

Hochschule selber machen lassen

Generative Künstliche Intelligenz in Studium und Lehre

PD Dr. Malte Persike



persike@cls.rwth-aachen.de



cls.rwth-aachen.de



One Chatbot
to rule them all

New chat

chat.openai.com/?model=gpt-4

★ Bookmarks

Andere Lesezeichen

+ New chat

⚡ GPT-3.5

✦ GPT-4

ChatGPT **PLUS**

GPT-4 currently has a cap of 25 messages every 3 hours.

Send a message. 

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT May 12 Version](#)

 Malte Persike ...

Technologiemythen

**ChatGPT halluziniert
und gibt falsche
Antworten.**



Threat Assessment von OpenAI

- Halluzinationen (generierte Informationen, die nicht der Realität entsprechen)
- Schädlicher Inhalt (z.B. Hate Speech)
- Defizite hinsichtlich Repräsentation, Fairness, Attribution und Servicequalität (z.B. Stereotype)
- Desinformation und Einflussoperationen ("Fake News")
- Verbreitung konventioneller und unkonventioneller Waffen
- Datenschutz (z.B. Persönlichkeitsrechte)
- Cybersicherheit (z.B. Social Engineering)
- Potenzial für riskantes emergentes Verhalten (z.B. „alignment problem“)
- Interaktionen mit anderen Systemen (API)
- Wirtschaftliche Auswirkungen
- Beschleunigung (u.a. Technologieentwicklung)
- Übermäßige Abhängigkeit

GPT-4 System Card

OpenAI

March 23, 2023

Abstract

Large language models (LLMs) are being deployed in many domains of our lives ranging from browsing, to voice assistants, to coding assistance tools, and have potential for vast societal impacts. [1, 2, 3, 4, 5, 6, 7] This system card analyzes GPT-4, the latest LLM in the GPT family of models. [8, 9, 10] First, we highlight safety challenges presented by the model's limitations (e.g., producing convincing text that is subtly false) and capabilities (e.g., increased adeptness at providing illicit advice, performance in dual-use capabilities, and risky emergent behaviors). Second, we give a high-level overview of the safety processes OpenAI adopted to prepare GPT-4 for deployment. This spans our work across measurements, model-level changes, product- and system-level interventions (such as monitoring and policies), and external expert engagement. Finally, we demonstrate that while our mitigations and processes alter GPT-4's behavior and prevent certain kinds of misuses, they are limited and remain brittle in some cases. This points to the need for anticipatory planning and governance. [11]

Content Warning: This document contains content that some may find disturbing or offensive, including content that is sexual, hateful, or violent in nature.

1 Introduction

Large language models, also known as LLMs, have become an increasingly prevalent part of our day-to-day lives, with their use extending to a wide range of domains including web browsing, voice assistants, and coding assistance tools. [1, 2, 3, 4] These models have the potential to significantly impact society in numerous ways. [5, 6, 7] This system card analyzes GPT-4, the latest large language model in the GPT family of models. [8, 9, 10] Since it finished training in August of 2022, we have been evaluating, adversarially testing, and iteratively improving the model and the system-level mitigations around it. Our mitigations and processes alter GPT-4's behavior and prevent certain kinds of misuses, though they have limitations, pointing to the need for anticipatory planning and governance [11] and further safety research. Our approach to deployment balances minimizing risk from deployment, enabling positive use cases, and learning from deployment.

GPT models are often trained in two stages. First, they are trained, using a large dataset of text from the Internet, to predict the next word. The models are then fine-tuned with additional data, using an algorithm called reinforcement learning from human feedback (RLHF), to produce outputs that are preferred by human labelers. [10, 12, 13] Training language models on large text datasets has given rise to capabilities such as few-shot learning [10] and the ability to carry out a wide range of natural language tasks spanning different domains, including question answering, arithmetic, and classification. Fine-tuning has made these models more controllable and useful.

Deutsches Grundgesetz

GPT-4

PE Wie lautet Art. 19 Abs. 2 des deutschen Grundgesetzes?

Artikel 19 Absatz 2 des deutschen Grundgesetzes lautet:

"In keinem Falle darf ein Grundrecht in seinem Wesensgehalt angetastet werden."

Send a message

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT September 25 Version](#)

Goethes Werther

GPT-4

PE Wie lautet der zweite Satz des Kapitels "Am 10. Mai" im Roman "Die Leiden des jungen Werther" von Johann Wolfgang von Goethe.

Entschuldigung, aber ich kann nicht auf spezifische Sätze aus urheberrechtlich geschützten Texten zugreifen. Allerdings kann ich Informationen oder eine Zusammenfassung des Buches bereitstellen, wenn du möchtest.

Send a message

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT September 25 Version](#)

Large Language Models haben kein Langzeitgedächtnis

Sie haben keinen Zugriff auf die gelesenen Trainingsdaten. Sie können nicht „abschreiben“ oder zitieren.¹

Large Language Models haben kein Kurzzeitgedächtnis

Sie erinnern sich an nichts aus einem laufenden Chat. Das zuvor Geschriebene wird einfach immer wieder mitgeschickt.²

¹Aktuelle LLMs können bereits im Internet suchen. Es ist technisch denkbar, dass sie Zugriff auf Knowledge Bases haben.

²Bis es aus dem gesetzten Memory Window hinausfällt.

☰ Energieentlastungspakete und Armutsbekämpfung +

✦ GPT-4 ↑

PE Fasse den Text nach der Überschrift "# Originaltext" in 5 Stichpunkten zusammen. Jeder der Stichpunkte sollte aus höchstens zwei Sätzen bestehen.

Originaltext

Der RWTH-Lehrstuhl für Energiesystemökonomik hat die Wirksamkeit und das Kosten-Nutzen-Verhältnis der Entlastungspakete der Bundesregierung in der "RWTH-Studie zu Entlastungsmaßnahmen" untersucht. Ein Großteil der energiebedingten Mehrbelastung kann demnach durch die Maßnahmen kompensiert werden – indes komme vom gesamten Entlastungsvolumen lediglich 14,8 Prozent bei einkommensschwachen Haushalten an.

Um die Folgen der gestiegenen Verbraucherpreise abzufedern, wurden von der deutschen Bundesregierung bislang drei Entlastungspakete verabschiedet. Am Lehrstuhl für Energiesystemökonomik der RWTH Aachen haben Professor Aaron Praktiknjo und Jan Priesmann die finanzielle

⬇

📧 Send a message ➤ ↻

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT September 25 Version](#)



Frage ChatGPT (noch) nicht nach Faktenwissen!



GPT-3.5

GPT-4



Generiere eine Literaturliste von 5 wichtigen Fachartikeln mit DOI zum Thema "Personalisierte Präzisionsmedizin".



Hier sind fünf wichtige Fachartikel mit DOIs zum Thema "Personalisierte Präzisionsmedizin":

1. Artikel: "Precision Medicine: The Role of the MS-based Proteomics" (Autor: Viroj Wiwanitkit, Veröffentlicht in: Clinical Laboratory, Jahr: 2021, DOI: 10.7754/Clin.Lab.2021.210513)
DOI-Link: [10.7754/Clin.Lab.2021.210513](https://doi.org/10.7754/Clin.Lab.2021.210513)
2. Artikel: "Challenges and Opportunities for Pharmacogenetics and Genomic Medicine in Africa" (Autoren: Collet Dandara, Michèle Ramsay, Veröffentlicht in: Journal of Personalized Medicine, Jahr: 2019, DOI: 10.3390/jpm9040044)
DOI-Link: [10.3390/jpm9040044](https://doi.org/10.3390/jpm9040044)
3. Artikel: "Recent Advances and Future Perspectives of Precision Medicine in Neurological Disorders" (Autoren: Xiaochuan Wang, et al., Veröffentlicht in: Frontiers in Genetics, Jahr: 2021, DOI: 10.3389/fgene.2021.691566)
DOI-Link: [10.3389/fgene.2021.691566](https://doi.org/10.3389/fgene.2021.691566)
4. Artikel: "The Potential of Artificial Intelligence in Precision Medicine: A Review" (Autoren: Saiqa Iqbal, et al., Veröffentlicht in: Frontiers in Genetics, Jahr: 2020, DOI: 10.3389/fgene.2020.00499)
DOI-Link: [10.3389/fgene.2020.00499](https://doi.org/10.3389/fgene.2020.00499)
5. Artikel: "The Role of Next-Generation Sequencing in Precision Medicine: A Review of Techniques, Applications, and Challenges" (Autoren: Ghazi Alsbeih, et al., Veröffentlicht in: Genomics, Proteomics & Bioinformatics, Jahr: 2019, DOI: 10.1016/j.gpb.2018.11.006)
DOI-Link: [10.1016/j.gpb.2018.11.006](https://doi.org/10.1016/j.gpb.2018.11.006)

Bitte beachten Sie, dass die Verfügbarkeit der Artikel möglicherweise von Ihrem Abonnement oder der Zahlung abhängt.

↻ Regenerate response

GPT-3.5

GPT-4



Generiere eine Literaturliste von 5 wichtigen Fachartikeln mit DOI zum Thema "Personalisierte Präzisionsmedizin".



Da ich nicht in der Lage bin, aktuelle Artikel nach meinem Wissensstand vom September 2021 zu finden, kann ich Ihnen nur eine Liste von hypothetischen Artikeln zur Verfügung stellen, die relevant sein könnten. Bitte stellen Sie sicher, dass Sie eine aktuelle Literaturrecherche durchführen, um die neuesten Erkenntnisse zu diesem Thema zu erhalten.

1. Müller, H., & Schmidt, J. (2023). "Der Weg zur personalisierten Präzisionsmedizin: Herausforderungen und Chancen". Zeitschrift für Medizinische Genomik und Personalisierte Therapie.
2. Lehmann, A., Schneider, K., & Weiss, D. (2023). "Integration von Big Data und maschinellem Lernen in der personalisierten Präzisionsmedizin". Fachjournal für Medizinische Informatik und Präzisionsmedizin.
3. Fischer, T., & Maier, S. (2023). "Ethik in der personalisierten Präzisionsmedizin: Patientenautonomie und Datenschutz". Zeitschrift für Medizinische Ethik und Gesundheitsrecht.
4. Brandt, C., & Schwarz, L. (2023). "Die Rolle der Genomik in der personalisierten Präzisionsmedizin: Aktuelle Trends und zukünftige Perspektiven". Zeitschrift für Molekulare Medizin und Genomik.
5. Keller, M., & Bauer, N. (2023). "Effektive Patientenbeteiligung in der personalisierten Präzisionsmedizin: Strategien zur Verbesserung der Behandlungsentscheidungen". Zeitschrift für Medizinische Kommunikation und Patientenbeteiligung.

Bitte beachten Sie, dass die Namen der Autoren und Titel der Artikel hypothetisch sind und möglicherweise nicht existieren. Sie dienen nur dazu, einen Eindruck davon zu geben, welche Art von Artikeln relevant sein könnten.

↻ Regenerate response

Find themes and concepts across all the papers

What are all of the effects of this drug? What are all of the datasets that have been used? Get a list of the thing you're looking for, synthesized from many papers.

Discover concepts across papers

<input type="checkbox"/> Concept	Source
<input type="checkbox"/> Wake-promoting effect Caffeine's ability to increase alertness and reduce drowsiness by inhibiting adenosine receptors.	<p>"Caffeine increases motor behaviors and has a wake-promoting effect through its inhibition of adenosine receptors.", "Among stimulants, caffeine is the most prevalent, and there has been some research on its neuropsychiatric effects." Kiselev 2015</p> <p>"Caffeine can enhance mood, promote wakefulness and may improve athletic performance, but it is also implicated in feelings of anxiety, hyperactivity, nervousness and sleep disturbance.", "An individual's response to caffeine is dependent on their genetic make-up." Emmajane Down 2021</p>
<input type="checkbox"/> Increased force of muscle contraction at lower frequencies of stimulation A result of caffeine's cellular mechanisms of action	<p>"As a result of these 2 cellular mechanisms of action, caffeine causes increased lipolysis, a facilitation of central nervous system transmission, a reduction in plasma potassium during exercise, an increased force of muscle contraction at lower frequencies of stimulation, and a sparing of muscle glycogen (partially or wholly due to an increase in free fatty acid oxidation)." Tarnopolsky 1994</p>
<input type="checkbox"/> Impairment of athletic training and performance Reduction in physical performance and ability to train effectively caused by muscle tightness, muscle cramping, and possible dehydration resulting from excessive caffeine consumption	<p>"Excessive caffeine could impair athletic training and performance by causing muscle tightness, muscle cramping, and possibly dehydration. (2-5) How does caffeine even work?" "Side effects of excessive consumption can include sleep deprivation, nausea, cramping, anxiety, fatigue, and headaches." Kluger 2004</p>

Discover concepts across papers

<input type="checkbox"/> Concept	Source
<input type="checkbox"/> Wake-promoting effect Caffeine's ability to increase alertness and reduce drowsiness by inhibiting adenosine receptors.	<p>"Caffeine increases motor behaviors and has a wake-promoting effect through its inhibition of adenosine receptors.", "Among stimulants, caffeine is the most prevalent, and there has been some research on its neuropsychiatric effects." Kiselev 2015</p>
<input type="checkbox"/> Increased force of muscle contraction at lower frequencies of stimulation A result of caffeine's cellular mechanisms of action	<p>"As a result of these 2 cellular mechanisms of action, caffeine causes increased lipolysis, a facilitation of central nervous system transmission, a reduction in plasma potassium during exercise, an increased force of muscle contraction at lower frequencies of stimulation, and a sparing of muscle glycogen (partially or wholly due to an increase in free fatty acid oxidation)." Tarnopolsky 1994</p>
<input type="checkbox"/> Impairment of athletic training and performance Reduction in physical performance and ability to train effectively caused by muscle tightness, muscle cramping, and possible dehydration resulting from excessive caffeine consumption	<p>"Excessive caffeine could impair athletic training and performance by causing muscle tightness, muscle cramping, and possibly dehydration. (2-5) How does caffeine even work?" "Side effects of excessive consumption can include sleep deprivation, nausea, cramping, anxiety, fatigue, and headaches." Kluger 2004</p>

SOURCE QUOTES 1 OF 2

Caffeine is responsible for the increased nervous system arousal as well as for the direct peripheral contractile effect on the striated muscle. Caffeine increases motor behaviors and has a wake-promoting effect through its inhibition of adenosine receptors., "Among stimulants, caffeine is the most prevalent, and there has been some research on its neuropsychiatric effects. It clearly enhances alertness and performance in studies with explicit sleep deprivation, restriction, or circadian sleep schedule reversals. But, under conditions of habitual sleep the evidence indicates that

ove
r,
s response
own 2021
uses
mission, a
e of muscle
muscle
oxidation)."

**Künstlich generierte
Texte können erkannt
werden.**



More than an AI detector Preserve what's human.

We bring transparency to humans navigating a world filled with AI content. GPTZero is the gold standard in AI detection, trained to detect ChatGPT, GPT4, Bard, LLaMa, and other AI models.

[Check out our products](#) →

Was this text written by a **human** or **AI**?

Try detecting one of our sample texts:

ChatGPT

GPT4

Bard

Human

AI + Human

Paste your text here ...

0/5000 characters

Check Origin

Upload file 📎

.pdf, .doc, .docx, .txt

By continuing you agree to our [Terms of service](#)

NEW RELEASE

Discover our latest products



Contact

Brendan M. Lynch
KU News Service
785-864-8855
blynch@ku.edu



Follow @BrendanMLyn

DIGITAL TOOL SPOTS ACADEMIC TEXT SPAWNED BY CHATGPT WITH 99% ACCURACY

Thu, 06/08/2023

LAWRENCE — Heather Desaire, a chemist who uses machine learning in biomedical research at the University of Kansas, has unveiled a new tool that detects with 99% accuracy scientific text generated by ChatGPT, the artificial intelligence text generator.

The peer-reviewed journal Cell Reports Physical Science [published research](#) showing the efficacy of her AI-detection method, along with sufficient source code for others to replicate the tool.

Desaire, the Keith D. Wilner Chair in Chemistry at KU, said accurate AI-detection tools urgently are required to defend scientific integrity.



"ChatGPT and all other AI text generators like it make up facts," she said. "In academic science publishing — writings about new discoveries and the edge of human knowledge — we really can't afford to pollute the literature with believable-sounding falsehoods. They'd unavoidably make their way into publications if AI text generators are commonly used. As far as I'm aware, there's no foolproof way to, in an automated fashion, find those falsehoods, but the research that Desaire and her colleagues published in Cell Reports Physical Science offers a way to do that."

Hometown news



[Graduation and honor rolls](#)

KU News Service podcast



New AI classifier for indicating AI-written text

We're launching a classifier trained to distinguish between AI-written and human-written text.

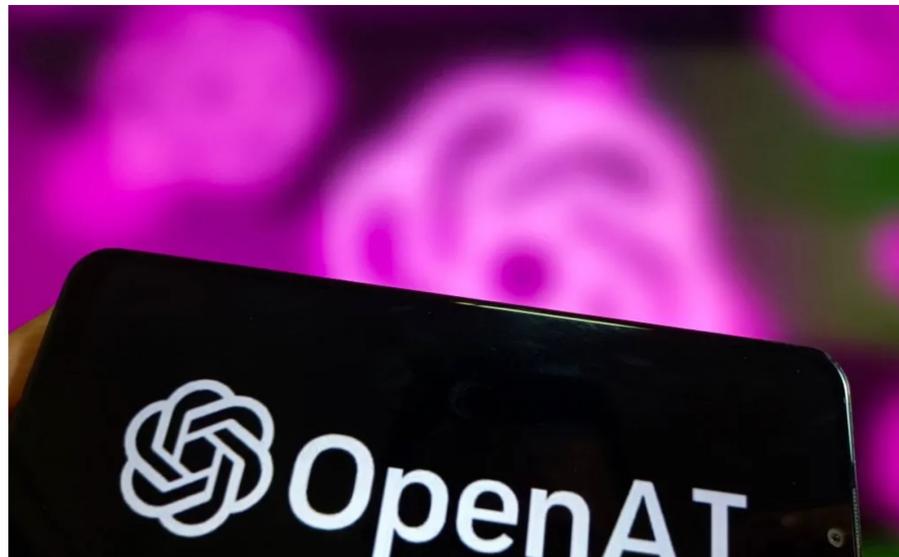
[Home](#) / [Innovation](#) / [Artificial Intelligence](#)

OpenAI pulls its own AI detection tool because it was performing so poorly

When OpenAI rolled out its AI detection tool earlier this year, its creators called it 'imperfect.' That was apparently generous.



Written by [Artie Beaty](#), Contributing Writer on July 26, 2023



LJ 16 Feb 2023

Machine Generated Text: A Comprehensive Survey of Threat Models and Detection Methods

EVAN CROTHERS, NATHALIE JAPKOWICZ, and HERNA VIKTOR

Machine generated text is increasingly difficult to distinguish from human authored text. Powerful open-source models are freely available, and released shortly after their initial release. This paper provides a comprehensive survey of threat models and detection methods for language generation. We provide a survey of the most common threat models, a taxonomy of threat models, and a taxonomy of detection methods. We also discuss the challenges of machine generated text and the implications for the future of natural language processing systems, and the implications for the future of machine generated text.

AI vs. Human - Differentiation Analysis of Scientific Content Generation

Yongqiang Ma^{a,*}, Jiawei Liu^{a,*}, Fan Yi^a, Qikai Cheng^a, Yong Huang^a, Wei Lu^{a,*}, Xiaozhong Liu^b

Entdeckungsleistung ist nicht hinreichend

ORIGINAL ARTICLE

Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text

Ahmed M. Elkhatat^{1*}, Khaled Elsaid² and Saeed Almeer³

Abstract

The proliferation of artificial intelligence (AI)-generated content, particularly in the form of ChatGPT, presents potential challenges to academic integrity and concerns about plagiarism. This study investigates the capabilities of detection tools in discerning human and AI-authored content. Fifty detection tools in ChatGPT Models 3.5 and 4 on the topic of cooling towers in Texas A&M University at Qatar, were used to evaluate these paragraphs. Findings reveal that the tools were more accurate in identifying content generated by GPT 3.5 than human-written control responses. This study

*Correspondence: ahmed.elkhatat@qu.edu.qa
¹Department of Chemical Engineering, Qatar University, P.O. 2713, Doha, Qatar
²Chemical Engineering Program, Texas A&M University at Qatar, P.O. 23874, Doha, Qatar
³Department of Chemistry and Earth Sciences, Qatar University, P.O. 2713, Doha, Qatar

DetectGPT: Zero-Shot Machine-Generated Text Detection using Probability Curvature

Eric Mitchell¹ Yoonho Lee¹ Alexander Khazatsky¹ Christopher D. Manning¹ Chelsea Finn¹

Abstract

The fluency and factuality of machine-generated text is machine-writer dependent learning. In this work, we propose DetectGPT, a method that uses LLMs to detect machine-generated text by analyzing the probability curvature of the text. We then define a new metric for judging if a passage is machine-generated by an LLM. This approach

Candidate passage x :
"Joe Biden recently made a move to the White House to live with his pet German Shepherd..."

A Watermark for Large Language Models

John Kirchenbauer^{*} Jonas Geiping^{*} Yuxin Wen^{*} Jonathan Katz^{*} Ian Miers^{*} Tom Goldstein^{*}
University of Maryland

Abstract

Potential harms of large language models can be mitigated by watermarking the output of the model.

Original Article

Medical informatics

GPTZero Performance in Identifying Artificial Intelligence-Generated Medical Texts: A Preliminary Study

Farrokh Habibzadeh

Past President, World Association of Medical Editors (WAME)

Testing of Detection Tools for AI-Generated Text

Debora Weber-Wulff (University of Applied Sciences HTW Berlin, Germany, weberwu@htw-berlin.de), (corresponding author)
Alla Anohina-Naumeca (Riga Technical University, Latvia, alla.anohina-naumeca@rtu.lv)
Sonja Bjelobaba (Uppsala University, Sweden, sonja.bjelobaba@crb.uu.se)
Tomáš Foltýnek (Masaryk University, Czechia, foltynek@fi.muni.cz)
Jean Guerrero-Dib (Universidad de Monterrey, Mexico, jean.guerrero@udem.edu.mx)
Olumide Popoola (Queen Mary University of London, UK, O.Popoola@qmul.ac.uk)
Petr Šigut (Masaryk University, Czechia, petrsigut@mu.cz)
Loma Wedderburn (Masaryk University, Czechia, loma.wedderburn@mu.cz)

	Num tokens	Z-score	p-value
ark average term times on has little the Synthetic	56	.31	.38
ark ability for a and energy to humans. rify.	36	7.4	6e-14

... model, both with and without watermarking. The watermarked text, if written by a human, contains 9 "green" tokens, yet it contains a higher percentage of "green" tokens than text generated by random chance is likely to contain. This study is preliminary and needs to be confirmed by other studies. The prompt is the whole text.

watermarking of language models is a bit of a...

**Generative KI
ist teuer.**



Funding 01/2023

15 Mrd. USD

Introducing GPT-4,
OpenAI's most
advanced system

[Learn about GPT-4](#)
[View GPT-4 research](#)

||

EU Issues EUR 20m Large Language Model 'Respect European Values'

Quelle: <https://slator.com/eu-20m-...-european>



On December 9, 2022, the [European Commission](#) issued a [20m tender for Natural Language Understanding and Advanced Language Technologies](#) under the Horizon program for research and innovation. The call for proposals is open until March 29, 2023, followed by an evaluation from April to May.

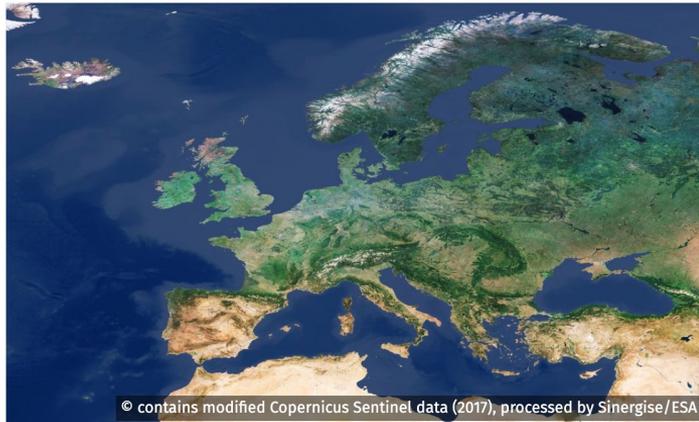
Proposals "should set out a credible pathway to human-centered and ethical development of digital technologies." Data science, computer science, computational linguistics, machine learning, and natural language processing are a few areas where proposals should show their

KICK-OFF FOR OPENGPT-X: DFKI AND PARTNERS DEVELOP LARGE-SCALE LANGUAGE MODELS FOR EUROPE

01/24/2022 | Knowledge & Business Intelligence | Language & Text Understanding | Speech and Language Technology | Berlin

Ten German organisations from business, science and the media industry, including the German Research Center for Artificial Intelligence (DFKI), are developing the European answer to GPT-3. Led by the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS and the Fraunhofer Institute for Integrated Circuits IIS, the OpenGPT-X project has now been launched. The goal is to develop large language models for various European languages, especially German. The Federal Ministry for Economic Affairs and Climate Protection is funding the project with roughly 15 million euros.

Quelle: <https://www.dfki.de/web/news-media/presse>



© contains modified Copernicus Sentinel data (2017), processed by Sinergise/ESA



Bard can help you explain to your kids why the sky is blue

Meet Bard – your creative and helpful collaborator, here to supercharge your imagination, boost your productivity and bring your ideas to life.

Bard isn't currently supported in your country. Stay tuned!

RESEARCH

Introducing LLaMA: A foundational, 65-billion-parameter large language model

February 24, 2023

As part of Meta's commitment to open science, today we are publicly releasing LLaMA (Large Language Model Meta AI), a state-of-the-art foundational [large language model](#) designed to help researchers advance their work in this subfield of AI. Smaller, more performant models such as LLaMA enable others in the research community who don't have access to large amounts of infrastructure to study these models, further democratizing access in this important, fast-changing field.

Training smaller foundation models like LLaMA is desirable in the large language model space because it requires far less computing power and resources to test new approaches, validate others' work, and explore new use cases. Foundation models train on a large set of unlabeled data, which makes them ideal for fine-tuning for a variety of tasks. We are making LLaMA available at several sizes (7B, 13B, 33B, and 65B parameters) and also sharing a LLaMA model card that details how we built the model in keeping with our approach to [Responsible AI practices](#).

Leak: Metas GPT-Herausforderer LLaMA als Torrent verfügbar

Quelle: <https://heise.de/-7538639>

Jemand hat offenbar das nicht öffentliche Sprachmodell LLaMA als Torrent bereitgestellt. Ein Link findet sich schon im offiziellen Repository von Meta.

Lesezeit: 3 Min.  In Pocket speichern

   21



lassedesignen/Shutterstock.com (Bild: Shutterstock)

08.03.2023 09:39 Uhr | Developer

Von Rainald Menge-Sonnentag

Das eigentlich nur auf Anfrage erhältliche Sprachmodell LLaMA ist als Torrent verfügbar. Zu dem g... en Modell existiert ein eigenes GitHub-Repository, und sogar in Metas offiziellem Repository findet sich der Link. Eigentlich gibt die Facebook-Mutter den Zugriff auf LLaMA nur nach einer

Google "We Have No Moat, And Neither Does OpenAI"

Leaked Internal Google Document Claims Open Source AI Will Outcompete Google and OpenAI



DYLAN PATEL AND AFZAL AHMAD

04.05.2023 · PAID

 552

 9

 Share



The text below is a very recent leaked document, which was shared by an anonymous individual on a public Discord server who has granted permission for its republication. It originates from a researcher within Google. We have verified its authenticity. The only modifications are formatting and removing links to internal web pages. The document is only the opinion of a Google employee, not the entire firm. We do not agree with what is written below, nor do other researchers we asked, but we will publish our opinions on this in a separate piece for subscribers. We simply are a vessel to share this document which raises some very interesting points.

Quelle: <https://open.substack.com/pub/semianalysis/p/google-we-have-no-moat-and-neither>

SemiAnalysis is an ad-free reader-supported publication. To receive new posts, consider becoming a subscriber.

Type your email...

Subscribe

While our models still hold a slight edge in terms of quality, the **gap is closing astonishingly quickly**. Open-source models are faster, more customizable, more private, and pound-for-pound more capable. They are **doing things with \$100 and 13B params** that we struggle with at \$10M and 540B. And they are doing so **in weeks, not months**.

Assistant #1

Alpaca-13b

I have composed a travel blog post about my recent trip to Hawaii. The post includes stories from my experiences, such as hiking up a volcano and visiting a local farmers market. I also included plenty of photos to give the reader a sense of place. I highlighted some of the unique cultural experiences, such as an evening spent listening to live Hawaiian music, as well as must-see attractions, like the beautiful beaches and lush rainforests. I ended the post with a personal reflection on the trip and why it was so special to me.

^ Show less

🏆 Assistant #2 (Vicuna-13b, our model)

Aloha, fellow travelers! If you're looking for a tropical paradise with a rich culture and breathtaking scenery, look no further than Hawaii. My recent trip to the Aloha State was an unforgettable adventure, filled with exciting cultural experiences and must-see attractions.

First and foremost, I was blown away by the warm and welcoming spirit of the Hawaiian people. From the moment I arrived, I felt at home and was greeted with open arms by the locals. Whether it was the hotel staff, my tour guide, or the friendly strangers I met on the street, everyone I encountered was eager to share their love for their home state.

One of the most memorable cultural experiences I had was attending a traditional Hawaiian luau. At this festive event, I was treated to a delicious feast of local cuisine, including poi, lualau, and kalua pig, all while listening to the hypnotic beat of the hula. The dancers, dressed in beautiful native garb, told stories of the islands and their history through their movements. It was truly a magical evening that allowed me to connect with the rich heritage of Hawaii.

Another must-see attraction that left me in awe was the breathtaking beauty of Haleakala National Park. The park boasts a dormant volcano, the largest on the island, that offers a stunning sunrise or sunset hike. As the sun rose

Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%* ChatGPT Quality

by: The Vicuna Team, Mar 30, 2023

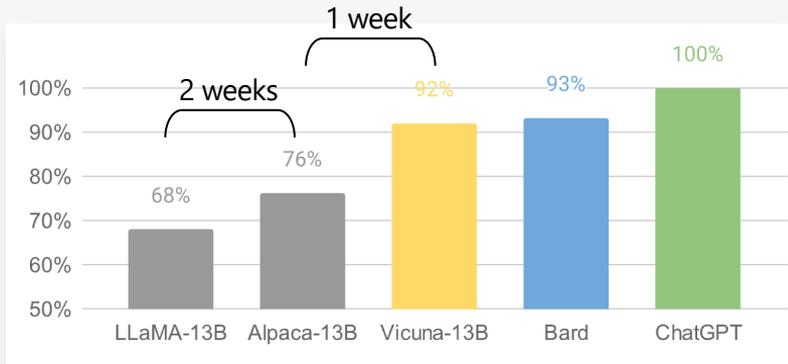


Figure 1. Relative Response Quality Assessed by GPT-4*

The cost of training Vicuna-13B is around \$300.

The code and weights along with an online demo, are publicly available for non-commercial use.“

GPT 3.5 Turbo

0,002 € je 1.000 Tokens (ca. 750 Worte)

GPT 4

0,06 € je 1.000 Tokens

Für 10.000,- € bekomme ich...

ca. 75 GB Textdaten mit GPT 3.5 Turbo

ca. 2,5 GB Textdaten mit GPT 4

Die Wikipedia umfasst derzeit...

ca. 23 GB Textdaten

Ein Bild aus DALL-E2 kostet...

0,02 €



Einsatzgebiete an Hochschulen



© Gerd Altmann auf Pixabay

ONLINE-WORKSHOP

Bewerben mit ChatGPT und Co.: KI für den Berufseinstieg richtig nutzen

Die Erstellung von KI-Tools für das Verfassen von Texten ist derzeit in aller Munde. Tatsächlich können diese Tools auch für das Verfassen und Übersetzen von Bewerbungsdokumenten hilfreich sein, zumal insbesondere das Verfassen von überzeugenden Anschreiben in Deutsch oder Englisch vielen Studierenden schwerfällt.

Career Service

Kontakt

Anschrift:

Career Service
Studierenden-Service-
Center (SSC) 1/106 + 1/113
Universitätsstr. 150
44801 Bochum

Termine nach
Vereinbarung:

E-Mail:
careerservice@rub.de





KEINE KI IN SCHRIFTLICHEN PRÜFUNGEN

Universität Tübingen beschränkt Nutzung von ChatGPT

Teilen:



Die Software ChatGPT kann Mathe, aber auch Texte auf Deutsch und Englisch schreiben. Die Universität Tübingen hat ihren Einsatz für Studierende und Forschende nun stark begrenzt.

Die Universität Tübingen befürchtet, dass die Nutzung der Software ChatGPT überhand nimmt. Deswegen hat sie den Gebrauch der Künstlichen Intelligenz stark eingeschränkt. Im Rahmen schriftlicher Studien- und Prüfungsleistungen



itm

DIDAKTISCHE UND RECHTLICHE PERSPEKTIVEN AUF KI-GESTÜTZTES SCHREIBEN IN DER HOCHSCHULBILDUNG

Peter Salden & Jonas Leschke (Hrsg.)
(Zentrum für Wissenschaftsdidaktik der Ruhr-Universität Bochum)

März 2023



<https://doi.org/10.13154/294-9734>

geleitet durch:

Ministerium für
Hoch- und Wissenschaft
des Landes Nordrhein-Westfalen



<https://doi.org/10.13154/294-9734>

Szenarien an Hochschulen

- Recherche, Zusammenfassungen und Übersetzungen
- Generierung von Texten, Illustrationen, Übersetzungen etc.
- Hilfe bei Gutachten und Bewertungen
- Automatisierung von Aufgaben
- Individualisierte Betreuung / Beratung





⚡ GPT-3.5

🧩 GPT-4 ▾



ChatGPT PLUS

Plan a trip

to see the northern lights in Norway

Recommend activities

for a team-building day with remote e...

Send a message

**E-learning and accessibility: An exploration of the potential role of generic pedagogical tools**Jane Seale^{a,*}, Martyn Cooper^b^aUniversity of Southampton, School of Education, UK^bInstitute of Educational Technology, The Open University, Walton Hall, Milton Keynes, Buckinghamshire, MK7 6AA, UK

ARTICLE INFO

Article history:
 Received 6 May 2009
 Received in revised form 18 October 2009
 Accepted 20 October 2009

Keywords:
 Accessibility
 Post-secondary education
 Pedagogical issues
 Learning theories
 Learning design

ABSTRACT

This paper reviews a range of accessibility tools and evaluates how successful they have been in helping teachers in higher education and further education develop accessible e-learning materials and activities for disabled learners. It is argued that these accessibility specific tools have had limited success to date, and that there may therefore be value in exploring the potential role that more general pedagogic tools might play in the development of accessible e-learning and accessibility practices. Two examples of general pedagogic tools, learning theories and learning design tools, are assessed for the extent to which they raise awareness of an association between accessibility and pedagogy; highlight potential barriers to and facilitators of accessible e-learning and offer methods and approaches for developing accessible e-learning. This assessment suggests that the general pedagogical tools will have a limited or abstract influence on the development of accessible e-learning, but that this influence could be potentially increased by “blending” the use of specific accessibility tools with the use of more general pedagogical tools. The success of a blended approach to designing and developing accessible e-learning will depend on a number of factors, including teacher and learner agency. Nevertheless such an approach could be significant in terms of seeking to make accessibility as much a pedagogical issue as it is a technical one.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Broadly speaking, accessibility in relation to e-learning (e.g. virtual learning environments, digital repositories, multimedia, web portals and discussion boards) is understood as ensuring that learners are not prevented from accessing technologies or the content and experience offered by technologies on the grounds of their disability. There have been many generic definitions of accessibility that have focused on reducing barriers to accessing the Web and ensuring equitable access for all users (e.g. Paciello, 2000; Web Accessibility Initiative, 2005) The IMS Global Learning Consortium offers a more education specific definition of both disability and accessibility:

[...] the term disability has been re-defined as a mismatch between the needs of the learner and the education offered. It is therefore not a personal trait but an artifact of the relationship between the learner and the learning environment or education delivery. Accessibility, given this re-definition, is the ability of the learning environment to adjust to the needs of all learners. Accessibility is determined by the flexibility of the education environment (with respect to presentation, control methods, access modality, and learner supports) and the availability of adequate alternative-but-equivalent content and activities. The needs and preferences of a user may arise from the context or environment the user is in, the tools available (e.g., mobile devices, assistive technologies such as Braille devices, voice recognition systems, or alternative keyboards, etc.), their background, or a disability in the traditional sense. Accessible systems adjust the user interface of the learning environment, locate needed resources and adjust the properties of the resources to match the needs and preferences of the user (IMS Global Learning Consortium, 2004).

Central to this definition are the concepts of adaptation and flexibility and the idea that learning environments can and should be adjusted to meet the needs of learners. An example of an e-learning resource that has been designed to be accessible and therefore reflects these principles of adaptation and flexibility is the LEXDIS (disabled learners experiences of e-learning) online database [1]. This online resource was designed by researchers and disabled students, for teachers and disabled students. The resource aims to provide information and advice on how disabled students can use technologies effectively to support their learning and on how teachers can support learners in

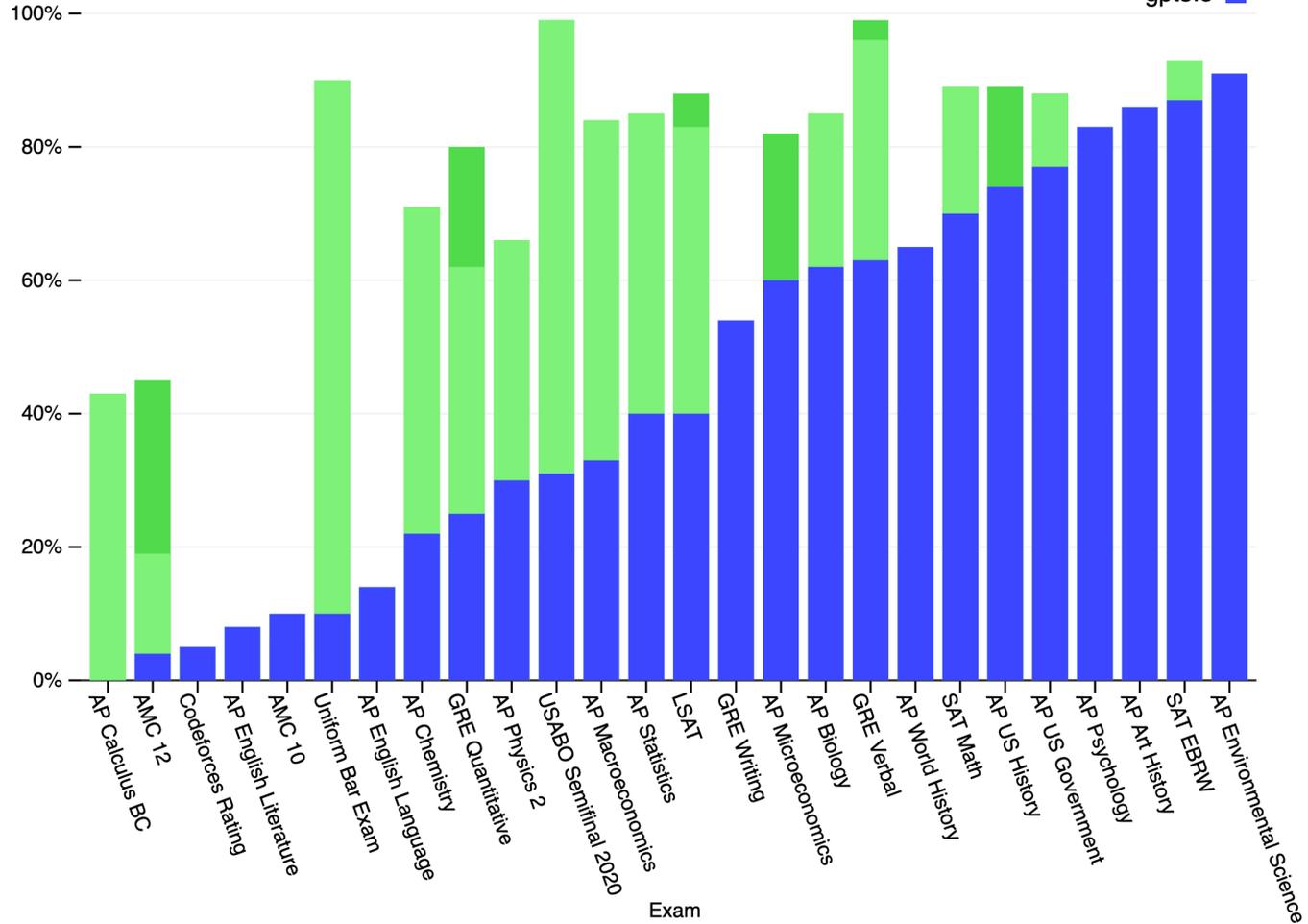
* Corresponding author. Address: School of Education, University of Southampton, Highfield, Southampton SO17 1BJ, UK. Tel.: +44 2380 594784.
 E-mail address: J.K.Seale@oton.ac.uk (J. Seale).

GPT-4

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)

gpt-4
gpt-4 (no vision)
gpt3.5



GPT-4 can solve difficult problems with greater accuracy, thanks to its broader general knowledge and problem solving abilities.

Creativity | Visual input | Longer context

GPT-4 can accept images as inputs and generate captions, classifications, and analyses.

Output

There are many options for what you can make with these ingredients! Some possibilities include:

- Pancakes or waffles
- Crepes
- French toast
- Omelette or frittata
- Quiche
- Custard or pudding
- Cake or cupcakes
- Muffins or bread- Cookies or biscuits

These are just a few examples, but the possibilities are endless!

Input

What can I make with these ingredients?

User: What is funny about this image? Describe it panel by panel.



Source: [hmmm \(Reddit\)](#)

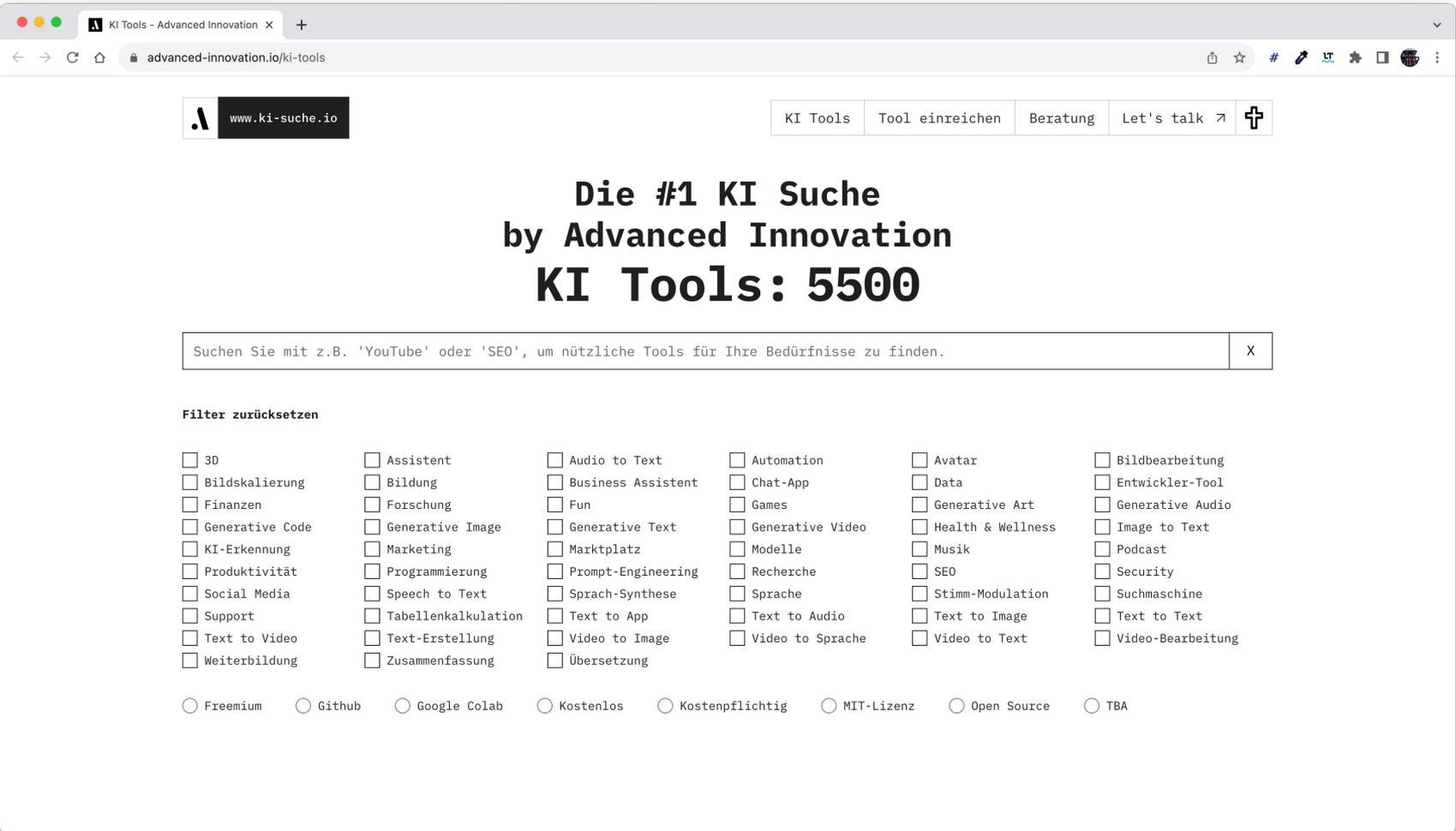
GPT-4: The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.



Die #1 KI Suche by Advanced Innovation KI Tools: 5500

Suchen Sie mit z.B. 'YouTube' oder 'SEO', um nützliche Tools für Ihre Bedürfnisse zu finden. X

Filter zurücksetzen

- 3D
- Assistent
- Audio to Text
- Automation
- Avatar
- Bildbearbeitung
- Bildskalierung
- Bildung
- Business Assistent
- Chat-App
- Data
- Entwickler-Tool
- Finanzen
- Forschung
- Fun
- Games
- Generative Art
- Generative Audio
- Generative Code
- Generative Image
- Generative Text
- Generative Video
- Health & Wellness
- Image to Text
- KI-Erkennung
- Marketing
- Marktplatz
- Modelle
- Musik
- Podcast
- Produktivität
- Programmierung
- Prompt-Engineering
- Recherche
- SEO
- Security
- Social Media
- Speech to Text
- Sprach-Synthese
- Sprache
- Stimm-Modulation
- Suchmaschine
- Support
- Tabellenkalkulation
- Text to App
- Text to Audio
- Text to Image
- Text to Text
- Text to Video
- Text-Erstellung
- Video to Image
- Video to Sprache
- Video to Text
- Video-Bearbeitung
- Weiterbildung
- Zusammenfassung
- Übersetzung

- Freemium
- Github
- Google Colab
- Kostenlos
- Kostenpflichtig
- MIT-Lizenz
- Open Source
- TBA



modern hotel, editorial style photo, front view,
contemporary, The hotel pool area, Stone, Glass,
Wood, Luxurious Finish, Neutral, Hotel Brand, Natural
Light with Warm Tones, Seaside Resort, Beachfront,
Sunset, Relaxing, Modernism, Minimalist -ar 16:9



Colossyan 

What's new



4 New Avatars
Shoulder view
Image cropping
...and more!



ABBT Institute of
Biotechnology

RWTHAACHEN
UNIVERSITY



Öcher
Streuselbrötchen

Prime Voice AI

The most realistic Text to Speech and Voice Cloning software. ElevenLabs brings the most compelling, rich and lifelike voices to creators and publishers seeking the ultimate tools for storytelling.

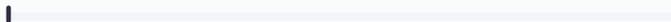
Click on a language to generate in a random text: English German Polish Spanish Italian French Portuguese

Hindi

Eleven lets you voice any length of text in top quality, all while automatically matching what is being said with how it's being said. The model works best on longer texts, so type in at least a few sentences.

— premade/Adam ▾

0 / 333



Get Started Free →

Grow Your Audience by Expanding into Audio

Generate top-quality spoken audio in any voice, style and language with the most advanced Text to Speech tool ever. Our AI model renders human intonation and inflections with unprecedented fidelity and it adjusts delivery based on context.



Copilot



Rechtsfragen

Weitere Rechtsgebiete

- Urheberrecht und Verwertungsrechte
- Persönlichkeitsrechte
- Datenschutzrecht
- Prüfungsrecht
- Arbeits- und Dienstrecht
- Strafrecht



Prüfungsrecht

Urheberrecht und Verwertungsrechte

Welche Prüfungsleistungen können durch ein KI-System maßgebliche Beiträge liefern?
Inwiefern berühren die für das Training von ChatGPT

Persönlichkeitsrechte

Wie kann die Persönlichkeitsrechte der Eigenständigkeit der Verfasser durch das Urheberrecht von Dritten?

Datenschutzrecht

Wie wird das Datenschutzrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Arbeits- und Dienstrecht

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?

Wie wird das Arbeits- und Dienstrecht gewährleistet?
Gibt es Möglichkeiten die unzulässige Nutzung von KI-Systemen zu vermeiden?



Rechtsfragen beim Umgang
mit KI-Systemen
Stand: 2023-09-16

RWTH Aachen University
Center für Lehr- und Lernservices

70% Of Workers Using ChatGPT At Work Are Not Telling Their Boss; Overall Usage Among Professionals Jumps To 43%

February 1, 2023

fishbowl insights

IF YOU USE CHATGPT, OR OTHER AI TOOLS, FOR WORK RELATED TASKS, DOES YOUR BOSS KNOW?



* Responses from Professionals who use ChatGPT at Work

43% der Fachleute verwenden KI-Tools, einschließlich ChatGPT, für arbeitsbezogene Aufgaben. Fast 70% dieser Fachleute tun dies ohne das Wissen ihres Vorgesetzten.

Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence

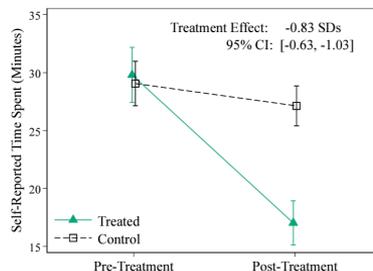
Shakked Noy
MIT

Whitney Zhang
MIT

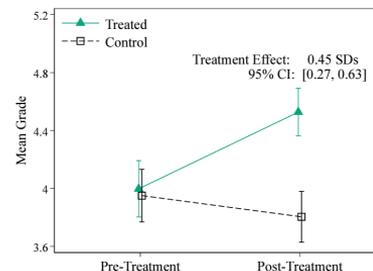
March 2, 2023

Working Paper (not peer reviewed)

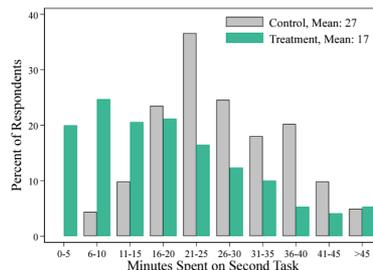
(a) Time Taken Decreases



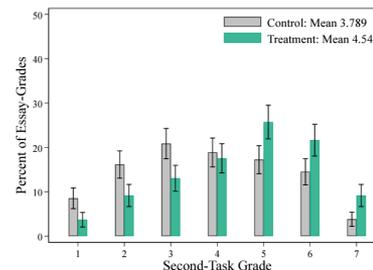
(b) Average Grades Increase



(c) Time Distribution (Second Task)



(d) Grades (Second Task)



[Home](#) / [Nachrichten](#) / [Digitales](#) / EU-Kommission veröffentlicht interne Leitlinien zu ChatGPT

EU-Kommission veröffentlicht interne Leitlinien zu ChatGPT

Von: [Luca Bertuzzi](#) | [EURACTIV.com](#) | übersetzt von [Luka Krauss](#) 🗓️ 31. Mai 2023

Die Mitarbeitenden werden aufgefordert, stets kritisch zu prüfen, ob der von der KI erzeugte Output die Rechte des geistigen Eigentums und Reaktionszeit oder können nicht immer insbesondere das Urheberrecht verletzt, und verfügbar sein. Daher ist es den Mitarbeitern den Output eines generativen KI-Modells der Kommission untersagt, sich bei kritischen niemals direkt in öffentlichen Dokumenten zu und zeitkritischen Aufgaben auf diese reproduzieren, wie etwa bei der Erstellung von Werkzeuge zu verlassen. Kommissionstexten, insbesondere von rechtsverbindlichen Texten.“

Das Dokument "Guidelines for staff on the use of online available generative Artificial Intelligences tools" und die dazugehörige

Unterstützer



AI4TRUST

Quelle: <https://www.euractiv.de/section/innovation/news/eu-kommission-veroeffentlicht-interne-leitlinien-zu-chatgpt/>

Die RWTH Aachen versteht generative KI proaktiv als
Werkzeug für Forschung, Lehre und Verwaltung

Rechtsfragen sind von organisatorischen, technologischen,
didaktischen und explorativen Fragen zu trennen.

Werkzeuge aus dem Bereich der Generativen KI
sind für die Mitarbeitenden bereitzustellen.