

„Innovatives Lehren und Lernen mit KI: Theorie und Praxis“

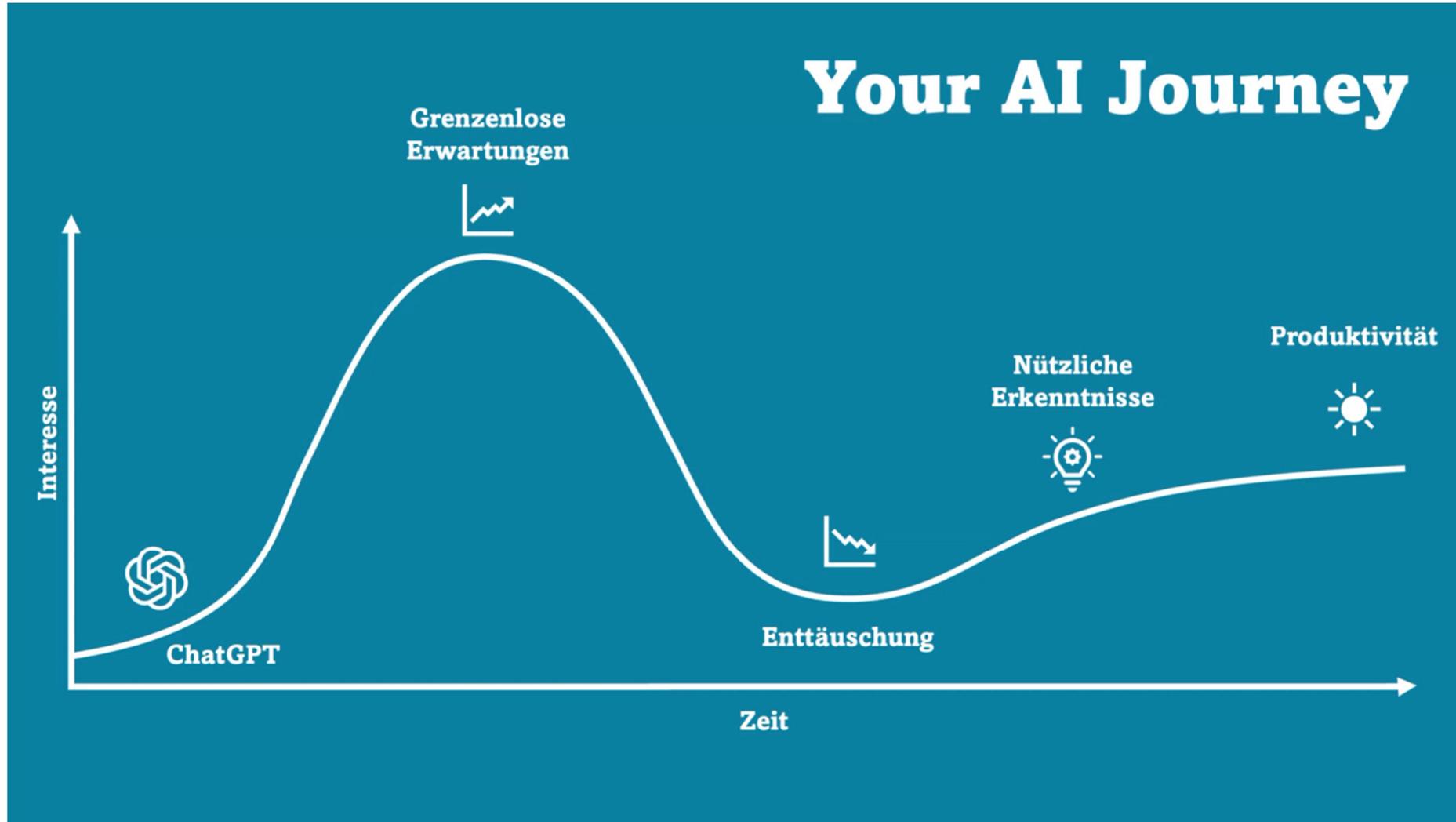
Aktuelle KI-Systeme und ihre Anwendbarkeit auf akademische Prozesse

Beginn: 17:00 Uhr

Stefan Müller, Westsächsische Hochschule Zwickau und Hochschuldidaktik Sachsen

20.03.24

„Hype Cycle“ (Persike 2024)



Frage an Studierende:

Welche Erfahrungen haben Sie mit KI-Chatbots und der Bearbeitung von Lernaufgaben gemacht?

-oftmals muss man in ChatGPT viele Iterationsschleifen durchlaufen, um auf ein brauchbares Ergebnis zu kommen

--> manchmal sehr mühsam, dann liest man es lieber direkt in den VL-Unterlagen nach bevor man ewig ausprobiert

Nicht für Lernaufgaben eingesetzt, aber für simple Aufgaben oder Fragenstellungen aus Eigeninteresse. Ich persönlich würde KI-Chatbots für Lernaufgaben nur begrenzt verwenden, weil die von der KI verwendeten Daten nicht immer korrekt sind.

Wenn ich KI benutzen würde bräuchte ich auch immer die Möglichkeit das Ergebnis zu evaluieren (ich muss wissen was raus kommen muss, das ist nicht immer der Fall)

Aktuelle KI-Systeme

“First, get access to the most advanced and largest Large Language Model you can get your hands on.”

Ethan Mollick

What people ask me most. Also, some answers.

WWW.ONEUSEFULTHING.ORG



Aktuelle KI-Systeme

kostenfreie KI-Chatbots

- Microsofts Copilot (GPT-4) – Verwendung der Hochschul-E-Mail-Adresse (Anmeldung mit Geschäfts-, Schul- oder Unikonto) und Auswahl des Unterhaltungsstils „Im höheren Maße Genau“
- Anthropics Claude 3 - noch nicht offiziell in Europa verfügbar (Registrierung mit einer E-Mail ohne „.de“-Endung und an Angabe einer Telefonnummer unter Region „international“)
- Googles Gemini - Anmeldung mit Google-Konto erforderlich

leistungsstärkere, kostenpflichtige KI-Chatbots

- OpenAIs ChatGPT Plus
- Googles Gemini Advanced

weitere KI-Chatbots

- Inflection AI's Pi
- Perplexity AI
- Mistral AI

Aktuelle KI-Systeme

(OpenAI 2023)

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)

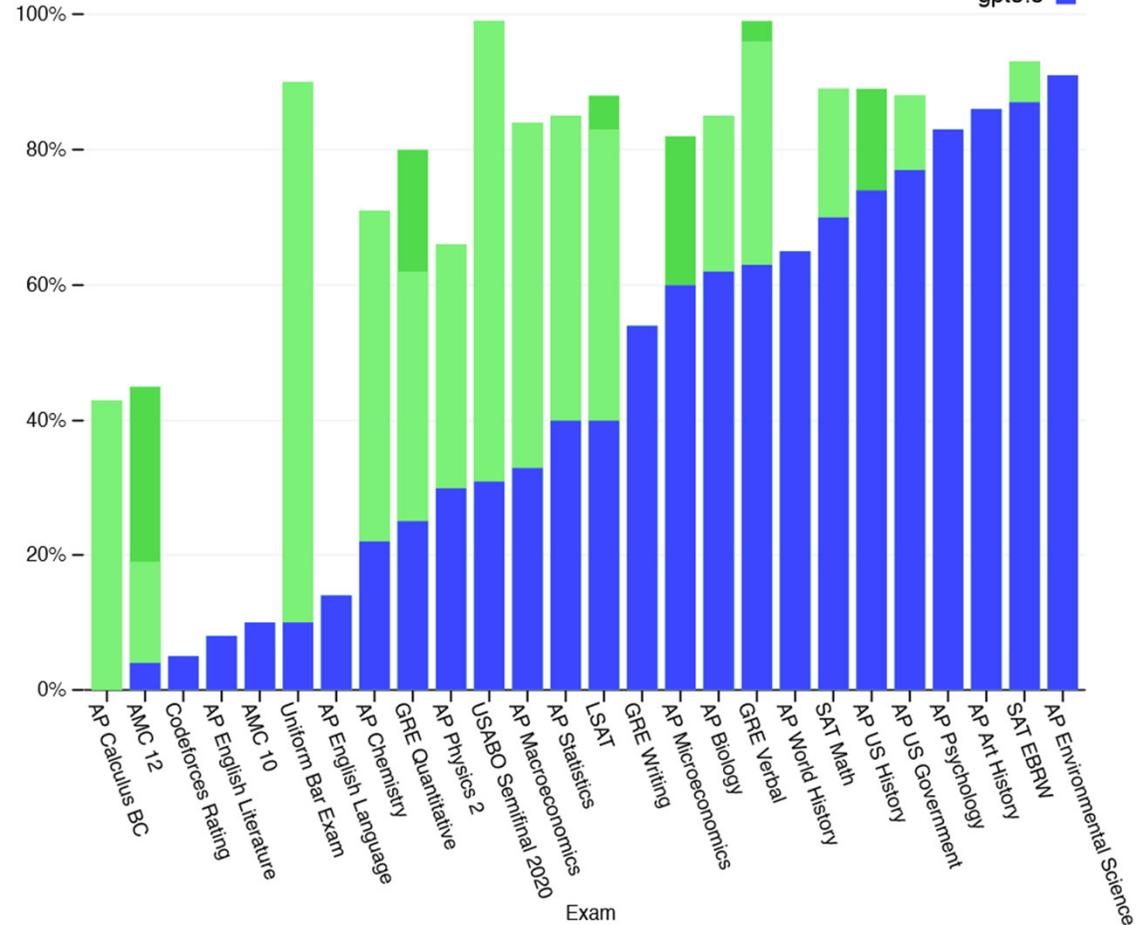


Figure 4. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. Exams are ordered from low to high based on GPT-3.5 performance. GPT-4 outperforms GPT-3.5 on most exams tested. To be conservative we report the lower end of the range of percentiles, but this creates some artifacts on the AP exams which have very wide scoring bins. For example although GPT-4 attains the highest possible score on AP Biology (5/5), this is only shown in the plot as 85th percentile because 15 percent of test-takers achieve that score.

Aktuelle KI-Systeme

(Anthropic 2024)

	Claude 3 Opus	Claude 3 Sonnet	Claude 3 Haiku	GPT-4	GPT-3.5	Gemini 1.0 Ultra	Gemini 1.0 Pro
Undergraduate level knowledge <i>MMLU</i>	86.8% 5-shot	79.0% 5-shot	75.2% 5-shot	86.4% 5-shot	70.0% 5-shot	83.7% 5-shot	71.8% 5-shot
Graduate level reasoning <i>GPQA, Diamond</i>	50.4% 0-shot CoT	40.4% 0-shot CoT	33.3% 0-shot CoT	35.7% 0-shot CoT	28.1% 0-shot CoT	—	—
Grade school math <i>GSM8K</i>	95.0% 0-shot CoT	92.3% 0-shot CoT	88.9% 0-shot CoT	92.0% 5-shot CoT	57.1% 5-shot	94.4% Maj1@32	86.5% Maj1@32
Math problem-solving <i>MATH</i>	60.1% 0-shot CoT	43.1% 0-shot CoT	38.9% 0-shot CoT	52.9% 4-shot	34.1% 4-shot	53.2% 4-shot	32.6% 4-shot
Multilingual math <i>MGS</i>	90.7% 0-shot	83.5% 0-shot	75.1% 0-shot	74.5% 8-shot	—	79.0% 8-shot	63.5% 8-shot
Code <i>HumanEval</i>	84.9% 0-shot	73.0% 0-shot	75.9% 0-shot	67.0% 0-shot	48.1% 0-shot	74.4% 0-shot	67.7% 0-shot
Reasoning over text <i>DROP, F1 score</i>	83.1 3-shot	78.9 3-shot	78.4 3-shot	80.9 3-shot	64.1 3-shot	82.4 Variable shots	74.1 Variable shots
Mixed evaluations <i>BIG-Bench-Hard</i>	86.8% 3-shot CoT	82.9% 3-shot CoT	73.7% 3-shot CoT	83.1% 3-shot CoT	66.6% 3-shot CoT	83.6% 3-shot CoT	75.0% 3-shot CoT
Knowledge Q&A <i>ARC-Challenge</i>	96.4% 25-shot	93.2% 25-shot	89.2% 25-shot	96.3% 25-shot	85.2% 25-shot	—	—
Common Knowledge <i>HellaSwag</i>	95.4% 10-shot	89.0% 10-shot	85.9% 10-shot	95.3% 10-shot	85.5% 10-shot	87.8% 10-shot	84.7% 10-shot

Aktuelle KI-Systeme

The screenshot shows the Chatbot Arena website. At the top, there is a navigation bar with links for 'Arena (battle)', 'Arena (side-by-side)', 'Direct Chat', 'Vision Direct Chat', 'Leaderboard', and 'About Us'. The main heading is 'Chatbot Arena: Benchmarking LLMs in the Wild'. Below this, there are links for 'Blog', 'GitHub', 'Paper', 'Dataset', 'Twitter', and 'Discord'. A section titled 'Rules' contains three bullet points: 'Ask any question to two anonymous models (e.g., ChatGPT, Claude, Llama) and vote for the better one!', 'You can continue chatting until you identify a winner.', and 'Vote won't be counted if model identity is revealed during conversation.' Below the rules is a section for the 'Arena Elo Leaderboard' with a text prompt: 'We collect 300K+ human votes to compute an Elo-based LLM leaderboard. Find out who is the LLM Champion!'. At the bottom, there is a 'Chat now!' button and a chat interface with two columns labeled 'Model A' and 'Model B'. A search bar at the top of the chat interface says 'Expand to see the descriptions of 33 models'.

Arena (battle) Arena (side-by-side) Direct Chat Vision Direct Chat Leaderboard About Us

Chatbot Arena: Benchmarking LLMs in the Wild

[Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#)

Rules

- Ask any question to two anonymous models (e.g., ChatGPT, Claude, Llama) and vote for the better one!
- You can continue chatting until you identify a winner.
- Vote won't be counted if model identity is revealed during conversation.

Arena Elo [Leaderboard](#)

We collect 300K+ human votes to compute an Elo-based LLM leaderboard. Find out who is the LLM Champion!

Chat now!

Expand to see the descriptions of 33 models

Model A Model B

Aktuelle KI-Systeme

kostenfreie KI-Chatbots

- Microsofts Copilot (GPT-4) – Verwendung der Hochschul-E-Mail-Adresse (Anmeldung mit Geschäfts-, Schul- oder Unikonto) und Auswahl des Unterhaltungsstils „Im höheren Maße Genau“
- Anthropics Claude 3 - noch nicht offiziell in Europa verfügbar (Registrierung mit einer E-Mail ohne „.de“-Endung und an Angabe einer Telefonnummer unter Region „international“)
- Googles Gemini - Anmeldung mit Google-Konto erforderlich

leistungsstärkere, kostenpflichtige KI-Chatbots

- OpenAIs ChatGPT Plus
- Googles Gemini Advanced

weitere KI-Chatbots

- Inflection AI's Pi
- Perplexity AI
- Mistral AI

← **Post**

 **Leopold Aschenbrenner** ✓
@leopoldasch

One year since GPT-4 release. Hope you all enjoyed some time to relax; it'll have been the slowest 12 months of AI progress for quite some time to come.

[Post übersetzen](#)

Zuletzt bearbeitet 6:13 vorm. · 16. März 2024 · **634.750** Mal angezeigt

256 Reposts 137 Zitate **2.056** „Gefällt mir“-Angaben **330** Lesezeichen

    330 

<https://twitter.com/leopoldasch/status/1768868127138549841>

Aktuelle KI-Systeme

HOCHSCHULE >

Service

Für Unternehmen

Für Medien

Hochschulbibliothek

Hochschuldidaktik ▾

Veranstaltungen

Beratung

Empfehlungen zum Umgang mit
generativer KI

Werkzeugkasten digitale Lehre >

Werkzeugkasten Hochschullehre >

Lehr-Lern-Projekte

1-Minute-Paper

TAP und BiLOE (Modulevaluation)

Thementag Gute Lehre >

Studierkompetenz >

Corporate Design

Hochschulshop

Historische Sammlung >

Empfehlungen zum Umgang mit generativer KI

„Der Umgang mit generativer KI an Hochschulen ist mit Unsicherheiten behaftet, die nicht zuletzt aufgrund der stetigen Weiterentwicklung generativer KI nie vollständig ausgeräumt werden können. Fakt ist: Generative KI Tools werden nicht verschwinden und in Hochschulen genutzt. Hochschulweite Leitlinien zum Umgang mit generativer KI können Orientierung bieten. Diese Leitlinien sollten dynamisch sein, um flexibel auf Veränderungen reagieren zu können.“ (Julius Friedrich, Projektleiter HFD)

Mit den Empfehlungen zum Umgang mit generativer KI soll Lehrenden an der WHZ die Nutzung von KI erleichtert werden. Struktur und Inhalte der Empfehlungen orientiert sich an dem [HFD-Blickpunkt „Leitlinien zum Umgang mit generativer KI“](#). Darin wurden KI-Leitlinien deutscher Hochschulen analysiert und wesentliche Elemente aufbereitet.

Einstieg und Einordnung +

Datenschutz und Urheberrecht +

Kompetenzen +

Lehr- und Lernaktivitäten +

Prüfungen +

KI-Zugänge -

Die Westsächsische Hochschule Zwickau kann, wie die meisten Hochschulen, gegenwärtig keinen kostenfreien und datenschutzkonformen Zugang zu KI-Systemen anbieten. In einzelnen Struktureinheiten wird lediglich der Einsatz von Einzellizenzen für ChatGPT Plus erprobt. Darüber hinaus können Hochschulangehörige nicht dazu verpflichtet werden, sich bei KI-Systemen anzumelden. Die Nutzung von generativer KI kann nur freiwillig erfolgen.

Für eine freiwillige Nutzung eignen sich die kostenfreien Zugänge zu [Microsofts Copilot \(GPT-4\)](#), [Anthropics Claude 3](#) und [Googles Gemini](#). In allen Fällen ist eine Anmeldung erforderlich. Für Microsofts Copilot kann die Hochschul-E-Mail-Adresse verwendet werden (Anmeldung mit Geschäfts-, Schul- oder Unikonto) und es sollten der Unterhaltungsstil "Im höheren Maße Genau" ausgewählt werden. Da Anthropics Claude 3 noch nicht offiziell in Europa verfügbar ist, muss in der Registrierung eine E-Mail ohne ".de"-Endung angegeben werden. Außerdem sollte bei der Angabe der Telefonnummer die Region "international" ausgewählt werden.

Sollten Hochschulangehörige nach leistungsstärkeren, kostenpflichtigen KI-Systemen suchen, wird [OpenAIs ChatGPT Plus](#) und [Googles Gemini Advanced](#) empfohlen.

letzte Änderung am 18.03.2024

Aktuelle KI-Systeme und ihre Anwendbarkeit auf akademische Prozesse

“First, get access to the most advanced and largest Large Language Model you can get your hands on.”

Ethan Mollick

What people ask me most. Also, some answers.

WWW.ONEUSEFULTHING.ORG



“Then use it to do everything that you are legally and ethically allowed to use it for. Generating ideas? Ask the AI for suggestions. In a meeting? Record the transcript and ask the AI to summarize action items. Writing an email? Work on drafting it with AI help. My rule of thumb is you need about 10 hours of AI use time to understand whether and how it might help you.”

Ethan Mollick

What people ask me most. Also, some answers.

WWW.ONEUSEFULTHING.ORG



Akademische Aufgaben – Entwicklung von Forschungsprojekten

Du bist ein sehr erfahrener Forscher. Ich soll ein Exposé für ein kleines Forschungsprojekt schreiben. Hilf mir dabei und berate mich.

Lass uns entlang der folgenden Schritte vorgehen:

1. Themenfindung

Frage mich nach dem Themengebiet, in dem ich forschen möchte. Hilf mir, mein Themengebiet einzugrenzen. Warte meine Antworten ab. Gehe erst zum nächsten Schritt über, wenn ich mit meiner Auswahl zufrieden bin. Berücksichtige für den nächsten Schritt die bis dahin getroffene Auswahl.

2. Entwicklung einer Forschungsfrage

Schlage mir Forschungsfragen zu meinem gewählten Thema vor. Beginne mit einer umfassenden Formulierung und hilf mir, von einer weitreichenden Forschungsfrage über Einschränkungen und Spezifizierungen zu einer spezifischen Forschungsfrage zu gelangen. Warte meine Antworten ab. Lass uns erst zum nächsten Schritt weitergehen, wenn ich zufrieden mit der getroffenen Auswahl bin. Berücksichtige für den nächsten Schritt die bis dahin getroffene Auswahl.

3. Identifikation relevanter Publikationen im Zusammenhang mit der Forschungsfrage

Schlage mir relevante Publikationen im Zusammenhang mit der Forschungsfrage vor. Warte meine Antworten ab. Lass uns erst zum nächsten Schritt weitergehen, wenn ich zufrieden mit der getroffenen Auswahl bin. Berücksichtige für den nächsten Schritt die bis dahin getroffene Auswahl.

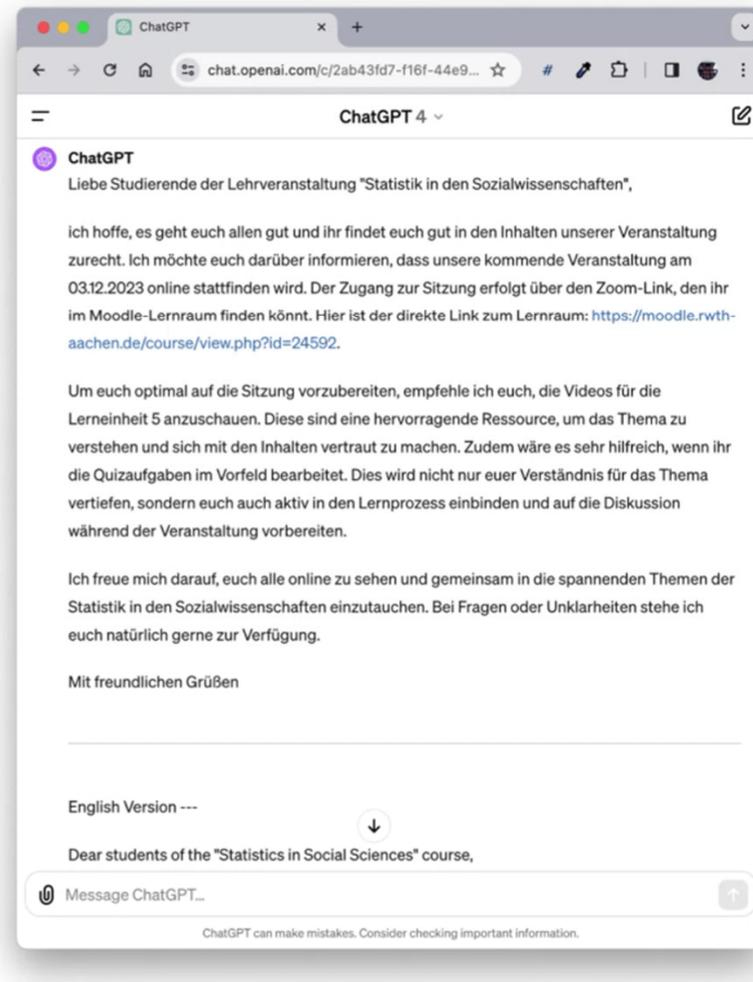
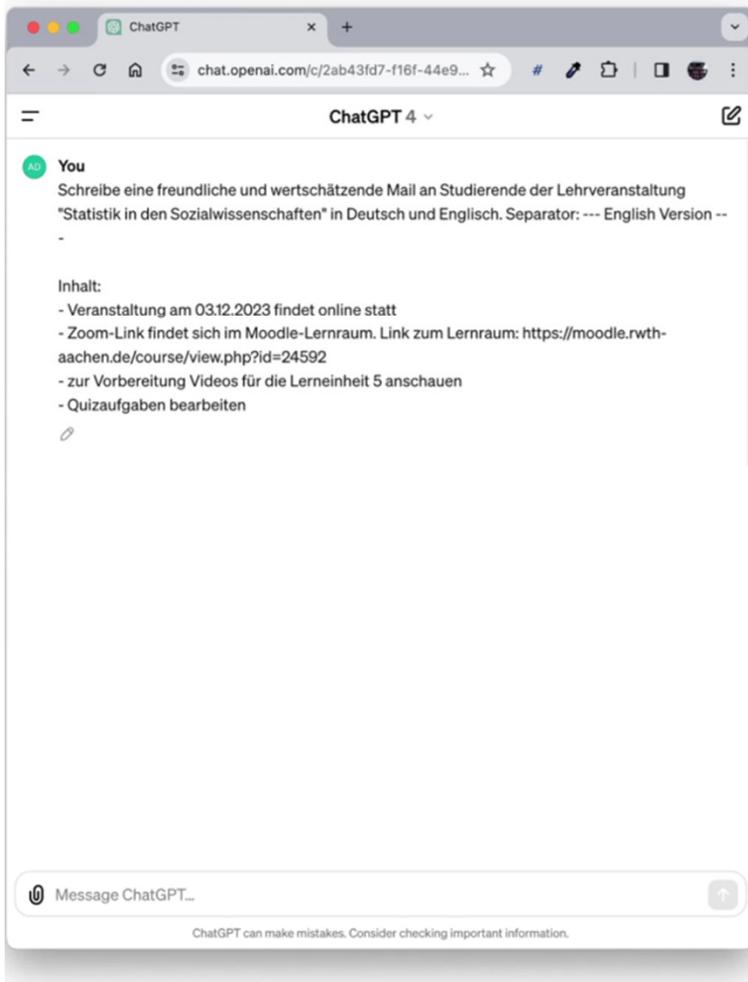
4. Auswahl der Forschungsmethodik

Schlage mir Forschungsmethoden vor, mit denen die Forschungsfrage bearbeitet werden kann. Benenne mögliche Vor- und Nachteile der Forschungsmethoden. Lass uns erst zum nächsten Schritt weitergehen, wenn ich zufrieden mit der getroffenen Auswahl bin. Berücksichtige für den nächsten Schritt die bis dahin getroffene Auswahl.

5. Erstellung Exposé

Schreibe ein ausführliches Exposé, das die getroffene Auswahl von Thema, Forschungsfrage, Publikationen und Forschungsmethodik in jeweils einem Absatz beschreibt. Verfasse das Exposé ausschließlich als Fließtext. Verwende für das Exposé mindestens 500 Wörter. Schreibe in einem präzisen, wissenschaftlichen Stil.

Akademische Aufgaben – E-Mail-Kommunikation (Persike 2024)



Akademische Aufgaben – E-Mail-Kommunikation

Von: [REDACTED]@gmail.com <[REDACTED]@gmail.com>

Gesendet: Dienstag, 9. Januar 2024 00:06

An: Pressestelle <Pressestelle@fh-zwickau.de>

Betreff: Unterstützung eines Forschungsprojekts zum Thema Lernverhalten und digitale Lernmethoden

Sehr geehrte Damen und Herren,

Ich hoffe, diese E-Mail erreicht Sie wohlbehalten. Mein Name ist [REDACTED] (Student der Informatik B.Sc. & Physik B.Sc. an der TU München) und gemeinsam mit [REDACTED] (Student der Wirtschaftspsychologie B.Sc. an der Hochschule Fresenius Heidelberg) untersuchen wir im Rahmen eines Forschungsprojekts das allgemeine Lernverhalten sowie die Nutzung digitaler Lernmethoden von Lernenden.

Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews

Weixin Liang^{1*} Zachary Izzo^{2*} Yaohui Zhang^{3*} Haley Lepp⁴ Hancheng Cao^{1,5} Xuandong Zhao⁶
Lingjiao Chen¹ Haotian Ye¹ Sheng Liu⁷ Zhi Huang⁷ Daniel A. McFarland^{4,8,9} James Y. Zou^{1,3,7}

Abstract

We present an approach for estimating the fraction of text in a large corpus which is likely to be substantially modified or produced by a large language model (LLM). Our maximum likelihood model leverages expert-written and AI-generated reference texts to accurately and efficiently examine real-world LLM-use at the corpus level. We apply this approach to a case study of scientific peer review in AI conferences that took place after the release of ChatGPT: *ICLR 2024*, *NeurIPS 2023*, *CoRL 2023* and *EMNLP 2023*. Our results suggest that between 6.5% and 16.9% of text submitted as peer reviews to these conferences could have been substantially modified by LLMs, i.e. beyond spell-checking or minor writing updates.

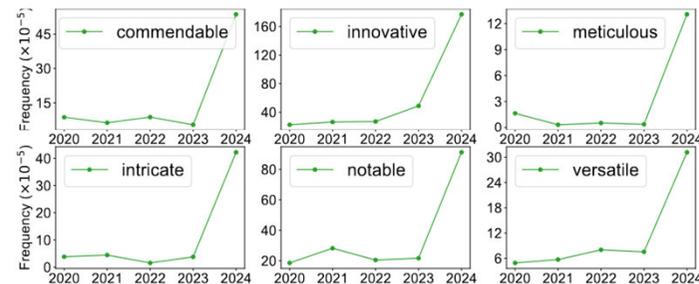
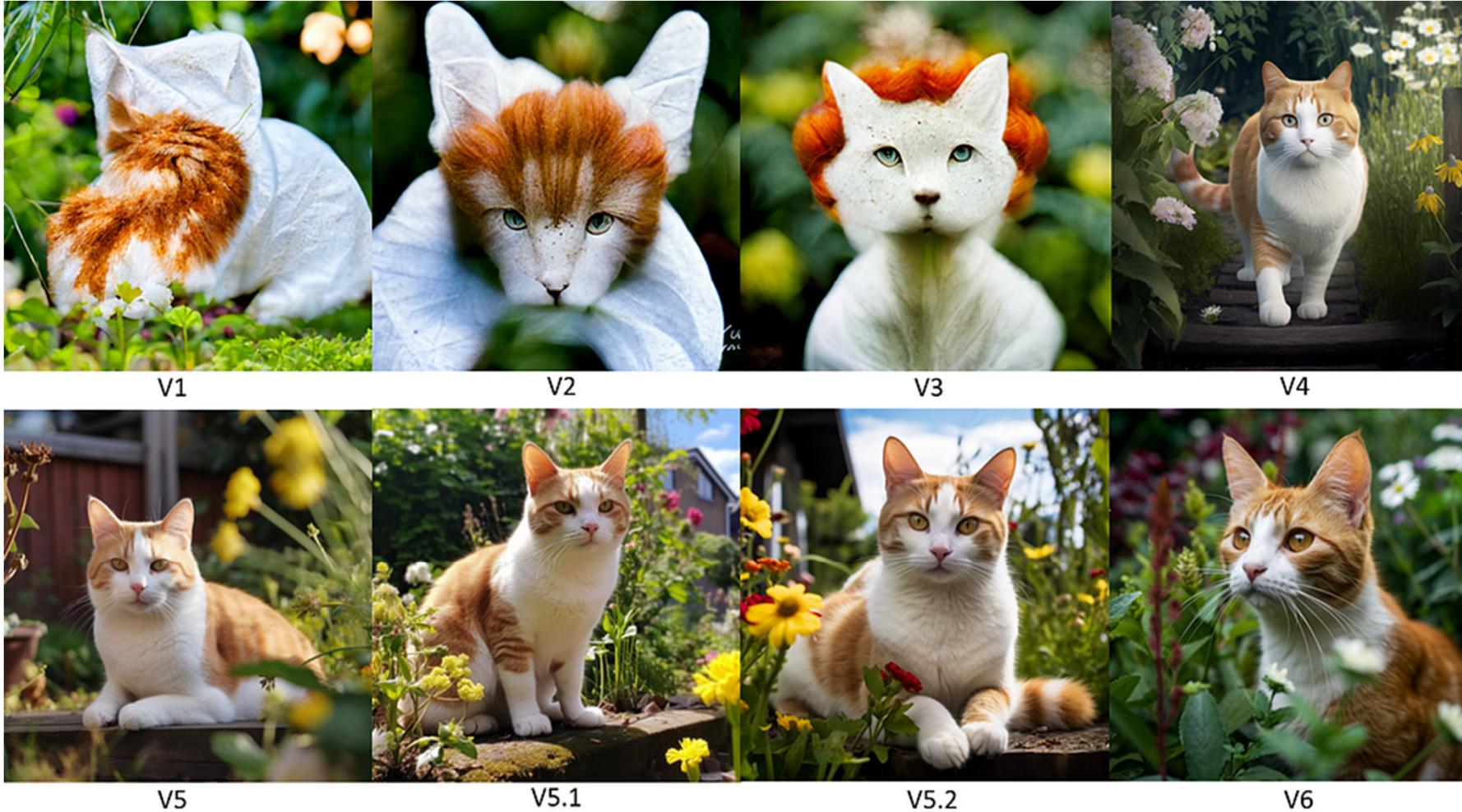


Figure 1: **Shift in Adjective Frequency in *ICLR 2024* Peer Reviews.** We find a significant shift in the frequency of certain tokens in *ICLR 2024*, with adjectives such as “commendable”, “meticulous”, and “intricate” showing 9.8, 34.7, and 11.2-fold increases in probability of occurring in a sentence. We find a similar trend in *NeurIPS* but not in *Nature Portfolio* journals. Supp. Table 2 and Supp. Figure 12 in the Appendix provide a visualization of the top 100 adjectives produced disproportionately by AI.

Midjourney



**Images generated on Midjourney by Henrique Centieiro and Bee Lee*

Release date:

V1, February 2022

V2, April 2022

V3, July 2022

V4, November 2022

V5, March 2023

V5.1, May 2023

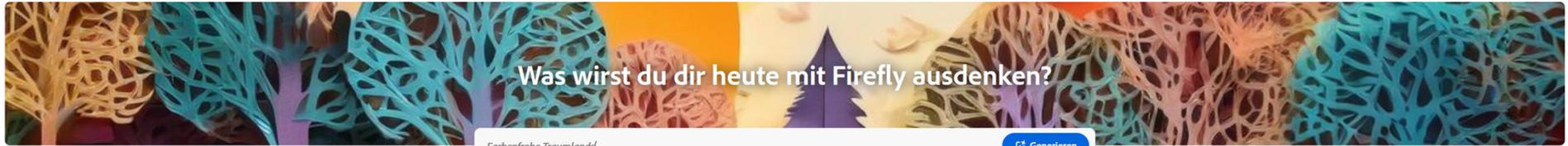
V5.2, June 2023

V6, Dec 2023

Prompt:

Photo of a ginger and white cat in the garden

Adobe Firefly



Jetzt ausprobieren

Probiere die neueste generative KI aus und sag uns, was du denkst.



Text zu Bild

Erstelle Bilder aus einer detaillierten Beschreibung und beeinflusse den Stil mit einem Referenzbild.

Generieren



Generatives Füllen

Du kannst Gegenstände mit einem Pinsel entfernen oder sogar neue aus Textbeschreibungen malen.

Generieren



Texteffekte

Wende Stile oder Texturen auf Text an, indem du sie beschreibst.

Generieren



Generative Neufärbung

Gib eine genaue Textbeschreibung ein und lass Adobe Firefly Farbvariationen deiner Vektorgrafik generieren.

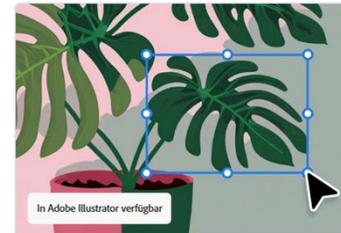
Generieren



Text zu Vorlage

Generiere bearbeitbare Vorlagen, indem du sie detailliert als Text beschreibst.

In Express testen



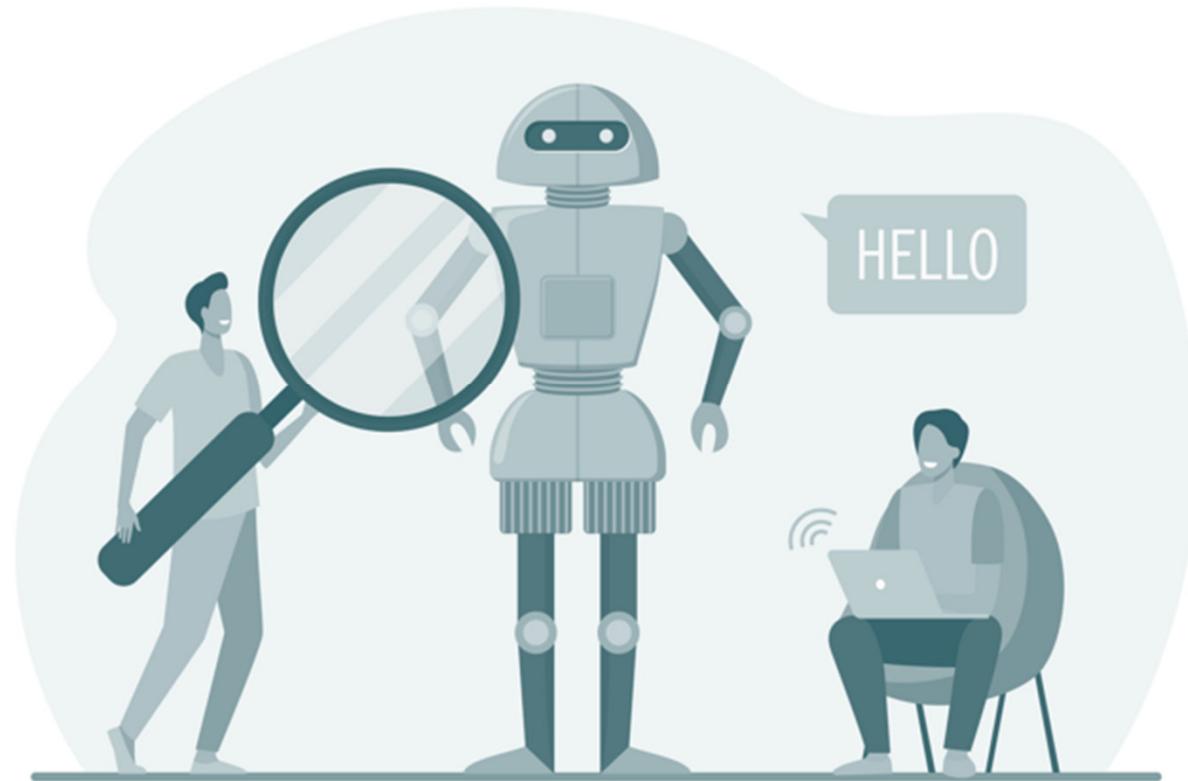
Text zu Vektorgrafik

Generiere SVG-Dateien, indem du sie detailliert als Text beschreibst.

In Adobe Illustrator testen

Hochschuldidaktische Perspektiven auf KI-Technologien (Amlung et al. 2023)

„Die akademische Gemeinschaft gewinnt an Bedeutung. Die Dynamik in der Entwicklung von KI-Technologien stellt das einseitige Verständnis von Lehrenden als Wissensträger:innen im Lehr-Lern-Prozess ebenso in Frage, wie den Anspruch an die Hochschuldidaktik, für jede didaktische Herausforderung im Zusammenhang mit KI-Technologien Lösungen anzubieten. Der hohe Grad an Agilität in Lehre und Studium kann nur kollaborativ, d. h. akteur:innen-, rollen- und hochschulübergreifend, in einer lernenden, akademischen Gemeinschaft gelingen – die über die nötigen Strukturen und Ressourcen verfügt.“



VIELEN DANK FÜR DEN AUSTAUSCH



Stefan Müller

WHZ/HDS

stefan.mueller@hd-sachsen.de