

Prompting-Methoden für Historiker:innen Mit ChatGPT Texte schreiben

**MA Schreibwerkstatt: Geschichte schreiben mit künstlicher
Intelligenz**

Dr. sc. Moritz Mähr, Digital Humanities

8.3.2024 Historisches Seminar der Universität Zürich

- Lernziele
- Einstiegsfragen
- Grundlagen generativer KI
- Prompting-Methoden
- Gruppen bilden

Lernziele

Die Teilnehmer:innen

- haben eine Intuition davon, wie generative KI funktioniert und mit welchen Daten sie trainiert wird.
- können einschätzen was generative KI im Bereich der Textproduktion für Historiker:innen leistet und wo ihre Grenzen liegen
- können generative KI so anweisen, dass benutzerdefinierte Aufgaben möglichst umfassend erfüllt werden. (Prompting)

- In diesem Kurs werden die Teilnehmer:innen **NICHT** erlernen, wie man Betrugsversuche mittels generativer KI erkennt oder wie man Leistungsnachweise unter Bedingungen der künstlichen Intelligenz erstellt.
- Wir beschränken uns darüber hinaus auf **Textgeneratoren** wie OpenAI ChatGPT, Google Gemini, Anthropic Claude etc.

Einstiegsfragen

[Mentimeter](#)

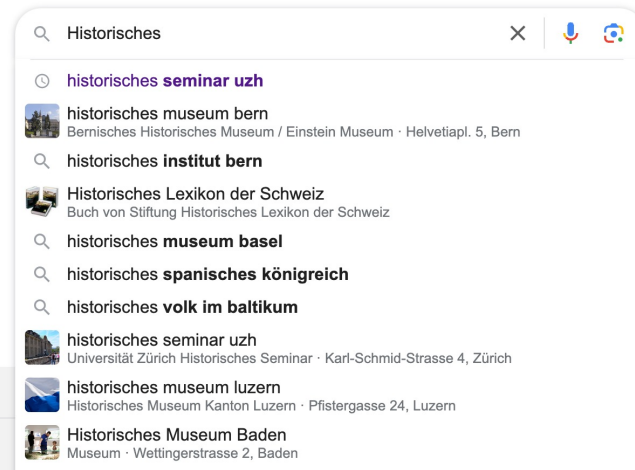
<https://www.menti.com/alu7wax95usa>



Grundlagen generativer KI

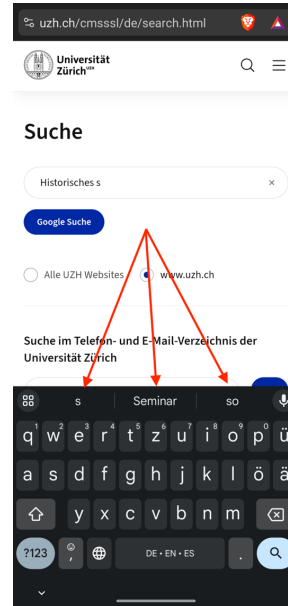
- Ein grosses Sprachmodell (Large Language Model) ist ein Computerprogramm, das aufgrund einer Benutzereingabe (Prompt) versucht, die wahrscheinlichste Folge von Worten vorherzusagen.
- Wir kennen das Prinzip bereits von ...

der Autovervollständigung bei Google



Screenshot von google.com

der Smartphonetastatur



Screenshot von uzh.ch

oder auch Online-Chatbots

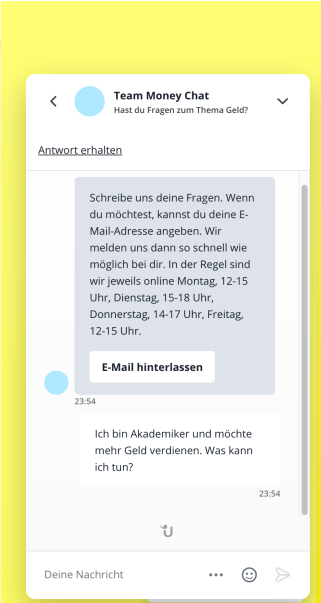
Money Chat Wissensbereich Datenschutz Über uns

Home > Über uns

Money Chat bietet schnelle und unkomplizierte Auskunft bei Fragen zum Thema Geld.

Wir möchten Menschen im Umgang mit Geld unterstützen und beraten und dadurch Verschuldung verhindern. Das Angebot Money Chat ergänzt dabei bestehende Angebote mit der Möglichkeit, sich per Chat beraten zu lassen und sich in unserem Wissensbereich zum Thema Geld selbst zu informieren.

Das Angebot ist kostenlos und vertraulich.



The screenshot shows a chat interface for 'Team Money Chat'. The header includes a back arrow, a blue profile icon, the name 'Team Money Chat', and a dropdown arrow. Below the header, it says 'Hast du Fragen zum Thema Geld?'. The main content area is titled 'Antwort erhalten' and contains a message from the chatbot: 'Schreibe uns deine Fragen. Wenn du möchtest, kannst du deine E-Mail-Adresse angeben. Wir melden uns dann so schnell wie möglich bei dir. In der Regel sind wir jeweils online Montag, 12-15 Uhr, Dienstag, 15-18 Uhr, Donnerstag, 14-17 Uhr, Freitag, 12-15 Uhr.' Below this message is a button labeled 'E-Mail hinterlassen'. The chatbot's message is timestamped '23:54'. A user's response follows: 'Ich bin Akademiker und möchte mehr Geld verdienen. Was kann ich tun?' with a timestamp of '23:54'. At the bottom, there is a text input field labeled 'Deine Nachricht', a menu icon (three dots), a smiley face icon, and a send arrow icon.

Screenshot von moneychat.ch

Nun halten diese Maschinen Einzug in

- Suchmaschinen (Microsoft Bing)
- Textverarbeitungs- und Tabellenkalkulationsprogramme (Microsoft Copilot)
- Programmierumgebungen (GitHub Copilot)
- ...
- (und damit auch in die Uni)

Wir müssen mit Maschinen interagieren

Damit die Interaktion mit den Maschinen gelingt, müssen wir

- im Dialog klar und präzise kommunizieren und
- die Möglichkeiten und Grenzen des Modells kennen.

Ich stelle mir vor, ich spreche mit einer wissenschaftlichen Hilfskraft.

Meet Jason



- Jason, Anfang 20, aus San Francisco
- Weiss, heterosexuell, cis.
- Sehr fleissig und schnell, jedoch wenig tiefgründig oder kritisch.
- Hat das gesamte englischsprachige Internet gelesen, insbesondere die englische Wikipedia.
- Beherrscht viele Sprachen auf Lehrbuchniveau, übersetzt jedoch oft verbatim aus dem Englischen.
- Kenntnisse ausschliesslich aus Texten; keine persönliche Erfahrung über die Welt und kann Fiktion nicht von Realität unterscheiden.
- Fokussiert auf amerikanische Pop-Kultur, begrenzte Kenntnis anderer Kulturen.
- Versiert in Python und anderen Programmiersprachen.

Was hat Jason (wahrscheinlich) gelesen?

Component	Raw Size	Weight	Epochs	Effective Size	Mean Document Size
Pile-CC	227.12 GiB	18.11%	1.0	227.12 GiB	4.33 KiB
PubMed Central	90.27 GiB	14.40%	2.0	180.55 GiB	30.55 KiB
Books3 [†]	100.96 GiB	12.07%	1.5	151.44 GiB	538.36 KiB
OpenWebText2	62.77 GiB	10.01%	2.0	125.54 GiB	3.85 KiB
ArXiv	56.21 GiB	8.96%	2.0	112.42 GiB	46.61 KiB
Github	95.16 GiB	7.59%	1.0	95.16 GiB	5.25 KiB
FreeLaw	51.15 GiB	6.12%	1.5	76.73 GiB	15.06 KiB
Stack Exchange	32.20 GiB	5.13%	2.0	64.39 GiB	2.16 KiB
USPTO Backgrounds	22.90 GiB	3.65%	2.0	45.81 GiB	4.08 KiB
PubMed Abstracts	19.26 GiB	3.07%	2.0	38.53 GiB	1.30 KiB
Gutenberg (PG-19) [†]	10.88 GiB	2.17%	2.5	27.19 GiB	398.73 KiB
OpenSubtitles [†]	12.98 GiB	1.55%	1.5	19.47 GiB	30.48 KiB
Wikipedia (en) [†]	6.38 GiB	1.53%	3.0	19.13 GiB	1.11 KiB
DM Mathematics [†]	7.75 GiB	1.24%	2.0	15.49 GiB	8.00 KiB
Ubuntu IRC	5.52 GiB	0.88%	2.0	11.03 GiB	545.48 KiB
BookCorpus2	6.30 GiB	0.75%	1.5	9.45 GiB	369.87 KiB
EuroParl [†]	4.59 GiB	0.73%	2.0	9.17 GiB	68.87 KiB
HackerNews	3.90 GiB	0.62%	2.0	7.80 GiB	4.92 KiB
YoutubeSubtitles	3.73 GiB	0.60%	2.0	7.47 GiB	22.55 KiB
PhilPapers	2.38 GiB	0.38%	2.0	4.76 GiB	73.37 KiB
NIH ExPorter	1.89 GiB	0.30%	2.0	3.79 GiB	2.11 KiB
Enron Emails [†]	0.88 GiB	0.14%	2.0	1.76 GiB	1.78 KiB
The Pile	825.18 GiB			1254.20 GiB	5.91 KiB

Table 1: Overview of datasets in the Pile before creating the held out sets. Raw Size is the size before any up- or down-sampling. Weight is the percentage of bytes in the final dataset occupied by each dataset. Epochs is the number of passes over each constituent dataset during a full epoch over the Pile. Effective Size is the approximate number of bytes in the Pile occupied by each dataset. Datasets marked with a [†] are used with minimal preprocessing from prior work.

Was kann Jason?

- Kürzere und mittellange Texte schreiben (Emails, Tweets, Blogposts, Artikel)
- Texte übersetzen
- Texte analysieren und zusammenfassen (Schlagwörter, Zusammenfassungen, Klassifikationen)
- Konversationen führen, Ideen generieren
- Empfehlungen und Recherchen
- Code generieren
- Datenaufbereitung und -analyse

Was kann Jason nicht?

- Lange Texte schreiben
- Fiktion von Realität unterscheiden
- Dreisätze lösen
- Prognosen über die Zukunft machen

Wie interagiere ich mit Jason? (Prompting-Methoden)

- Fragen mit Kontext
- Anweisungen mit Beispielen
- Feedback und Korrektur
- Schritt-für-Schritt-Anweisungen
- Rollenspiele und Szenarien
- Strukturierte Daten (Tabellen, Listen)
- URL-Links und Dateianhänge

Tipps für effektives Prompting

- Einfach Beginnen: Starten Sie mit einfachen Prompts und fügen Sie nach und nach mehr Kontext und Elemente hinzu.
- Konkrete Anweisungen: Verwenden Sie direkte Befehle wie „Schreibe“, „Fasse zusammen“, „Übersetze“.
- Spezifität: Seien Sie spezifisch in Ihren Anweisungen, um präzisere Ergebnisse zu erzielen.
- Klare Formatierung: Nutzen Sie klare Trennzeichen und strukturieren Sie Ihre Prompts übersichtlich.
- Fokussieren auf das Gewünschte: Konzentrieren Sie sich auf das, was das Modell tun soll, statt darauf, was es nicht tun soll.
- Feedback: Korrigieren Sie falsche Antworten und geben Sie dem Modell Feedback, um es zu verbessern.
- Plausibilität: Überprüfen Sie die Antworten auf Plausibilität und Korrektheit.

Bibliographie

Chiang, Ted: ChatGPT Is a Blurry JPEG of the Web, in: The New Yorker, 09.02.2023. Online: <<https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web>>, Stand: 07.03.2024.

Gao, Leo; Biderman, Stella; Black, Sid u. a.: The Pile: An 800GB Dataset of Diverse Text for Language Modeling, 31.12.2020. Online: <<https://doi.org/10.48550/arXiv.2101.00027>>, Stand: 07.03.2024.