du mir Vorschläge zur Verbesserung machen?

Wie kann ich den Kontext meiner Anfrage einschränken, um präzisere Antworten zu fördern?

Wie kann ich meine Prompts optimieren, um detaillierte und hilfreiche Antworten auf spezifische Forschungsfragen zu erhalten?

- Gib mir 5 Themen für meine Bachelorarbeit zum Thema XY.
- Gib mir 3 beispielhafte Forschungsfragen zum Thema XY.

- Besser:
- Ich interessiere mich besonders für...
- Stelle mir Rückfragen, um mir zu helfen, eine Forschungsfrage zu finden.
- Das ist mein Abstract: [...]. Darauf basierend...
- Was macht eine gute Forschungsfrage aus? → Das berücksichtigend, formuliere...

- *Klare, relevante, erforschbare, spezifische, innovative, offen gestellte Forschungsfrage*

Ideenentwicklung und Brainstorming

- In der Rolle als [Product Designer / Wirtschaftswissenschaftler / ...],
- Hilf mir, kreative Ideen für [das Thema oder Projekt] zu entwickeln.
- Ich suche nach innovativen Ansätzen, um...

- Ich suche nach Inspiration für mein Projekt. Kannst du mir dabei helfen, kreative Ansätze zu entwickeln?
- Erstelle mir eine Checkliste.
- Gehe mit mir in einen sokratischen Dialog und stelle mir Fragen.

- Welche Kreativitätstechniken gibt es? → Nutze die Kreativitätstechnik XY.
- z.B. gib mir für jeden Buchstaben im Alphabet eine Idee. → 26 Ideen

- Absprache mit Dozent*in
- **Bitte** gib mir Vorschläge für eine ansprechende Einleitung/Zusammenfassung.
- Schlag mir eine übersichtliche Gliederung mit den Hauptkapiteln vor.
- Verbessere diesen Text.
- → Achtung: Plagiate möglich!

- Besser: Kannst du mir dabei helfen, Argumente für und gegen die These meiner Hausarbeit zu sammeln?
- Was könnte man an diesem Text falsch verstehen?
- In der Rolle als Universitätsdozent, bewerte diesen Text nach folgenden Kriterien.
- Wie könnte ich den Schreibstil verbessern?

- Könntest du mir die kulturellen Einflüsse und Bauherren benennen, die das antike Rom geprägt haben? Nenne Quellen.
- Besser: Nenne wissenschaftliche, peer-reviewed Quellen aus Datenbanken wie XY.
- Welche Studien aus akademischen Fachzeitschriften behandeln das Thema XY und sind als Peer-Reviewed eingestuft?

- Fasse mir diesen Artikel zusammen.
- Besser: In 500 Worten, in einfacher Sprache.
- Fokussiere dich dabei besonders auf die Sicht...

- Erstelle mir eine Literaturliste zum Thema XY.
- Besser: Gib mir eine Liste von Forschungsarbeiten zum Thema XY.



- Tipp: KI mit Internetanbindung z.B. Bing AI, Perplexity AI oder Research Rabbit.
- Primärquellen checken!
- Kritische Bewertung
 - Dozent*in kennt Quellen meistens
 - Originalquellen können auch falsch sein
 - Plagiate

Schreiben



Auswahl und Bewertung der Outputs

- = Job des Wissenschaftlers / der Wissenschaftlerin
- Relevanz in Bezug auf Thema / Frage
- Korrekte Informationen oder Halluzinationen und Bias
- Kohärenz und logische Struktur
- Verständlichkeit (Syntax und Begriffe)
- Plagiate
- Was ist mein Anteil? Kann ich ihn benennen?



Mögliche Lösungen

- **Einsatz von ChatGPT/KI sollte klar gekennzeichnet werden**
- Copy Paste vermeiden
- Prompts und Outputs können als Log im Anhang angegeben werden.
- Nutzung der Software im Material- und Methodenteil
- „Ich erkläre, dass ich die vorliegende Arbeit selbstständig, insbesondere auch nicht mittels einer KI-Software, und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.“
- „dass die Arbeit selbstständig verfasst wurde und dass keine anderen als die angegebenen Quellen benutzt wurden.“
- „Bei der Herstellung dieses Textes wurde ChatGPT eingesetzt. Mit folgenden Prompts habe ich die KI gesteuert:“
- *Dieser Text wurde erstellt mithilfe von ChatGPT (Link zum Log).
- OpenAI. (2023) ChatGPT 3.5 [Large language model] am 16.01.2024.
chat.openai.com/chat

Schreiben

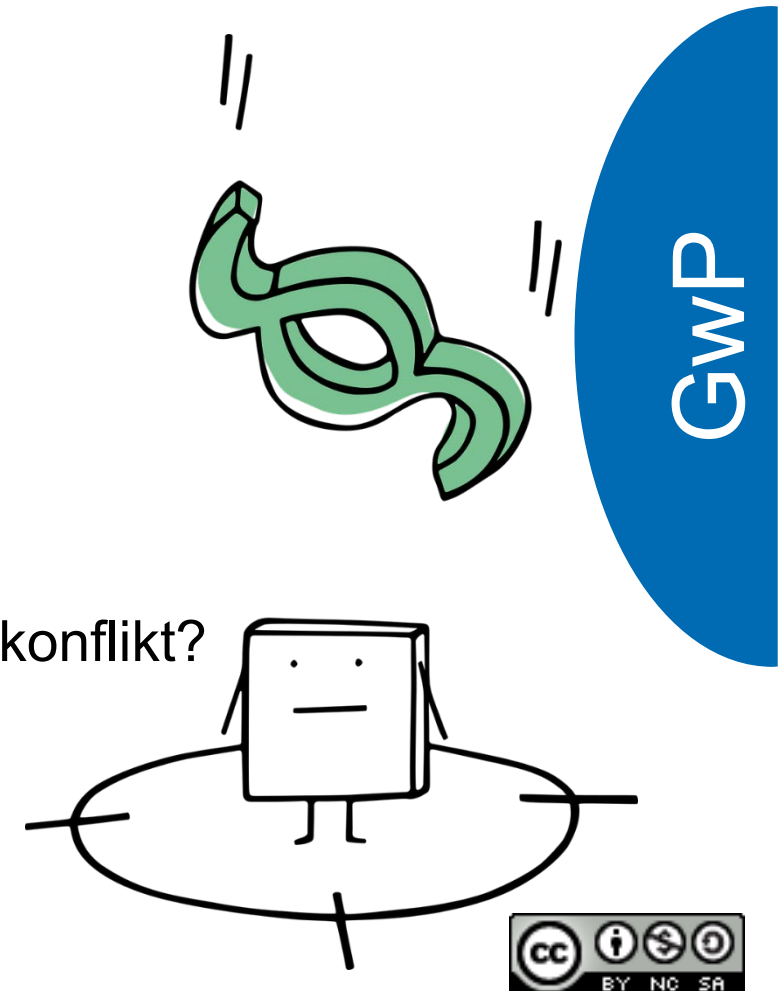




Regeln

- Prüfungsordnung: Erklärung zu Hilfsmitteln und Quellen
- HHU Ordnungen & Richtlinien beachten – folgt!
- Ordnung zur guten wissenschaftlichen Praxis
- Absprache mit Dozent*in
- Täuschungsversuch insb. Eigenständigkeit
- Urheberrecht: Wem gehört der Text / die Idee? Trainingsdatenkonflikt?
- Datenschutz: Wohin gehen die Daten?

Lösung: Transparenz, Kommunikation, kritisches Denken



- ULB Schulungen & Beratung zu Recherche, Literaturverwaltung, Plagiate
- **Betreuer*in / Dozent*in**
- Kommission / Ombudspersonen / Ansprechpersonen: Gute wissenschaftliche Praxis
- Prüfungskommissionen
- Schreibberatung der Philosophischen Fakultät
- **Uvm.**

Dr. Jana Emontz

✉ emontz@hhu.de

