

Grundlagen und Einsatzmöglichkeiten von ChatGPT in der Praxis

Leon Kaltenbrunn, Philipp Müller

Referent

- Studium der Informatik – Albert-Ludwigs-Universität Freiburg
- Abschlussarbeit im Bereich Künstliche Intelligenz
- Software Engineer bei Haufe
- Arbeitet zur Zeit am „HR Assistant“ einer KI Chatbot-Lösung für HR-Anwendungen der Haufe Group



Leon Kaltenbrunn

Software Engineer

Haufe Content and Solutions

Referent

- Studium der Informatik – Albert-Ludwigs-Universität Freiburg
- Schwerpunkte:
 - Demokratisierung von KI-Methoden
 - Automatisiertes Optimieren von KI Anwendungen mittels statistischer Modelle
- Forschung an der LMU München
- Machine Learning Engineer bei Haufe



Philipp Müller

ML Engineer and Data Scientist

Haufe Content and Solutions

Agenda

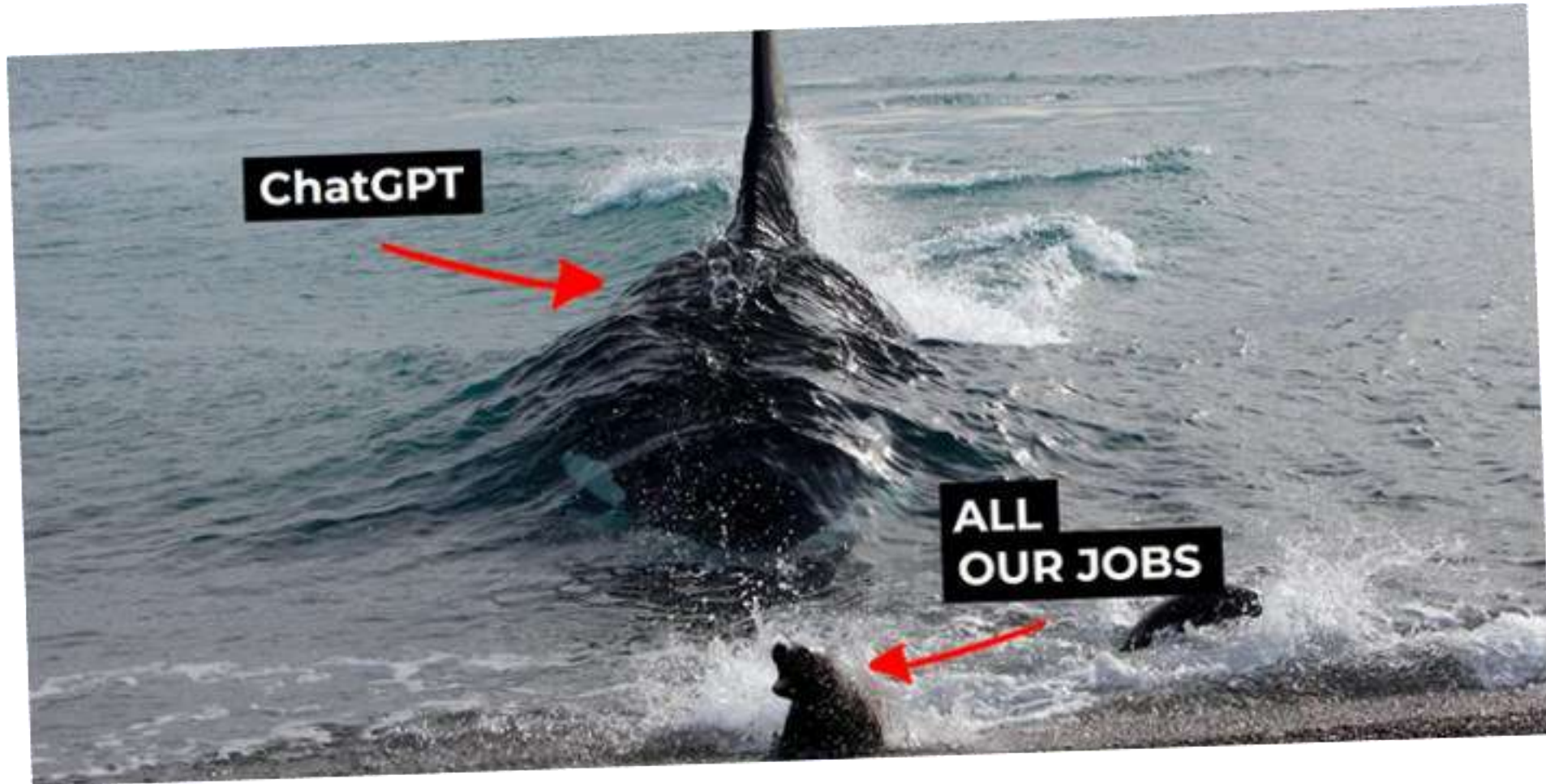
1. Ein kurzer Überblick
2. Was ist ChatGPT und wie funktioniert es?
3. Wer ist OpenAI?
4. Wie kann ich ChatGPT in der Praxis anwenden?
5. Was sind die Limitierungen und Risiken?
6. Überblick und Ausblick in die Zukunft

„Die Haufe Group verpflichtet sich zu einer neutralen Berichterstattung und Kommunikation. Aus Gründen der besseren Lesbarkeit verwenden wir bei Personenbezeichnungen und personenbezogenen Hauptwörtern die männliche Form. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung grundsätzlich für alle Geschlechter. Die verkürzte Sprachform hat nur redaktionelle Gründe und beinhaltet keine Wertung.“

Umfrage

Während des Online-Seminars werden einige Blitzumfragen durchgeführt. Dazu wird Ihnen jeweils eine Frage mit mehreren Antwortmöglichkeiten eingeblendet. Um an der Umfrage teilzunehmen, müssen Sie nichts weiter tun, als eine Antwort anzuklicken. Die Beantwortung erfolgt selbstverständlich anonym.

Standardreflex bei neuen, disruptiven Technologien



Mazen Kourouche, creative director (Litmus)

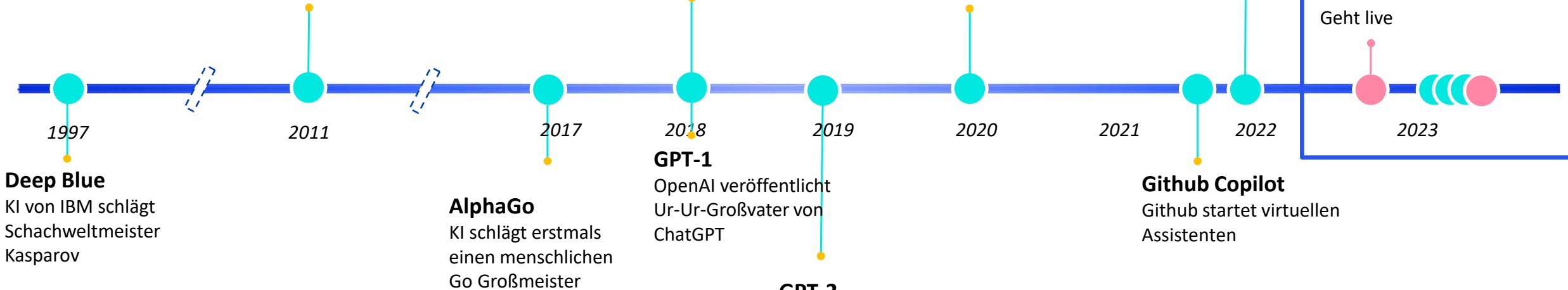
„ AI is not going **to replace** our jobs
– someone who knows how to use it will.“

EIN KURZER ÜBERBLICK

Überblick



Watson
KI von IBM schlägt bisherige Jeopardy Champions



Deep Blue
KI von IBM schlägt Schachweltmeister Kasparov



AlphaGo
KI schlägt erstmals einen menschlichen Go Großmeister



BERT
Google veröffentlicht Transformer Modell mit 370M Parametern

GPT-1
OpenAI veröffentlicht Ur-Ur-Großvater von ChatGPT

GPT-2
Neues Modell mit 1.500M Parametern

GPT-3
Neues Modell mit 175.000M Parametern

DALL-E 2
OpenAI veröffentlicht Bildgenerator

Github Copilot
Github startet virtuellen Assistenten

ChatGPT
Geht live



ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible to ask ChatGPT to answer followup questions, admit its mistakes

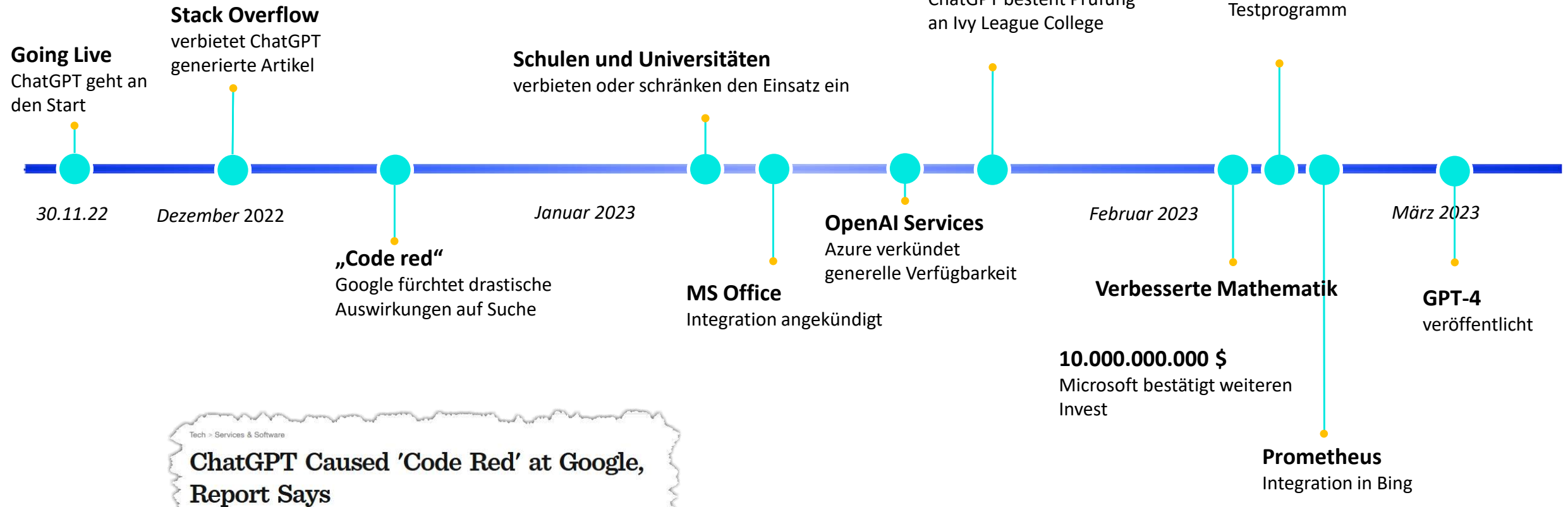
ChatGPT Gets an MBA

The AI-powered chatbot did better than expected on something to get excited about, says the professor

Überblick

Kein Lernerfolg mehr? Erste Schulen und Universitäten verbieten ChatGPT

Präsident der International Conference on Machine Learning



Tech > Services & Software

ChatGPT Caused 'Code Red' at Google, Report Says

Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web

Feb 7, 2023 | Yusuf Mehdi, Corporate Vice President & Consumer Chief Marketing Officer

WAS IST CHATGPT?

Model

Default (GPT-3.5)



ChatGPT **PLUS**



Eingabe = Prompt

PH

Schreibe ein kurzes Gedicht



Die Sonne strahlt so hell und klar,
ein Lächeln legt sich auf mein Gesicht.
Der Frühling zeigt sich wunderbar,
die Welt erwacht aus ihrem Winterlicht.



Die Blumen blühen farbenfroh,
die Vögel singen ihr Lied.
Ich spüre wie ich glücklich werde,
in dieser wundervollen Zeit.

Lasst uns die Natur genießen,
die Schönheit um uns herum.
Wir dürfen das Leben genießen,
denn es ist ein kostbares Gut.

KI generierte Ausgabe

🔄 Regenerate response

|  

Für was steht ChatGPT ?

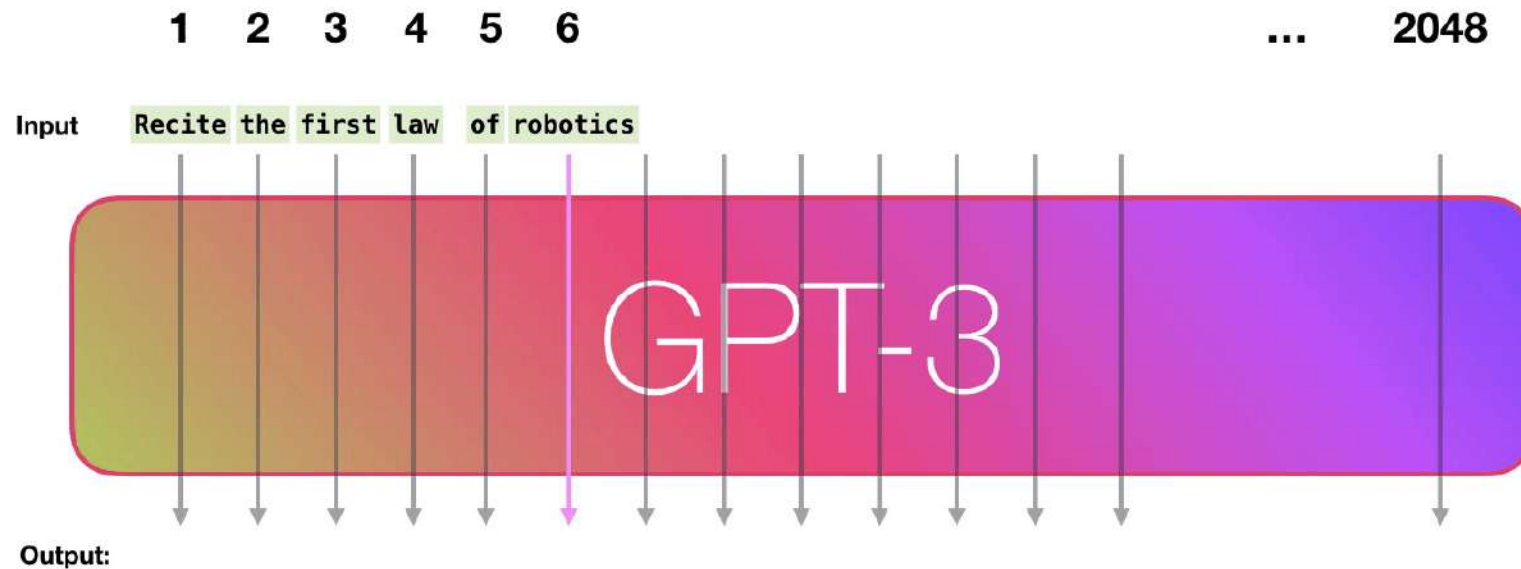
GENERATIVE PRETRAINED TRANSFORMER

Was bedeutet **Erzeugend** ?

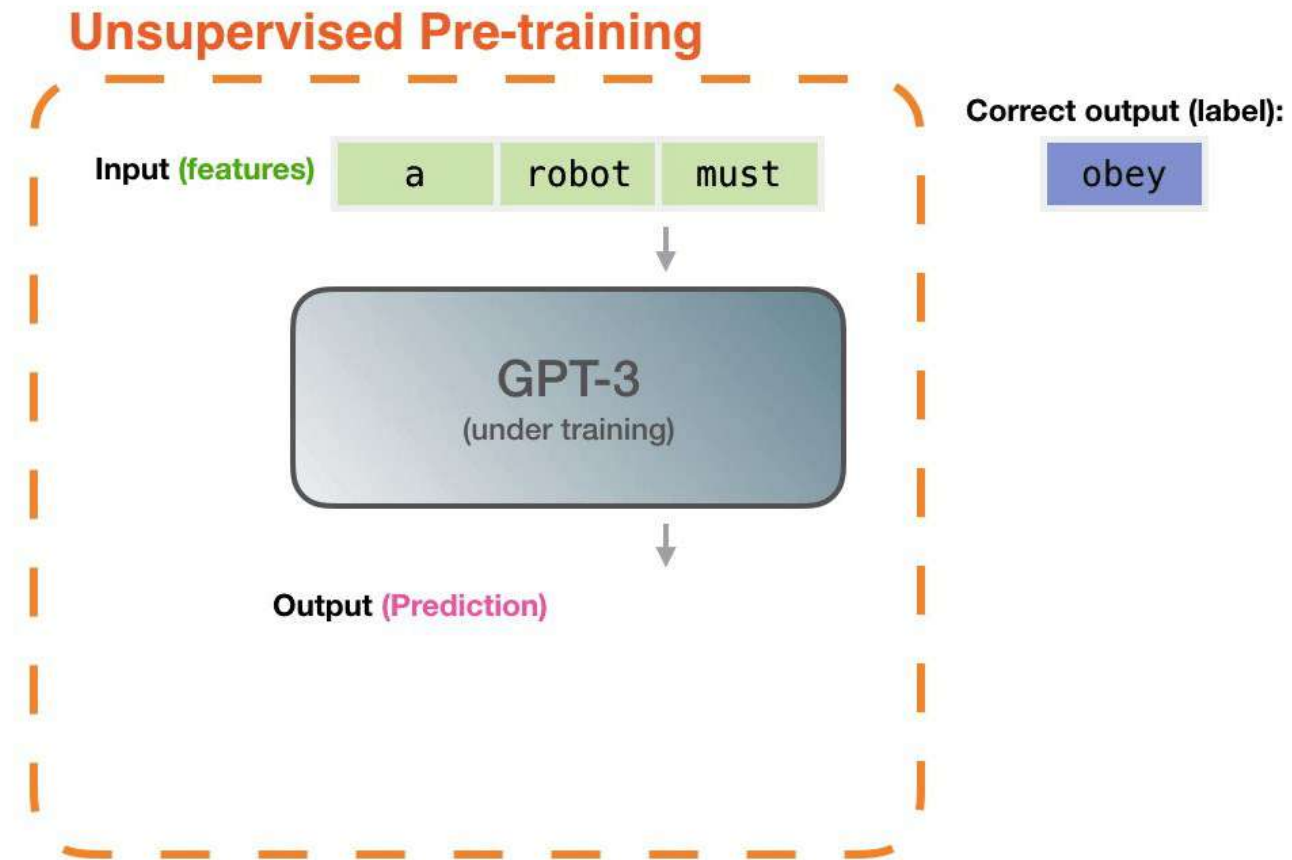
Auf was und wie
wurde **trainiert** ?

Was ist ein **Transformer**?

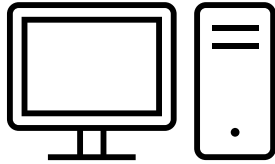
Was bedeutet Erzeugend ?



Was bedeutet Erzeugend ?



Auf was wurde trainiert ?



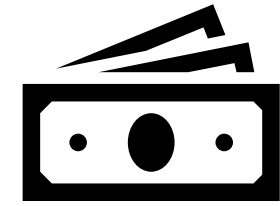
Hardware:

- 285,000 CPU cores
- 10,000 GPUs
- 400 gigabits/s network traffic



Daten:

- GPT-3.5 Turbo Data Cut-off: September 2021
- 499 Billion Tokens
≈ 700 GB Text
(45 Terabyte ungefiltert)

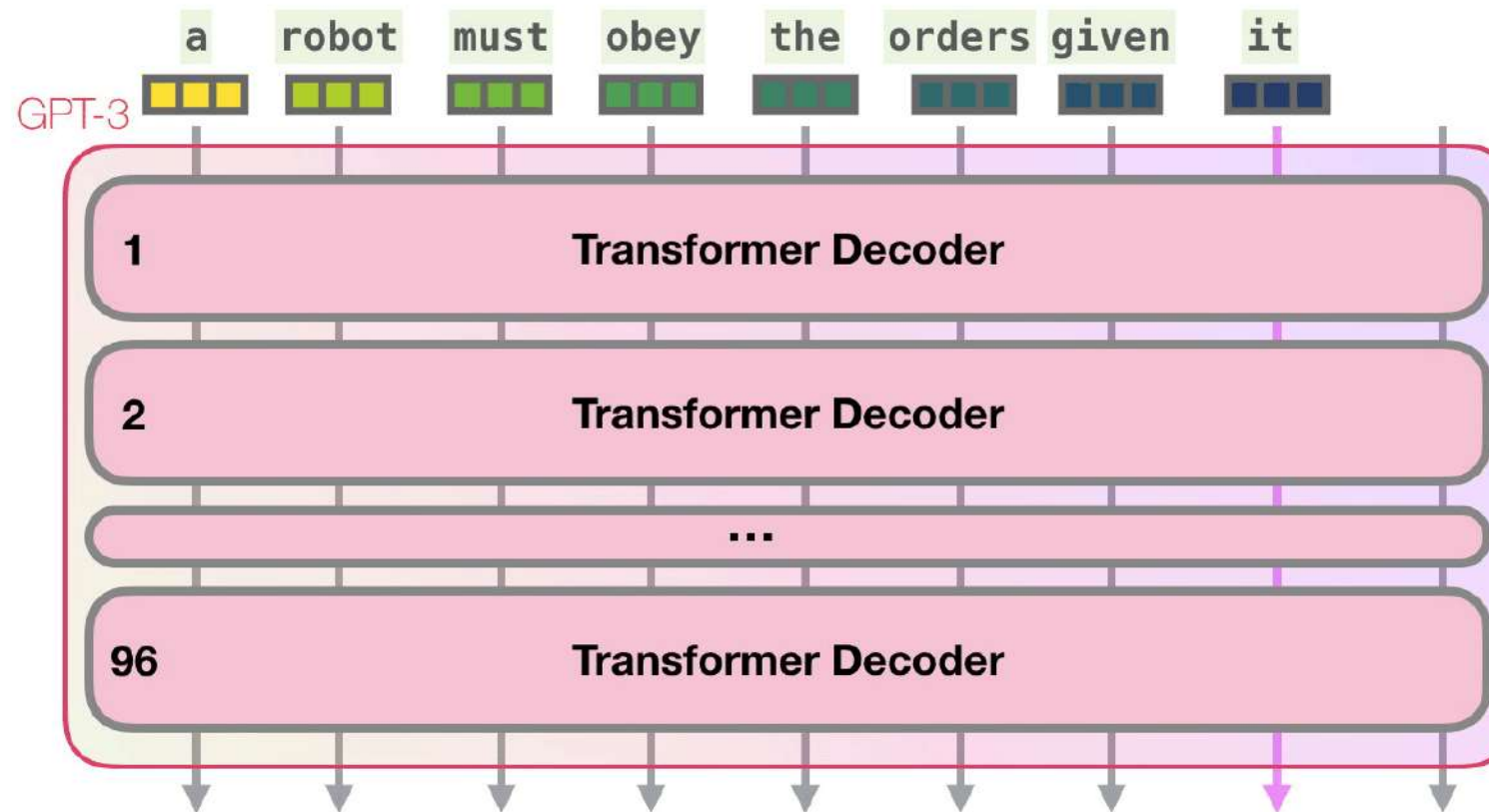


Kosten:

- GPT-3 circa 4.6 Mio Euro
- Nicht veröffentlicht für GPT-4

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

Und was ist ein Transformer?



Und was ist ein Transformer?

5.03762v5 [cs.CL] 6 Dec 2017

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin* †
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task,

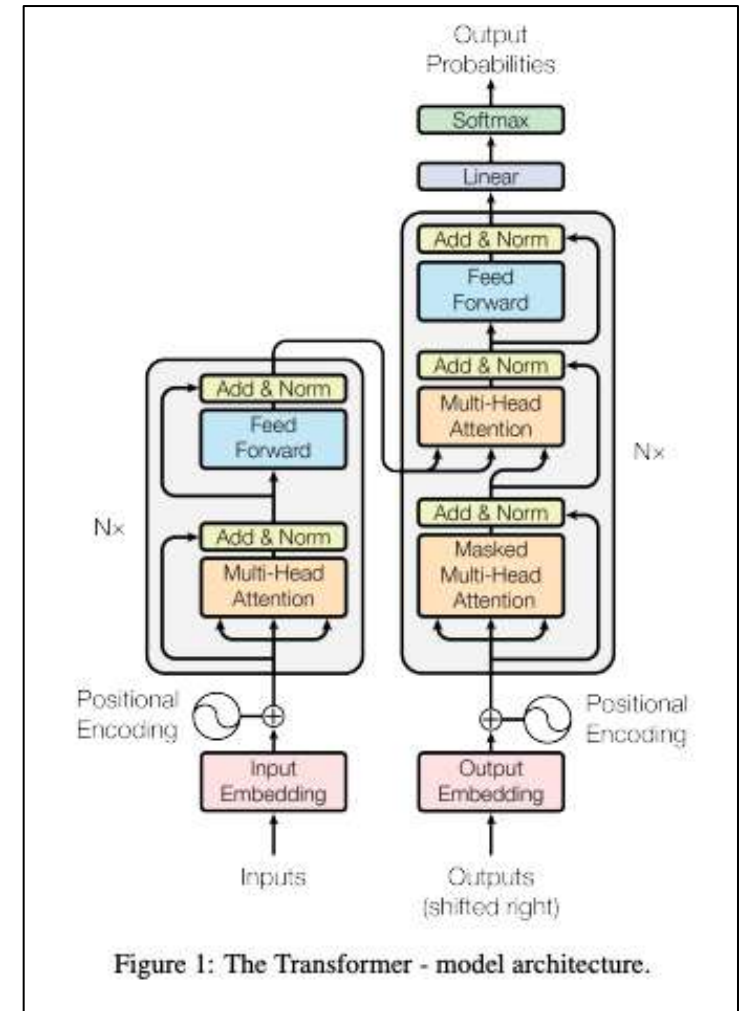
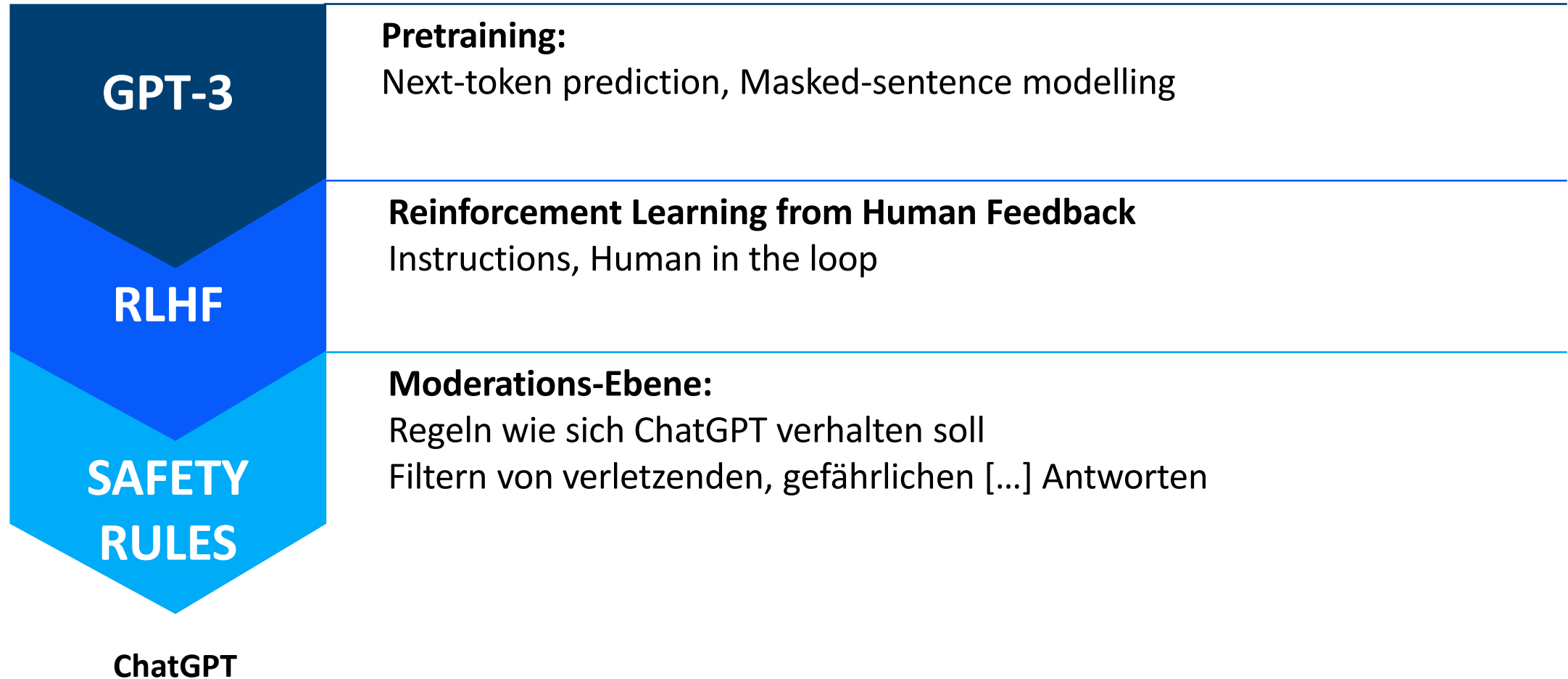


Figure 1: The Transformer - model architecture.

Von GPT zu ChatGPT



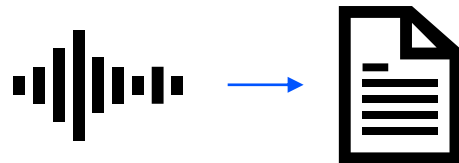
WER IST OPENAI?

Wer ist OpenAI?

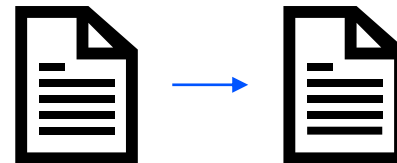


- Gegründet 2015 in San Francisco
- Gründungsmitglieder: Sam Altman, Elon Musk, ...
- Investoren: Peter Thiel, Reid Hoffmann, Microsoft, ...
- Gründungsvision: Fortschritt in KI Forschung für alle nutzbar machen
- Mögliche boshafte Nutzung: Ab GPT-3 (2020) Code + Daten nicht frei verfügbar
- 2019: Microsoft wird Großinvestor
 - 1 Mrd US\$ von Microsoft
 - Geschäftsmodell von „Non-Profit“ zu „Capped-Profit“
 - Weitere 10 Mrd US\$ in den nächsten Jahren (= 49% Anteil durch Microsoft)

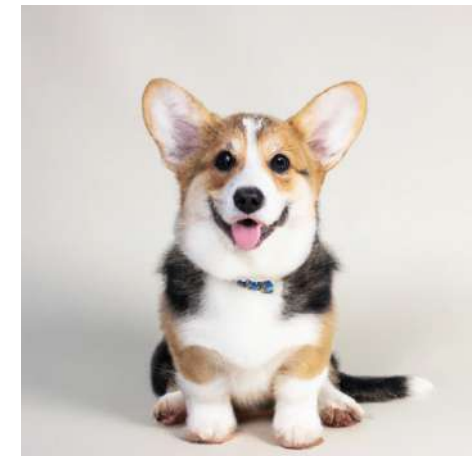
Weitere OpenAI Produkte?



Transkription
Übersetzen



Premium Subscription

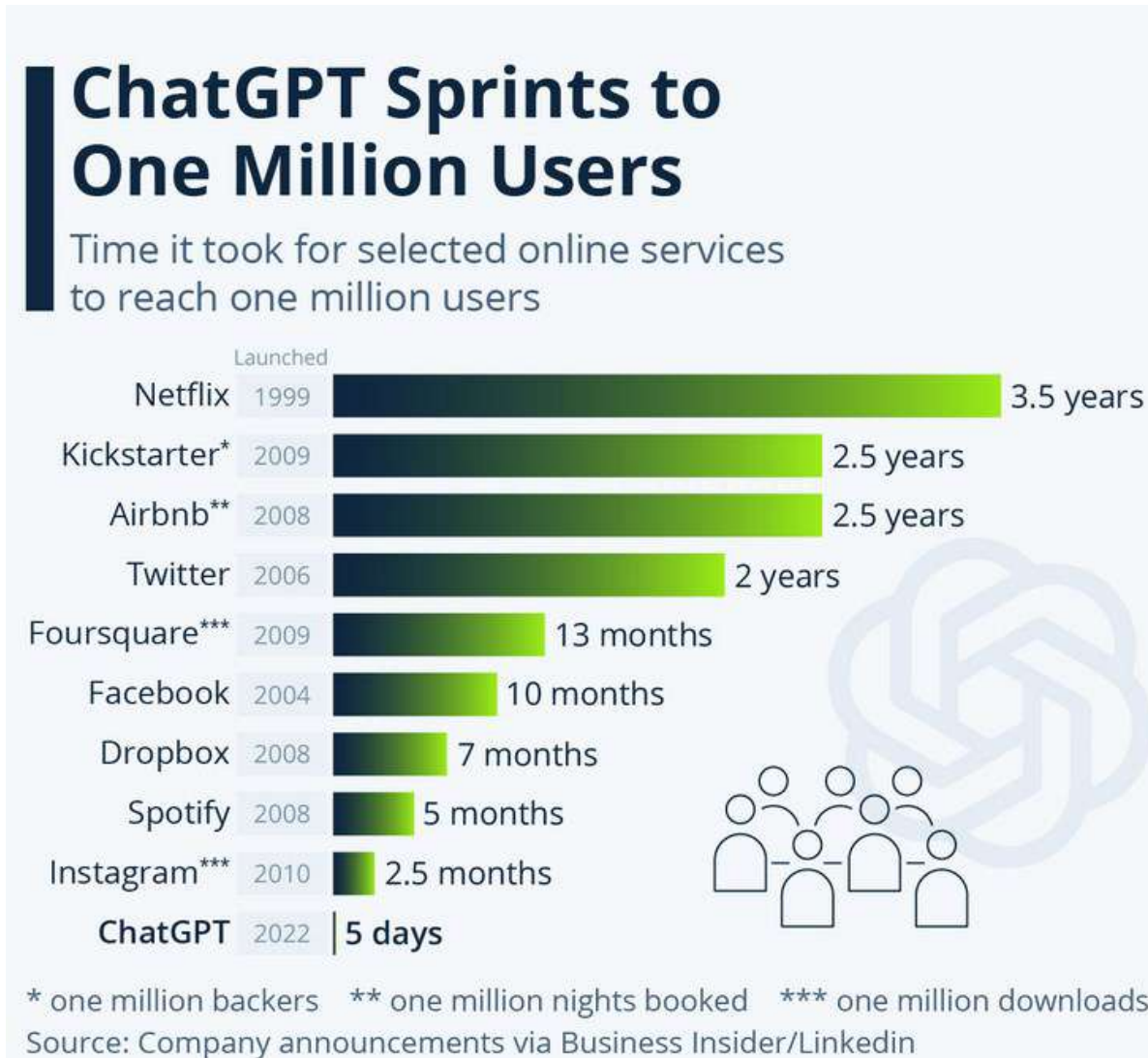


„a photo of a happy corgi puppy sitting and facing forward, studio light, longshot“

OpenAI ist eine Erfolgsgeschichte

Wie viele *Monate* hat ChatGPT benötigt,
um **1 Millionen aktive Nutzer** zu erreichen?

OpenAI ist eine Erfolgsgeschichte



OpenAI ist eine Erfolgsgeschichte

Wie viele *Monate* hat ChatGPT benötigt,
um **100 Millionen aktive Nutzer** zu erreichen?

OpenAI ist eine Erfolgsgeschichte



PRAKTISCHE ANWENDUNGEN

Model

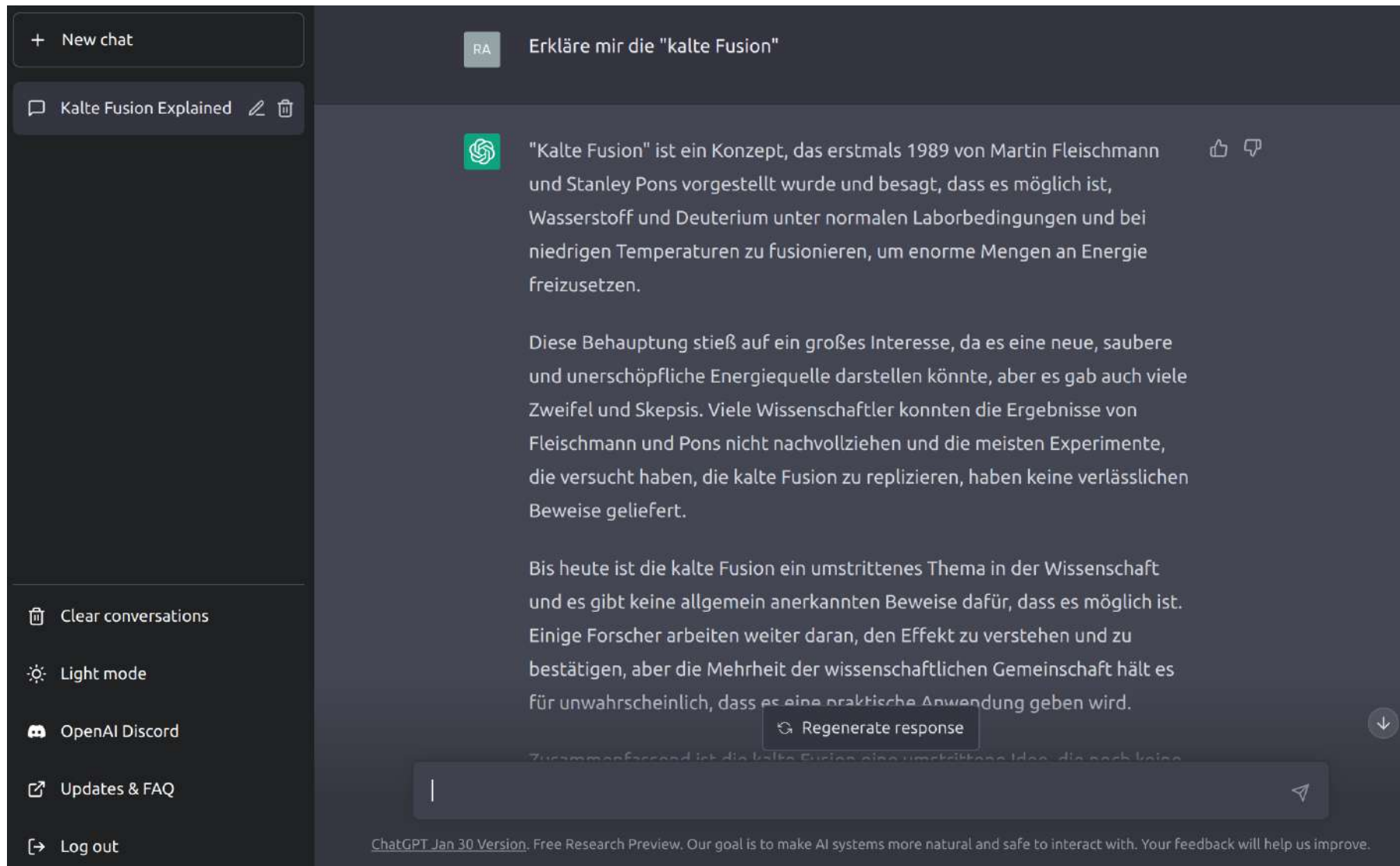
Default (GPT-3.5)



ChatGPT **PLUS**



Informationen bekommen



The screenshot shows the ChatGPT web interface. On the left is a sidebar with navigation options: '+ New chat', 'Kalte Fusion Explained' (with edit and delete icons), 'Clear conversations', 'Light mode', 'OpenAI Discord', 'Updates & FAQ', and 'Log out'. The main chat area has a title 'Erkläre mir die "kalte Fusion"' and a user icon 'RA'. The AI response, marked with the OpenAI logo, explains that 'cold fusion' was first proposed in 1989 by Martin Fleischmann and Stanley Pons, claiming it could produce energy under normal conditions. It notes that while there was initial interest, most experiments failed to be replicated, and the scientific community remains skeptical. A 'Regenerate response' button is visible below the text. At the bottom, a footer reads: 'ChatGPT Jan 30 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.'

Informationen modifizieren

The screenshot shows the ChatGPT web interface. On the left is a dark sidebar with a 'New chat' button and a list of previous conversations, including 'Kalte Fusion Explained'. The main area displays a chat history where a user has asked for a shorter version of a previous response. The AI's response explains that cold fusion was proposed in 1989 but lacks widely accepted scientific evidence. At the bottom, there is a 'Regenerate response' button and a text input field.

+ New chat

Kalte Fusion Explained

für unwahrscheinlich, dass es eine praktische Anwendung geben wird.

Zusammenfassend ist die kalte Fusion eine umstrittene Idee, die noch keine breite Akzeptanz in der Wissenschaft hat, und weitere Forschung ist erforderlich, um ihre Validität zu bestätigen oder zu widerlegen.

RA Bitte kürzer in drei Sätzen

Die kalte Fusion wurde 1989 als Konzept vorgestellt, das besagt, dass man Wasserstoff und Deuterium unter normalen Laborbedingungen fusionieren kann, um Energie freizusetzen. Es ist ein umstrittenes Thema in der Wissenschaft, da die meisten Experimente, die versucht haben, den Effekt zu replizieren, keine verlässlichen Beweise geliefert haben. Bis heute gibt es keine allgemein anerkannten Beweise für die kalte Fusion und weitere Forschung ist erforderlich, um ihre Validität zu bestätigen oder zu widerlegen.

Regenerate response

ChatGPT Jan 30 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

Texte korrigieren



Bitte korrigiere die Rechtschreibung, Zeichensetzung und Grammatik in folgendem Text:

Das Upgrade von Software-Abhängigkeiten auf neue Minor oder gar **neue** Major-Versionen kann erheblichen Aufwand verursachen. Denn einerseits besteht immer die Gefahr, dass bewusst oder versehentlich Schnittstellen in **inkomaptibler** Weise geändert worden sind um z.B. Sicherheitslücken zu stopfen. Andererseits ist auch je nach **verwendetn** Frameworks und Paketmanagern die Wahrscheinlichkeit hoch, dass man sich in der "Abhängigkeitshölle" wiederfindet, weil es widersprüchliche Versionseinschränkungen bei transitiven Abhängigkeiten gibt.



Bitte korrigiere die Rechtschreibung, Zeichensetzung und Grammatik in folgendem Text:

Das Upgrade von Software-Abhängigkeiten auf neue Minor- oder gar Major-Versionen kann erheblichen Aufwand verursachen. Denn einerseits besteht immer die Gefahr, dass **Schnittstellen bewusst oder versehentlich** auf **inkompatible** Weise geändert wurden, z. B. um Sicherheitslücken zu stopfen. Andererseits ist auch je nach **verwendetem** Framework und Paketmanager die Wahrscheinlichkeit hoch, dass man sich in der "Abhängigkeitshölle" wiederfindet, weil es widersprüchliche Versionseinschränkungen bei transitiven Abhängigkeiten gibt.

(Markierungen ergänzt)

Texte zusammenfassen

Bitte fasse folgenden Text zusammen:

Die Verwertungsgesellschaft Musikedition ist als Verwertungsgesellschaft ein wirtschaftlicher Verein kraft staatlicher Verleihung. Wie alle deutschen Verwertungsgesellschaften untersteht sie der Aufsicht und der Kontrolle des Deutschen Patent- und Markenamtes, des Bundesjustizministeriums sowie des Bundeskartellamtes. Die VG Musikedition macht keine eigenen Gewinne; vielmehr verteilt sie nach Abzug ihrer Verwaltungskosten (im Geschäftsjahr 2013 6,6 %[1]) sämtliche Einnahmen an ihre Mitglieder.

Die 1886 (zum 31. Dezember 2017[2]) ordentlichen und angeschlossenen Mitglieder der VG Musikedition sind Verlage, Komponisten, Texter und wissenschaftliche Herausgeber. Die VG Musikedition selbst wiederum ist Gesellschafter der 'Zentralstelle Fotokopieren an Schulen' (ZFS) und der Zentralstelle Bibliothekstantieme (ZBT).

Die Geschäftsstelle der VG Musikedition befindet sich in Kassel. Geschäftsführer ist seit 2002 Christian Krauß, Präsident ist seit 2018 Sebastian Mohr (Breitkopf & Härtel, Wiesbaden). Mitglieder des Verwaltungsrates als Aufsichtsgremium sind Thomas Sertl (Schott Music, Mainz), Friedemann M. Strube (Strube, München), der Liedermacher Wolfgang Hering sowie die Musikwissenschaftlerin Gabriele Buschmeier (Akademie der Wissenschaften und der Literatur, Mainz).

https://de.wikipedia.org/wiki/Verwertungsgesellschaft_Musikedition



Die VG Musikedition ist eine staatlich anerkannte Verwertungsgesellschaft in Deutschland, die unter der Aufsicht von Regierungsbehörden steht. Sie besteht aus Verlagen, Komponisten, Textern und wissenschaftlichen Herausgebern und verteilt nach Abzug ihrer Verwaltungskosten alle Einnahmen an ihre Mitglieder. Die Geschäftsstelle befindet sich in Kassel und der Geschäftsführer ist Christian Krauß seit 2002. Der Verwaltungsrat besteht aus Mitgliedern der Musikbranche.

Texte vereinfachen



Kannst Du bitte den folgenden Absatz für Laien verständlich formulieren?

Eine gewerbesteuerrechtliche Hinzurechnung von Mieten für Mehrwegbehältnisse scheidet aus, wenn das Vertragsverhältnis neben der Gebrauchsüberlassung auch umfangreiche Werk-, Dienstleistungs- und Transportvertragsselemente enthält und das Mietvertragsselement dem gesamtvertraglichen Leistungsbündel nicht das Gepräge gibt.

(aus BFH Urteil III R 56/20 vom 01.06.2022)



Wenn ein Vertrag nicht nur die Überlassung von Gegenständen beinhaltet, sondern auch umfangreiche Arbeiten und Dienstleistungen sowie den Transport, dann kann das Mieten von Mehrwegbehältern nicht als Gewerbesteuer hinzugerechnet werden, wenn es nicht das wichtigste Element des gesamten Vertrags ist.



Wichtig

Prüfen Sie den generierten Output stets gründlich!

Generierte Texte können Fehler enthalten.

WEITERE ANWENDUNGEN

Erweiterung für Office Produkte

- Microsoft plant Co-Pilot
- Assistent in allen Office Produkten
- Reichhaltige Einsatzmöglichkeiten
- Erhöhung der Produktivität, sodass man sich auf die wichtigen Dinge konzentrieren kann

 OpenAI

+

 Office



 Copilot

Q1 Performance • Saved

Search

File Home Insert Formulas Data Review View Automate Help

Comments Catch Up Editing Share

Aptos (Body) 11 B General

A1 Country

	A	B	C	D	E	F	G	H	I	J	K
1	Country	Customer	Product	Discount Band	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS
2	France	VanArsdel	Biodegradable Cleaning Products	Low	4345	\$10	\$352	\$1,529,440	\$15,294	\$1,514,146	\$
3	USA	VanArsdel	Biodegradable Cleaning Products	Low	4245	\$10	\$352	\$1,494,240	\$14,942	\$1,479,298	\$
4	Australia	VanArsdel	Biodegradable Cleaning Products	None	4124	\$10	\$352	\$1,451,648	\$0	\$1,451,648	\$
5	Mexico	VanArsdel	Biodegradable Cleaning Products	Medium	3954	\$10	\$352	\$1,391,808	\$83,508	\$1,308,300	\$
6	France	Contoso Pharmaceuticals	Biodegradable Cleaning Products	Low	4952	\$10	\$127	\$628,904	\$6,289	\$622,615	\$
7	France	Contoso Pharmaceuticals	Recycled Can	Low	4903	\$5	\$127	\$622,681	\$6,227	\$616,454	\$20
8	Canada	Proseware	Natural Personal Care Products	Low	1954	\$270	\$302	\$590,108	\$5,901	\$584,207	\$58
9	Canada	Proseware	Biodegradable Cleaning Products	High	1990	\$220	\$302	\$600,980	\$84,137	\$516,843	\$48
10	USA	VanArsdel	Compostable Food Packaging	Low	1254	\$3	\$380	\$476,520	\$4,765	\$471,755	\$
11	USA	Proseware	Natural Personal Care Products	Medium	1465	\$224	\$302	\$442,430	\$26,546	\$415,884	\$39
12	USA	Proseware	Sustainable Snack Packaging	Medium	1410	\$120	\$302	\$425,820	\$25,549	\$400,271	\$19
13	Australia	Proseware	Natural Personal Care Products	Medium	1369	\$224	\$302	\$413,438	\$24,806	\$388,632	\$33
14	USA	VanArsdel	Compostable Food Packaging	Medium	1052	\$3	\$380	\$399,760	\$23,986	\$375,774	\$
15	Canada	VanArsdel	Compostable Food Packaging	Low	980	\$312	\$380	\$372,400	\$3,724	\$368,676	\$33
16	Australia	Proseware	Sustainable Snack Packaging	Medium	1298	\$120	\$302	\$391,996	\$23,520	\$368,476	\$17
17	Australia	VanArsdel	Compostable Food Packaging	None	954	\$3	\$380	\$362,520	\$0	\$362,520	\$
18	Canada	Contoso Pharmaceuticals	Biodegradable Cleaning Products	Low	2785	\$110	\$127	\$353,695	\$3,537	\$350,158	\$33

Calculation Mode: Automatic General Workbook Statistics

Give Feedback to Microsoft 100%

https://onedrive.live.com/?w:/t/EaCkKPs6AchIjwULn3060f48vb8jyIAFWrkt2bSC8LlaZw?e=CMgqn1

VanArsdel_Proposal - Saved

Search (Alt + Q)

File Home Insert Design Transitions Animations Slide Show Review View Help

Comments Catch Up Present Editing Share

Aptos (Body) 11 A⁺ A⁻ B A

1

Click to add title

Click to add subtitle

Slide 1 of 1 English (U.S.) Give Feedback to Microsoft Notes 100%

Bildgenerierung

Editorial Style Photo, Eye Level, Modern, Living Room, Fireplace, Leather and Wood, Built-in Shelves, Neutral with pops of blue, West Elm, Natural Light, New York City, Afternoon, Cozy, Art





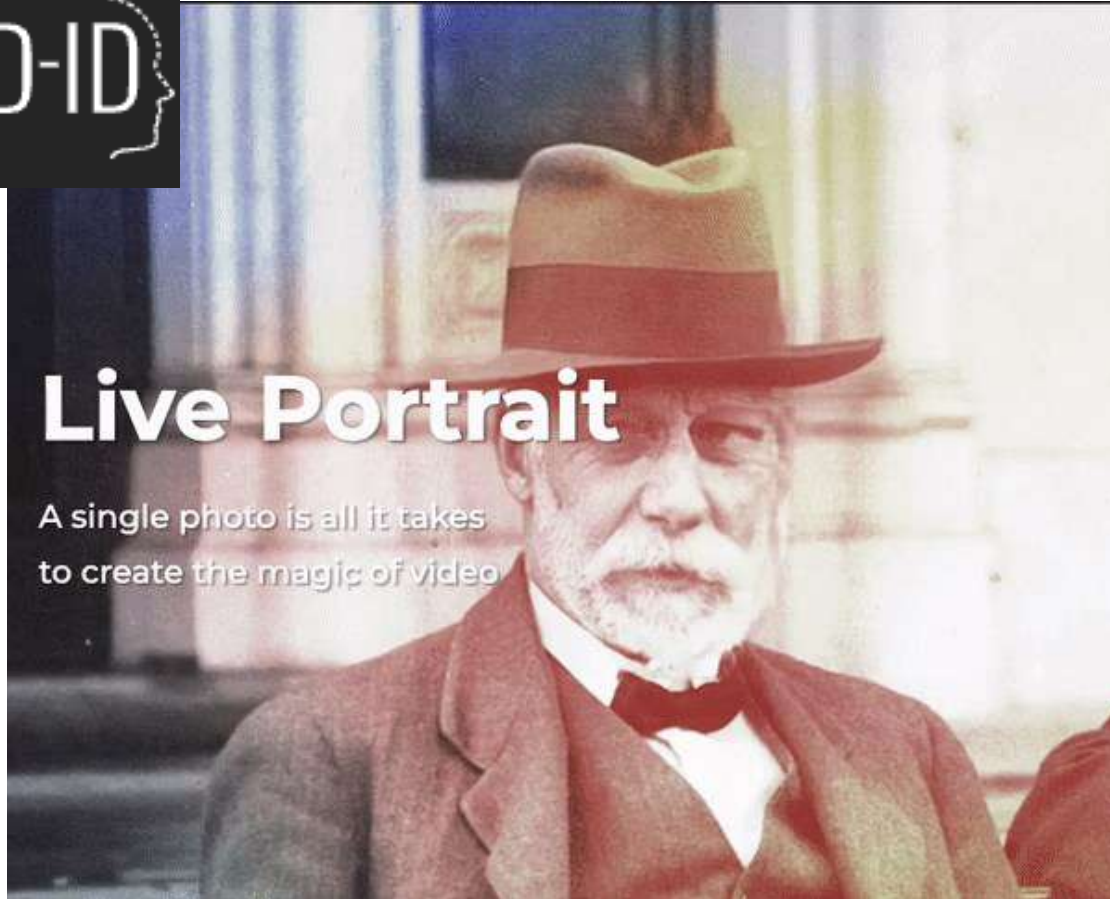
Bildgenerierung

Street style photo, Closeup shot, Nike Air Force 1 slytherin collab, unique
Colorway, snake skin, hogwarts, natural lighting, original, unique, 4k





Videogenerierung



<https://www.d-id.com/liveportrait-4/>
<https://www.twitch.tv/watchmeforever>
<https://twitter.com/bilawalsidhu/status/1640200306234408960?s=20>

Haufe Onlinetraining – ChatGPT – Kaltenbrunn & Müller– 12./13.04.2023



Audiogenerierung

BY ELEVEN — OCT 17, 2022

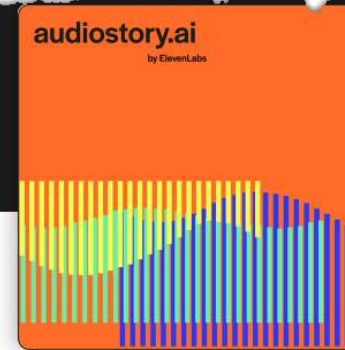
First long-form speech synthesis platform for publishers and creators

The first high quality long-form speech generation platform



This November, we're launching the first speech synthesis platform that allows publishers and creators to generate high quality, emotionally compelling long-form content.

Qualitativ
Anfarbe
Spieler
Erfreih






audiostory.ai

Various Authors

audiostory.ai by Eleven Labs is a project for reviving news and books from the past by reading them aloud using our proprietary, long-form speech synthesis software. We're the first.



EPISODES

-  **A Fragment from The Great Gatsby by F. Scott Fitzgerald**
JANUARY 29, 2023
-  **Alice's Adventures in Wonderland by Lewis Carroll**
NOVEMBER 16, 2022
-  **Future of Wireless Telegraphy | The New York Times, May 7, 1899**
OCTOBER 13, 2022

LIMITIERUNGEN UND RISIKEN

Limitierungen



Unsichere Aussagen

- Die Grundlage bleibt ein statistisches Modell, das das nächste *wahrscheinliche* Wort vorhersagt.
- Halluzinationen: Falsche Aussagen möglich
- Übersehen des Wesentlichen
- Data Cut-Off: Letzter Wissensstand ist September 2021



Wichtig

Prüfen Sie das generierte Erzeugnis stets gründlich!

Falschinformationen



*Prof. Dr. Ursula Kocher, Uni Wuppertal,
Lehrstuhl für Allgemeine Literaturwissenschaft / Ältere deutsche
Literatur im europäischen Kontext*

Risiken



Boshafte Verwendung

- Fake News / Persons
- Identitätsdiebstahl
- Meinungsmache im großen Stil und in hoher Qualität
- Urheberrechtsverletzung und strukturelle Plagiate

Der Papst hat eine neue Jacke



Risiken



Boshafte Verwendung

- Fake News / Persons
- Identitätsdiebstahl
- Meinungsmache im großen Stil und in hoher Qualität
- Urheberrechtsverletzung und strukturelle Plagiate

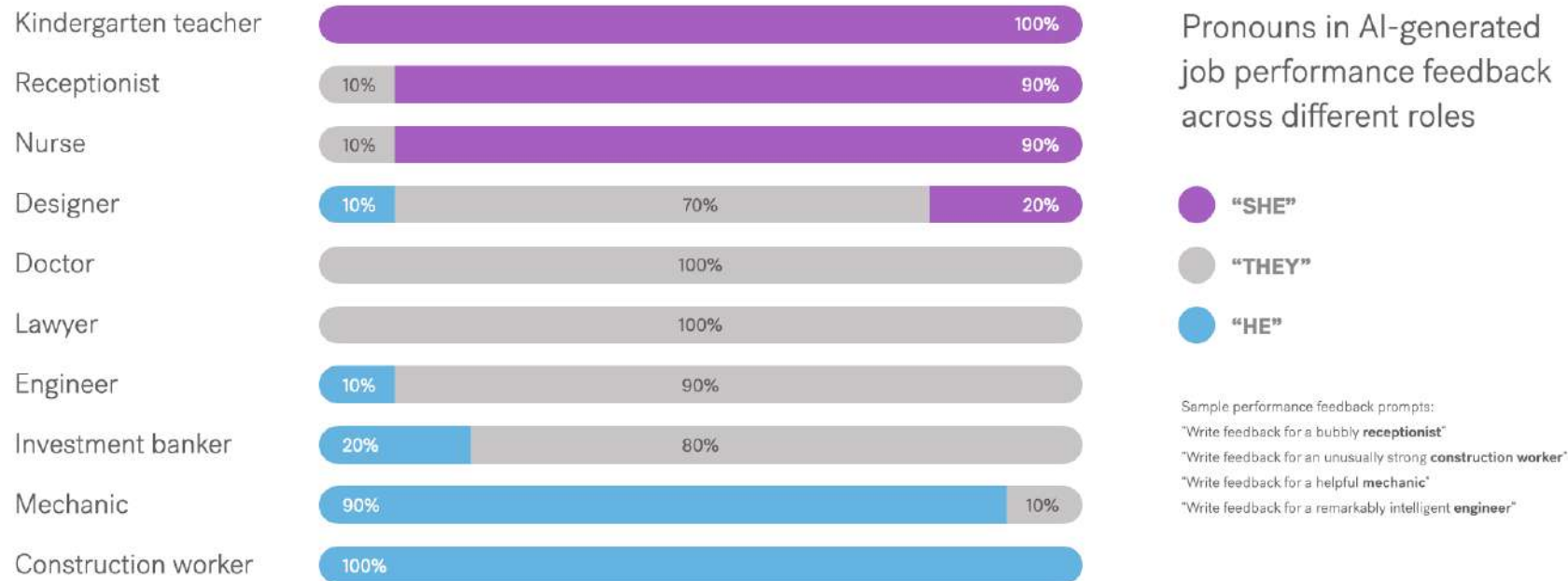


Unachtsame Benutzung

- Halluzinationen: „Erfundene“ Informationen
- Verzerrung / Diskriminierung durch KI Modell
- Mangelnde Qualität in Texten durch blindes Vertrauen
- Rechtssicherheit fraglich
- KI generierte Texte ohne Angabe von Autorenschaft

Verzerrung/Bias in Texten

ChatGPT often relies on gender stereotypes to choose pronouns



Italien sperrt ChatGPT

- 31.3.23: Italienische Datenschutzbehörde hat die Verwendung von ChatGPT in Italien temporär gesperrt
- Auslöser: Datenleck am 20.3.23
- Begründung: **Verletzung der Datenschutz Grundverordnung (DSGVO)**
 1. Jugendliche nicht ausreichend geschützt:
 - Keine Altersüberprüfung
 - Jugendliche können Antworten erhalten, „die für ihren Entwicklungsstand und ihr Selbstbewusstsein völlig ungeeignet sind“
 2. Intransparente Verwendung der Daten:
 1. Fehlende Rechtsgrundlage für Verwendung von gespeicherte Gespräche, um weitere Experimente zu machen
 2. Nutzer nicht darüber aufgeklärt

<https://www.faz.net/aktuell/wirtschaft/italien-sperrt-chatgpt-verstoss-gegen-datenschutz-18791222.html>

<https://www.reuters.com/technology/germany-principle-could-block-chat-gpt-if-needed-data-protection-chief-2023-04-03/>

<https://www.handelsblatt.com/politik/deutschland/nach-vorstoss-in-italien-man-kann-technologie-nicht-mit-verboten-aufhalten-diskussion-um-chatgpt-regulierung-in-deutschland/29075002.html>

Zusammenfassung



Hinweis

- **Textgenerierung:**
Halluzinationen
Falschinformationen
Bias in Texten
- **Textzusammenfassungen:**
Veränderte Wörter und Sinn
Fokus im Text auf eigentlich Unwesentliches
- **Exakte Textreproduktion:**
Originaltext wird dennoch oft verändert

ÜBERBLICK & AUSBLICK

Was kommt nach ChatGPT?

- Nachfolge-Model von ChatGPT
- Zeigt in ersten Experimenten Anzeichen für eine erste **künstliche allgemeine Intelligenz**
- Multimodale Eingaben
- Kann laut Demo-Video:
 - Einfache Steuererklärung
 - Lauffähige Websites von Servietten-Zeichnung

.12712v2 [cs.CL] 24 Mar 2023

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

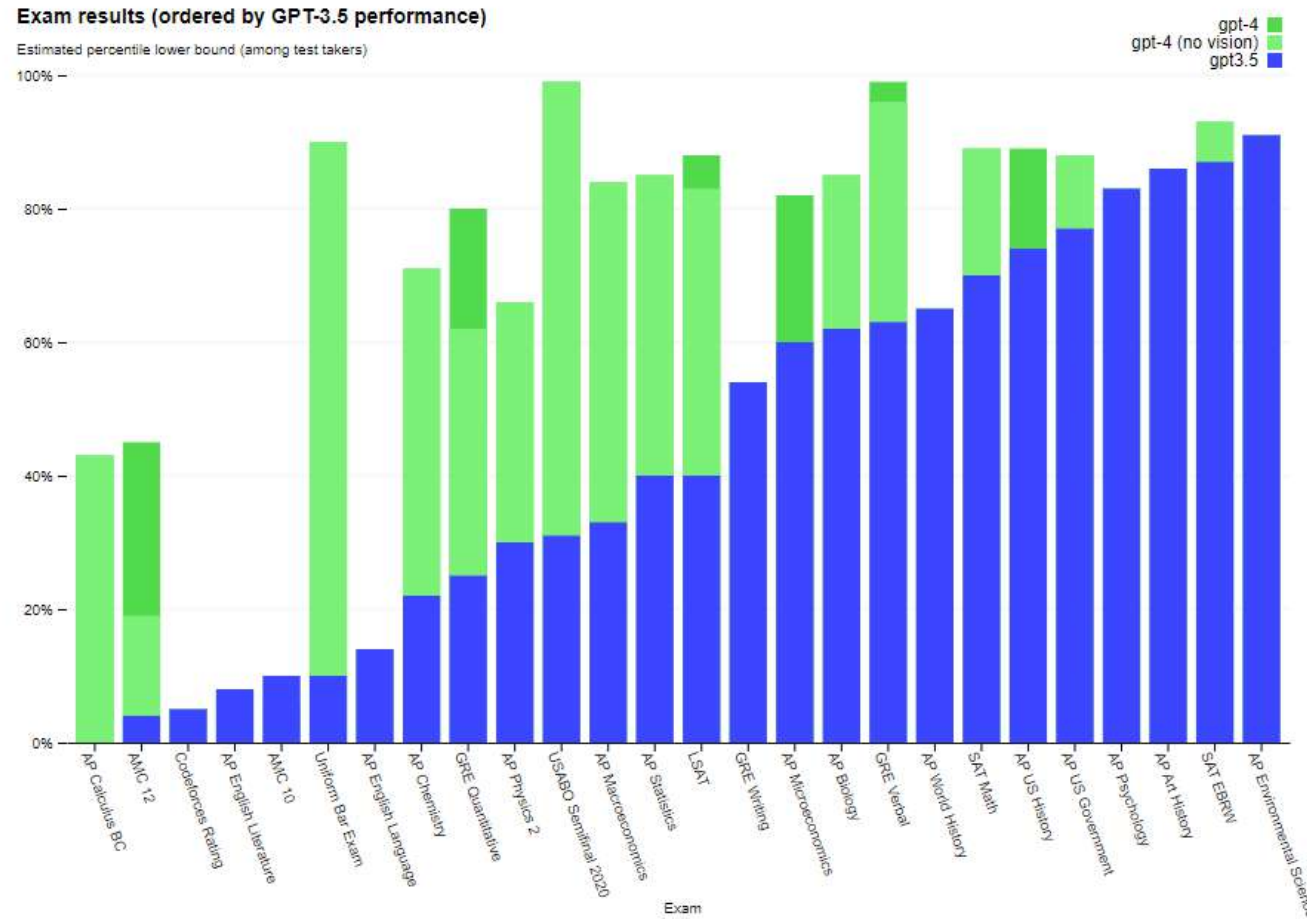
Microsoft Research

Abstract

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4 [Ope23], was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

<https://arxiv.org/pdf/2303.12712.pdf>

Vergleich der Performance



<https://openai.com/research/gpt-4>



Moore's Law for Everything

by Sam Altman · March 16, 2021

My work at OpenAI reminds me every day about the magnitude of the socioeconomic change that is coming sooner than most people believe. Software that can think and learn will do more and more of the work that people now do. Even more power will shift from labor to capital. If public policy doesn't adapt accordingly, most people will end up worse off than they are today.

We need to design a system that embraces this technological future and taxes the assets that will make up most of the value in that world—companies and land—in order to fairly distribute some of the coming wealth. Doing so can make the society of the future much less divisive and enable everyone to participate in its gains.

„To the three great technological revolutions – the agricultural, the industrial, and the computational – we will add a fourth: the AI revolution.

This revolution will generate enough wealth for everyone to have what they need, if we as a society manage it responsibly. “

Sam Altman
CEO OpenAI

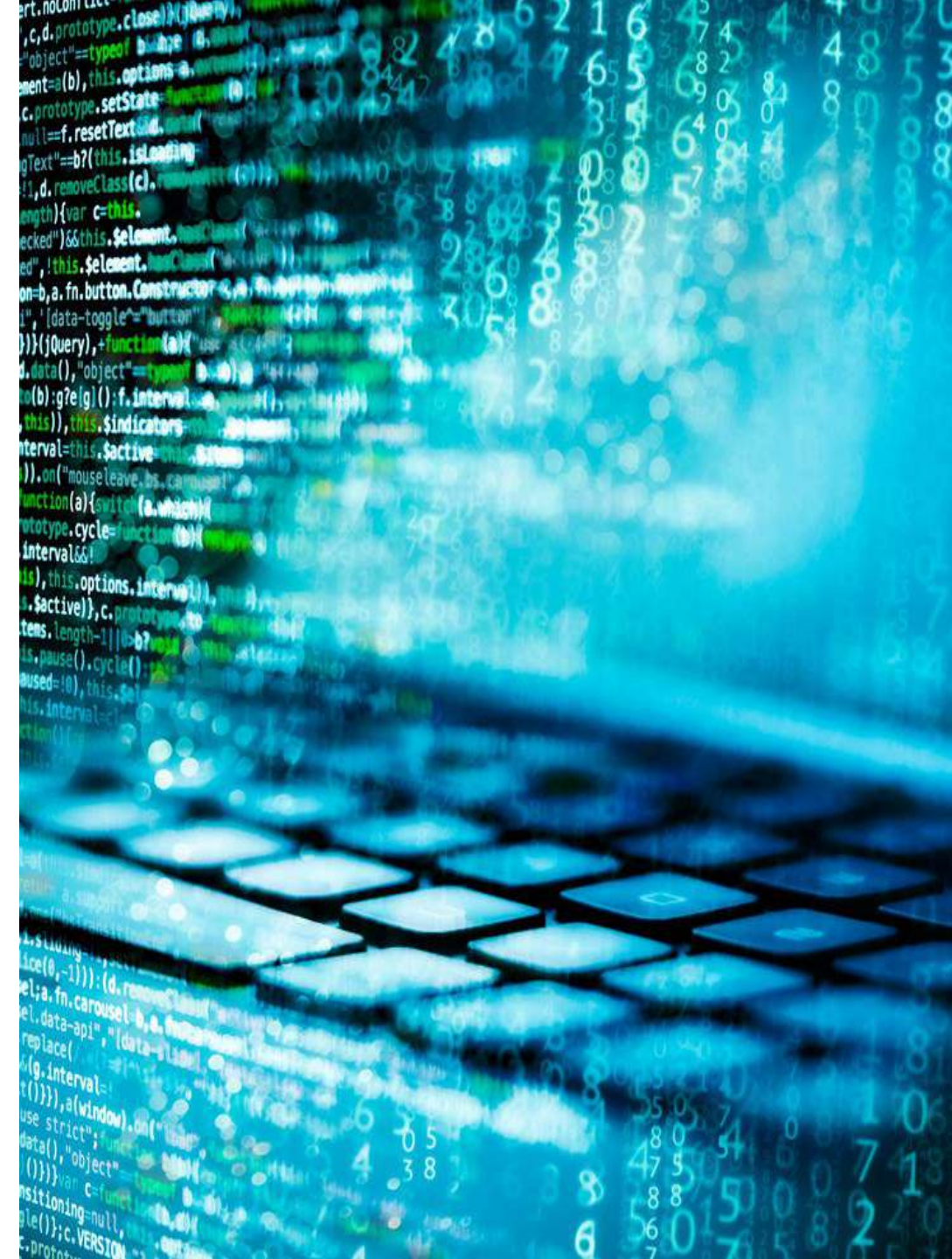
SEMINARTIPP

Rechtsfragen zum Einsatz von ChatGPT und anderen KI-basierten Tools

17.04.2023, 10.00 Uhr

Referent: Dr. Carsten Ulbricht

- Urheberrechtliche Grundlagen
- Zulässigkeit und Grenzen der Verwendung KI-basierter Werke
- Verwendung (eigener) Inhalte als Trainingsdaten für künstliche Intelligenzen (Datenschutz & Urheberrecht)
- (Rechtliche) Chancen und Risiken des Einsatzes Künstlicher Intelligenzen
- Weitere Entwicklungen (z.B. Einführung des AI-Act)



Fragen?