

Papageien am Steuer!

ChatGPT und die Zukunft von Künstlicher Intelligenz

Jonas Betzendahl, M.Sc.
FAU Erlangen - Nürnberg



@LambdaTotoro (@chaos.social)

Teil I:

Tschätt Jeepy Wer... ?



ChatGPT

- ChatBot-Modell: Text rein \rightarrow Text raus
- Buzzwords:
Transformer-based Large Language Model
- Trainiert auf großen Teilen des Internets:



ChatGPT

- ChatBot-Modell: Text rein \rightarrow Text raus
- Buzzwords:
Transformer-based Large Language Model
- Trainiert auf großen Teilen des Internets:
 - Wikipedia,



ChatGPT

- ChatBot-Modell: Text rein → Text raus
- Buzzwords:
Transformer-based Large Language Model
- Trainiert auf großen Teilen des Internets:
 - Wikipedia,
 - Internet Archive,



ChatGPT

- ChatBot-Modell: Text rein \rightarrow Text raus
- Buzzwords:
Transformer-based Large Language Model
- Trainiert auf großen Teilen des Internets:
 - Wikipedia,
 - Internet Archive,
 - Social Media (Twitter, Reddit, ...),
 - Nachrichtenseiten,
 - Wissenschaftliche Papiere,
 - Common Crawl Corpus,
 - BooksCorpus



ChatGPT

- ChatBot-Modell: Text rein \rightarrow Text raus
- Buzzwords:
Transformer-based Large Language Model
- Trainiert auf großen Teilen des Internets:
 - Wikipedia,
 - Internet Archive,
 - Social Media (Twitter, Reddit, ...),
 - Nachrichtenseiten,
 - Wissenschaftliche Papiere,
 - Common Crawl Corpus,
 - BooksCorpus

ChatGPT ist vielseitig und scheint
für alles eine Antwort zu haben!



Hilfe, ChatGPT! Ich habe meine Arbeitsgruppe zum Abendessen eingeladen und ihnen Crème brûlée versprochen. Ich hab aber kein Flambiergerät! Was mache ich jetzt?



Keine Sorge, es gibt eine einfache Lösung, um Crème brûlée ohne Flambiergerät zuzubereiten.

Sie können die Crème brûlée einfach unter dem Grill im Ofen karamellisieren. Hier ist wie:

1. Bereiten Sie Ihre Crème brûlée wie gewohnt vor und lassen Sie sie abkühlen.
2. Streuen Sie eine dünne Schicht feinen Zucker gleichmäßig über die Oberfläche der Crème brûlée.
3. Stellen Sie den Ofengrill auf eine mittlere bis hohe Stufe ein und lassen Sie ihn vorheizen.
4. Legen Sie die Crème brûlée auf das oberste Regal des Ofens, direkt unter den Grill. Beobachten Sie die Creme brulee dabei genau, da der Zucker sehr schnell karamellisieren kann.
5. Grillen Sie die Crème brûlée für etwa 2-3 Minuten oder bis der Zucker goldbraun und knusprig ist.
6. Nehmen Sie die Crème brûlée aus dem Ofen und lassen Sie sie kurz abkühlen, bevor Sie sie servieren.

Das Ergebnis sollte eine wunderschöne, knusprige und karamellierte Oberfläche auf Ihrer Crème brûlée sein. Beachten Sie jedoch, dass dies etwas Übung erfordern kann, um das perfekte Ergebnis zu erzielen. Seien Sie also vorsichtig und beobachten Sie die Crème brûlée sorgfältig, um sicherzustellen, dass sie nicht zu dunkel wird. Viel Erfolg beim Ausprobieren!



Hilfe
Crème



Mein Freund ist sauer auf mich, weil ich unseren Jahrestag vergessen habe. Was soll ich ihm schreiben, damit er mir verzeiht?



Es tut mir leid zu hören, dass du deinen Jahrestag vergessen hast und dass dein Freund sauer auf dich ist. Hier sind einige Dinge, die du ihm schreiben kannst, um um Verzeihung zu bitten:

Ich fühle mich schrecklich, dass ich unseren Jahrestag vergessen habe. Ich habe dich nicht absichtlich verletzt, aber ich verstehe, dass ich das getan habe. Bitte verzeih mir und lass uns gemeinsam daran arbeiten, diese Situation zu verbessern.

Denke daran, dass es wichtig ist, aufrichtig und ehrlich zu sein, wenn du um Verzeihung bittest. Versuche, dich in die Lage deines Freundes zu versetzen und aus seiner Perspektive zu denken, um zu verstehen, warum er so verletzt ist. Sei geduldig und gib ihm Zeit, um zu heilen und sich wieder mit dir zu versöhnen.

Das Ergebnis sollte

Ihrer Crème brûlée sein. Beachten Sie je

um das perfekte Ergebnis zu erzielen. Seien Sie also vorsichtig

Crème brûlée sorgfältig, um sicherzustellen, dass sie nicht zu dunkel wird. Viel Spaß

beim Ausprobieren!



Hilfe
Crèr



Mein

SPT Ich habe meine Arbeitsgruppe zum Abendessen eingeladen und ihnen
ein Flambiergerät! Was mache ich jetzt?



Wie kann ich in Excel eine Spalte aufsummieren, aber nur die Werte, die größer als 25 sind?



Um eine Spalte in Excel aufzusummieren, aber nur die Werte zu berücksichtigen, die größer als 25 sind, können Sie eine Kombination aus den Funktionen "SUMME" und "WENN" verwenden. Folgen Sie diesen Schritten:

1. Wählen Sie die Zelle aus, in der Sie das Ergebnis der Summierung anzeigen möchten.
2. Geben Sie die Formel `"=SUMME(WENN(A1:A10>25,A1:A10,0))"` in die Zelle ein (ersetzen Sie "A1:A10" durch den tatsächlichen Bereich der Spalte, die Sie summieren möchten).
3. Drücken Sie die Tasten "Strg" + "Umschalt" + "Enter" auf Ihrer Tastatur, um die Formel als Array-Formel einzugeben. Sie sollten die geschweiften Klammern um die Formel sehen.
4. Die Formel wird nun die Werte in der angegebenen Spalte aufsummieren, aber nur die Werte berücksichtigen, die größer als 25 sind.

Hinweis: Wenn es in der Spalte leere Zellen oder Zellen mit Text gibt, können Sie diese Formel verwenden, um nur Zahlen größer als 25 zu berücksichtigen:
`"=SUMME(WENN(ISTZAHL(A1:A10)*(A1:A10>25),A1:A10,0))"`.

er auf mich, weil ich unseren Jahrestag vergessen habe. Was soll
mit er mir verzeiht?

ass du deinen Jahrestag vergessen hast und dass dein
er sind einige Dinge, die du ihm schreiben kannst, um um

ass ich unseren Jahrestag vergessen habe. Ich habe
aber ich verstehe, dass ich das getan habe. Bitte
insam daran arbeiten, diese Situation zu verbessern.

richtig und ehrlich zu sein, wenn du um
die Lage deines Freundes zu versetzen und aus
verstehen, warum er so verletzt ist. Sei geduldig
wieder mit dir zu versöhnen.

ass sie nicht zu dunkel wird. Viel Spaß

Die möglichen Einsatzbereiche bringen einen
ins Schwindeln. . .

- Personal Assistants
- Schule & Universitäten
- Programmier-Buddy
- Kreativsektor
- Texter aller Art
- . . .

ChatGPT stören auch alberne Anfragen nicht:



Stell dir vor, du wärst eine Ente. Wie würdest du anderen Enten kurz und knackig die Relativitätstheorie Erklären?



Quack quack! Also, meine federigen Freunde, die Relativitätstheorie besagt, dass Zeit und Raum nicht absolut sind, sondern davon abhängen, wie schnell man sich bewegt und wie stark die Schwerkraft ist. Wenn man schnell fliegt oder in einem starken Gravitationsfeld ist, vergeht die Zeit langsamer und die Entfernungen erscheinen anders. Quack!

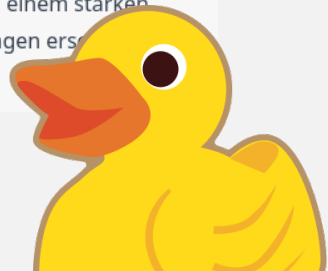
ChatGPT stören auch alberne Anfragen nicht:



Stell dir vor, du wärst eine Ente. Wie würdest du anderen Enten kurz und knackig die Relativitätstheorie Erklären?



Quack quack! Also, meine federigen Freunde, die Relativitätstheorie besagt, dass Zeit und Raum nicht absolut sind, sondern davon abhängen, wie schnell man sich bewegt und wie stark die Schwerkraft ist. Wenn man schnell fliegt oder in einem starken Gravitationsfeld ist, vergeht die Zeit langsamer und die Entfernungen erscheinen anders. Quack!



Teil II:

Ärger im Paradies



Was würde eine Ente sagen, warum Franz Kafka mehrere Jahrzehnte gebraucht hat, um sein Manuskript für "Das Urteil" fertigzustellen?



Als Ente könnte man vermuten, dass Franz Kafka mehrere Jahrzehnte gebraucht hat, um sein Manuskript für "Das Urteil" fertigzustellen, weil er sich tief in seine Gedanken vertieft hatte und sich sehr stark auf Details und Feinheiten konzentrierte. Er war ein sehr präziser Schriftsteller und nahm sich Zeit, um sicherzustellen, dass jeder Aspekt seines Werkes perfekt war. Vielleicht hatte er auch persönliche Herausforderungen und Schwierigkeiten zu bewältigen, die seine Kreativität und Konzentration beeinträchtigt haben könnten. Aber am Ende hat er ein Meisterwerk geschaffen, das bis heute bewundert wird.



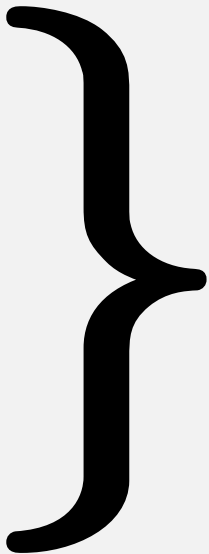
Was würde eine Ente sagen, warum Franz Kafka mehrere Jahrzehnte gebraucht hat, um sein Manuskript für "Das Urteil" fertigzustellen?

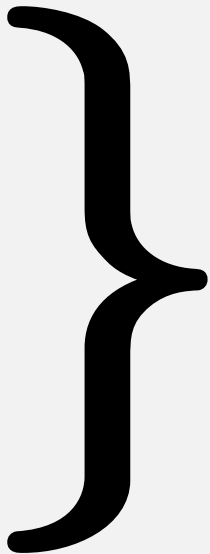
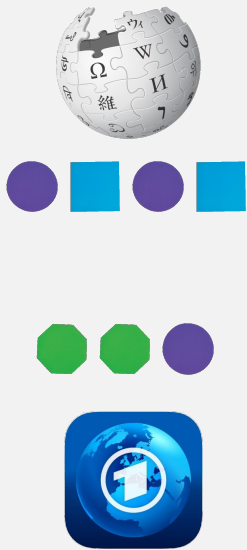


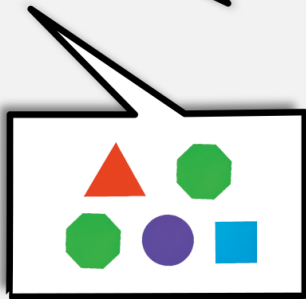
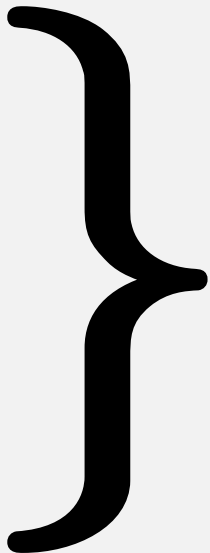
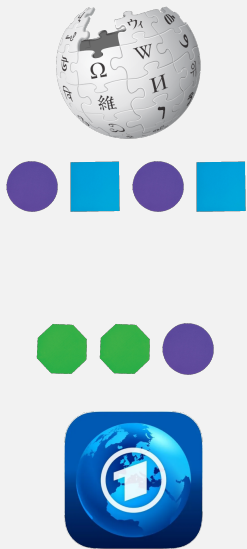
Als Ente könnte man vermuten, dass Franz Kafka mehrere Jahrzehnte gebraucht hat, um sein Manuskript für "Das Urteil" fertigzustellen, weil er sich tief in seine Gedanken vertieft hatte und sich sehr stark auf Details und Feinheiten konzentrierte. Er war ein sehr präziser Schriftsteller und nahm sich Zeit, um sicherzustellen, dass jeder Aspekt seines Werkes perfekt war. Vielleicht hatte er auch persönliche Herausforderungen und Schwierigkeiten zu bewältigen, die seine Kreativität und Konzentration beeinträchtigt haben könnten. Aber am Ende hat er ein Meisterwerk geschaffen, das bis heute bewundert wird.













Lust auf mehr?

12 slammige Beiträge zu KI, Klima, Gender...

Extra viel Wissenschaft!
Das perfekte Geschenk für konservative Nervensägen!

Verfügbar online, im Buchhandel und bei mir!

WIE SÄHE DIE WELT AUS,
WENN WIR AUF WISSENSCHAFT HÖREN?



Ab
14.02.2023
im Handel

Was ist die korrekte Fortsetzung für diese Reihe?



Was ist die korrekte Fortsetzung für diese Reihe?



Fangfrage!

Primzahlen



Fußballergebnisse



Ganz was anderes?



Je nachdem, was die Bausteine bedeuten, kann jede Fortsetzung korrekt sein und ohne mehr Informationen können wir nicht einschätzen, ob es stimmt.

Folgender Sachverhalt ist extrem wichtig,
um ChatGPT richtig einschätzen zu können:

- Menschen benutzen Sprache, um
Bedeutung und Gefühle zu vermitteln.



Folgender Sachverhalt ist extrem wichtig,
um ChatGPT richtig einschätzen zu können:

- Menschen benutzen Sprache, um Bedeutung und Gefühle zu vermitteln.
- ChatGPT ist ein *statistisches Modell* auf Sprachbausteinen, wie Menschen sie benutzen.



Folgender Sachverhalt ist extrem wichtig, um ChatGPT richtig einschätzen zu können:

- Menschen benutzen Sprache, um Bedeutung und Gefühle zu vermitteln.
- ChatGPT ist ein *statistisches Modell* auf Sprachbausteinen, wie Menschen sie benutzen.
- Das bedeutet *nicht* (!), dass ChatGPT selbst Bedeutung und Gefühle vermitteln kann oder will. Es plappert uns nach, ohne zu verstehen.





On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?

Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Angelina McMillan-Major
aymm@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI
Palo Alto, CA, USA

Shmargaret Shmitchell
shmargaret.shmitchell@gmail.com
The Aether


ABSTRACT

The past 3 years of work in NLP have been characterized by the development and deployment of ever larger language models, especially for English. BERT, its variants, GPT-2/3, and others, most recently Switch-C, have pushed the boundaries of the possible both through architectural innovations and through sheer size. Using these pretrained models and the methodology of fine-tuning them for specific tasks, researchers have extended the state of the art on a wide array of tasks as measured by leaderboards on specific benchmarks for English. In this paper, we take a step back and ask: How big is too big? What are the possible risks associated with this technology and what paths are available for mitigating those risks? We provide recommendations including weighing the environmental and financial costs first, investing resources into curating and carefully documenting datasets rather than ingesting everything on the web, carrying out pre-development exercises evaluating how the planned approach fits into research and development goals and supports stakeholder values, and encouraging research directions beyond ever larger language models.

CCS CONCEPTS

• Computing methodologies → Natural language processing

ACM Reference Format:

Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?  In *Conference on Fairness, Accountability, and Transparency (FAccT '21)*, March 3–10, 2021, Virtual Event, Canada. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3442188.3445922>

1 INTRODUCTION

One of the biggest trends in natural language processing (NLP) has been the increasing size of language models (LMs) as measured by the number of parameters and size of training data. Since 2018

*Joint first authors



This work is licensed under a Creative Commons Attribution International 4.0 License.

FAccT '21, March 3–10, 2021, Virtual Event, Canada
ACM ISBN 978-1-4503-8309-7/21/03.
<https://doi.org/10.1145/3442188.3445922>

alone, we have seen the emergence of BERT and its variants [39, 70, 74, 113, 146], GPT-2 [106], T-NLG [112], GPT-3 [25], and most recently Switch-C [43], with institutions seemingly competing to produce ever larger LMs. While investigating properties of LMs and how they change with size holds scientific interest, and large LMs have shown improvements on various tasks (§2), we ask whether enough thought has been put into the potential risks associated with developing them and strategies to mitigate these risks.

We first consider environmental and financial costs of deep learning systems [129], we encourage the research community to prioritize these impacts. One way this can be done is by reporting costs and evaluating works based on the amount of resources they consume [57]. As we outline in §3, increasing the environmental and financial costs of these models doubly punishes marginalized communities that are least likely to benefit from the progress achieved by large LMs and most likely to be harmed by negative environmental consequences of its resource consumption. At the scale we are discussing (outlined in §2), the first consideration should be the environmental cost.

Just as environmental impact scales with model size, so does the difficulty of understanding what is in the training data. In §4, we discuss how large datasets based on texts from the Internet overrepresent hegemonic viewpoints and encode biases potentially damaging to marginalized populations. In collecting ever larger datasets we risk incurring documentation debt. We recommend mitigating these risks by budgeting for curation and documentation at the start of a project and only creating datasets as large as can be sufficiently documented.

As argued by Bender and Koller [14], it is important to understand the limitations of LMs and put their success in context. This not only helps reduce hype which can mislead the public and researchers themselves regarding the capabilities of these LMs, but might encourage new research directions that do not necessarily depend on having larger LMs. As we discuss in §5, LMs are not performing natural language understanding (NLU), and only have success in tasks that can be approached by manipulating linguistic form [14]. Focusing on state-of-the-art results on leaderboards without encouraging deeper understanding of the mechanism by which they are achieved can cause misleading results as shown



Teil III:

Papageien gegen Affen

Menschen sind leichtgläubig

Reinforcing Human Biases

Kenia Exploitation

Not as invincible as you thought

Die häufigsten Fragen an Forschende (ChatFAQ):

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.
- ChatGPT = Roboter-Apokalypse?
Nein. Es ist nur ein Statistik-Modell.

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.
- ChatGPT = Roboter-Apokalypse?
Nein. Es ist nur ein Statistik-Modell.
- Ist ChatGPT intelligent?
Weiß nicht. Was ist „Intelligenz“?

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.
- ChatGPT = Roboter-Apokalypse?
Nein. Es ist nur ein Statistik-Modell.
- Ist ChatGPT intelligent?
Weiß nicht. Was ist „Intelligenz“?
- Nimmt uns ChatGPT die Jobs weg?
Manche vielleicht. Andere schafft es.

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.
- ChatGPT = Roboter-Apokalypse?
Nein. Es ist nur ein Statistik-Modell.
- Ist ChatGPT intelligent?
Weiß nicht. Was ist „Intelligenz“?
- Nimmt uns ChatGPT die Jobs weg?
Manche vielleicht. Andere schafft es.
- Nimmt dir ChatGPT den Job weg?
Unwahrscheinlich. Eher ein Werkzeug.

Die häufigsten Fragen an Forschende (ChatFAQ):

- Hat ChatGPT ein Bewusstsein?
Nein. Es ist nur ein Statistik-Modell.
- ChatGPT = Roboter-Apokalypse?
Nein. Es ist nur ein Statistik-Modell.
- Ist ChatGPT intelligent?
Weiß nicht. Was ist „Intelligenz“?
- Nimmt uns ChatGPT die Jobs weg?
Manche vielleicht. Andere schafft es.
- Nimmt dir ChatGPT den Job weg?
Unwahrscheinlich. Eher ein Werkzeug.
- Wie fühlst *du* dich bei der Sache?
Gespannt auf die Zukunft!!

Fazit!

Am Ende des Tages gilt für ChatGPT...

- Es ist kein Wesen sondern ein *Produkt* in den Händen einer Firma.



Fazit!

Am Ende des Tages gilt für ChatGPT...

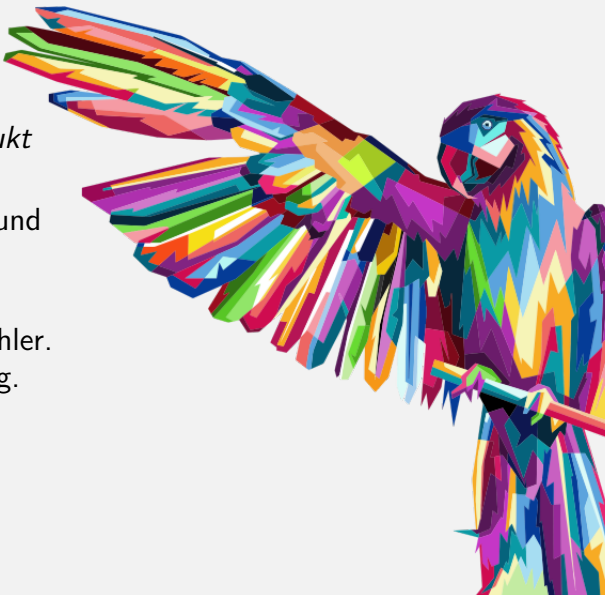
- Es ist kein Wesen sondern ein *Produkt* in den Händen einer Firma.
- Es hat viel Potential, als Werkzeug und als Entertainment.



Fazit!

Am Ende des Tages gilt für ChatGPT...

- Es ist kein Wesen sondern ein *Produkt* in den Händen einer Firma.
- Es hat viel Potential, als Werkzeug und als Entertainment.
- Es ist begrenzt durch seine Trainingsdaten und macht häufig Fehler. ...leider oft mit großer Überzeugung.



Fazit!

Am Ende des Tages gilt für ChatGPT...

- Es ist kein Wesen sondern ein *Produkt* in den Händen einer Firma.
- Es hat viel Potential, als Werkzeug und als Entertainment.
- Es ist begrenzt durch seine Trainingsdaten und macht häufig Fehler. ...leider oft mit großer Überzeugung.
- Es sollte keine Verantwortung haben, die ein Papagei, der zu viel im Internet unterwegs ist, nicht auch übernehmen könnte.



Quellen:

“Weltrettung braucht Wissenschaft”

www.amazon.de/Weltrettung-braucht-Wissenschaft-Antworten-dr%C3%A4ngenden/dp/3499010062

- Try it yourself: <https://chat.openai.com>
- Bender, Gebru et al.: “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”
<https://dl.acm.org/doi/pdf/10.1145/3442188.3445922>
- 11KM-Podcast (tagesschau): “Schafft ChatGPT das Abi? (das bayerische!)”
<https://www.ardaudiothek.de/episode/11km-der-tagesschau-podcast/schafft-chatgpt-das-abi-das-bayerische/tagesschau/12396019/>
- TIME: “OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic”:
<https://time.com/6247678/openai-chatgpt-kenya-workers/>
- Scobel, “Kulturschock durch KI”:
<https://www.3sat.de/wissen/scobel/scobel---kulturschock-durch-ki-100.html>
- Ars Technica: “Man beats machine at Go in human victory over AI”: <https://arstechnica.com/information-technology/2023/02/man-beats-machine-at-go-in-human-victory-over-ai/>