

# Prompting-Methoden für Historiker:innen Mit ChatGPT Texte schreiben

## Dozierenden-Workshop

**Dr. sc. Moritz Mähr, Digital Humanities**

15.3.2024 Historisches Seminar der Universität Zürich

# Überblick

---

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

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 (Prompting)**, dass benutzerdefinierte Aufgaben möglichst umfassend erfüllt werden.

# Disclaimer

---

- Wir konzentrieren uns auf weit verbreitete (kommerzielle) **Textgeneratoren** wie OpenAI ChatGPT, Google Gemini, Anthropic Claude etc.
- Die Erkennung von **Betrugsversuchen** mittels generativer KI oder die Erstellung von **Leistungsnachweisen** unter den Bedingungen künstlicher Intelligenz sind nicht Gegenstand dieses Workshops.

# Einstiegsfragen

---

[Mentimeter](#)

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

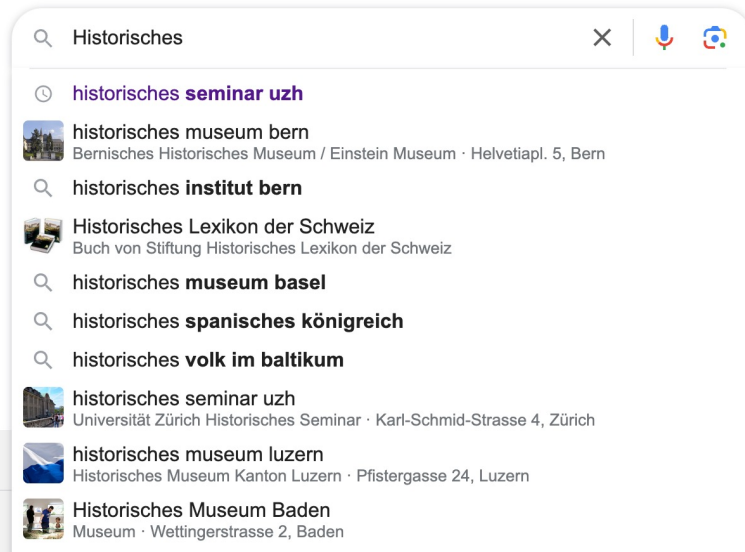


# Grundlagen generativer KI

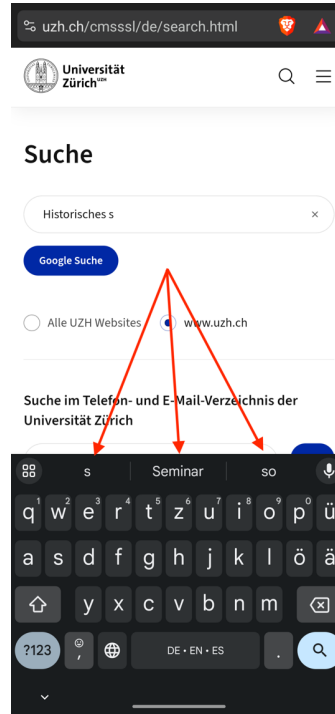
---

- Ein Sprachmodell ist ein Computerprogramm, das aufgrund einer Benutzereingabe (Prompt) versucht, die wahrscheinlichste Folge von Worten vorherzusagen. Es ist ein **Textgenerator**.
- Wir kennen das Prinzip von **Textgeneratoren** bereits von ...

# der Autovervollständigung bei Google



# und der Smartphonetastatur





# Nun halten dialogische Textgeneratoren Einzug in

---

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

# Wie funktioniert ein dialogischer Textgenerator?

---

- Erklärt am Beispiel GPT3 (Brown u. a. 2020), das erste Sprachmodell von ChatGPT von 2020.

# Der Textgenerator wird trainiert



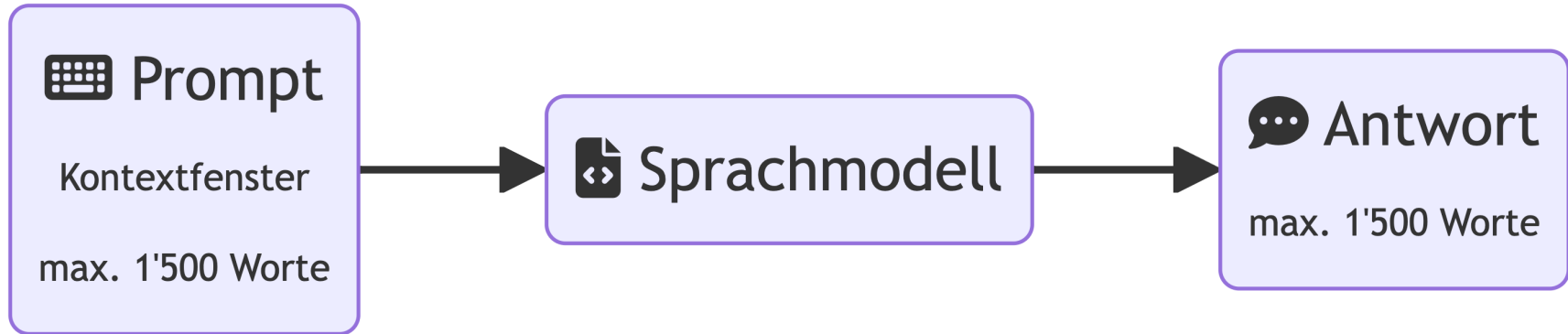
Mit GPT wird das Internet in eine verlustbehaftete mathematische Repräsentation “komprimiert”. (Siehe Chiang 2023)

# Der Textgenerator wird trainiert

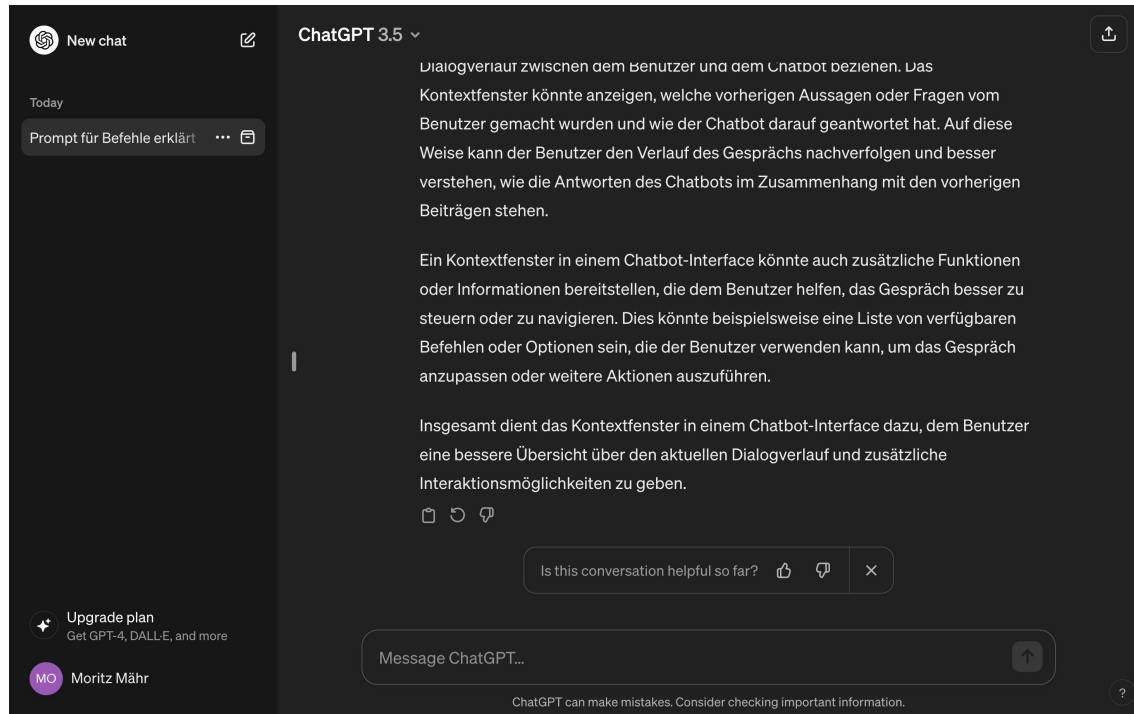
Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

**Table 2.2: Datasets used to train GPT-3.** “Weight in training mix” refers to the fraction of examples during training that are drawn from a given dataset, which we intentionally do not make proportional to the size of the dataset. As a result, when we train for 300 billion tokens, some datasets are seen up to 3.4 times during training while other datasets are seen less than once.

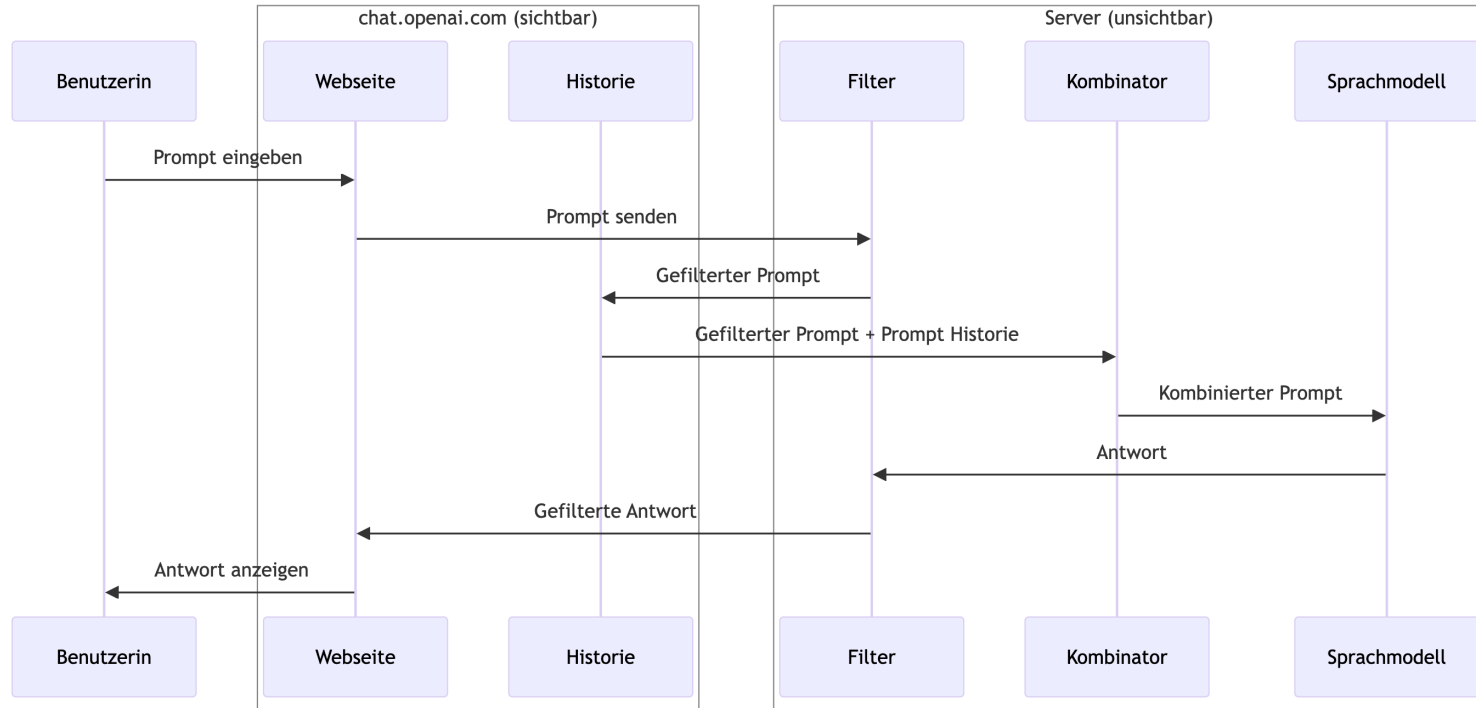
# Der Textgenerator wird ausgeführt



# Der Textgenerator im Internet



# Der Textgenerator im Internet



# Der Textgenerator im Internet

The image shows a screenshot of the ChatGPT web interface. On the left is a sidebar with a 'New chat' button and a 'Today' section containing a prompt: 'Prompt für Befehle erklärt'. Below this is an 'Upgrade plan' section and a user profile for 'Moritz Mähr'. The main chat area is titled 'ChatGPT 3.5' and contains a text block highlighted with a yellow border. The text in the highlighted block discusses context windows in chatbot interfaces. At the bottom of the chat area is a text input field with the placeholder 'Message ChatGPT...' and a 'Send' button. A feedback prompt 'Is this conversation helpful so far?' is visible above the input field. A small disclaimer at the bottom of the chat area reads 'ChatGPT can make mistakes. Consider checking important information.'

**Historie**

Dialogverlauf zwischen dem Benutzer und dem Chatbot beziehen. Das Kontextfenster könnte anzeigen, welche vorherigen Aussagen oder Fragen vom Benutzer gemacht wurden und wie der Chatbot darauf geantwortet hat. Auf diese Weise kann der Benutzer den Verlauf des Gesprächs nachverfolgen und besser verstehen, wie die Antworten des Chatbots im Zusammenhang mit den vorherigen Beiträgen stehen.

Ein Kontextfenster in einem Chatbot-Interface könnte auch zusätzliche Funktionen oder Informationen bereitstellen, die dem Benutzer helfen, das Gespräch besser zu steuern oder zu navigieren. Dies könnte beispielsweise eine Liste von verfügbaren Befehlen oder Optionen sein, die der Benutzer verwenden kann, um das Gespräch anzupassen oder weitere Aktionen auszuführen.

Insgesamt dient das Kontextfenster in einem Chatbot-Interface dazu, dem Benutzer eine bessere Übersicht über den aktuellen Dialogverlauf und zusätzliche Interaktionsmöglichkeiten zu geben.

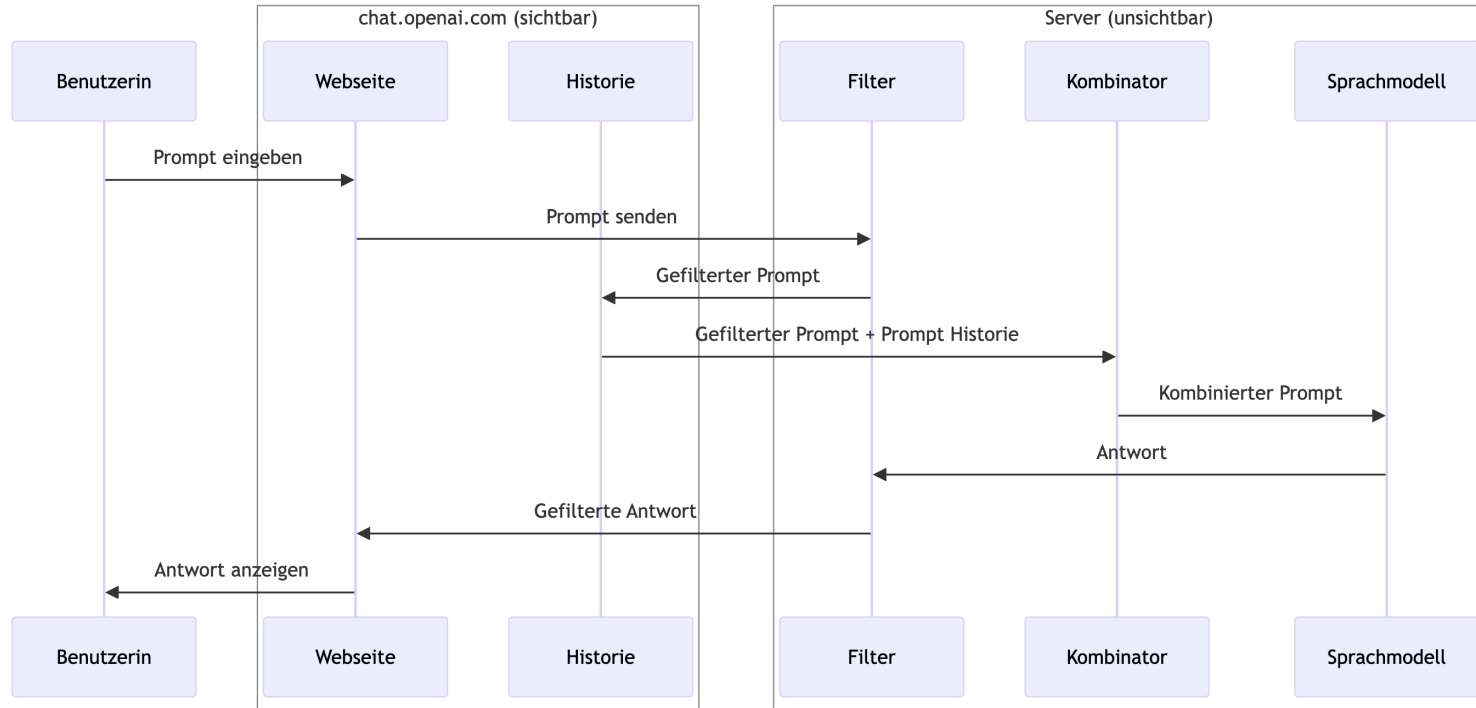
**Prompt**

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.



# Der Textgenerator im Internet



# Wie können wir mit Textgeneratoren 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 ist ein weisser, heterosexueller Cis-Mittzwanziger aus San Francisco.
- Sehr fleissig und schnell, aber nicht sehr tiefgründig oder kritisch
- Hat das gesamte englischsprachige Internet gelesen, insbesondere die englische Wikipedia, hat aber ein sehr schlechtes Kurzzeitgedächtnis.
- Beherrscht viele Sprachen auf Lehrbuchniveau, übersetzt aber oft wörtlich aus dem Englischen.
- Kenntnisse ausschliesslich aus Texten; hat keine persönlichen Erfahrungen mit der Welt und kann nicht zwischen Fiktion und Realität unterscheiden.
- Fokus auf angloamerikanische Popkultur, wenig Wissen über andere Kulturen
- Beherrscht Python und andere Programmiersprachen

# Was kann Jason?

---

- Kurze und mittellange Texte schreiben (E-Mails, Tweets, Blogposts, Artikel)
- Texte übersetzen
- Texte analysieren und zusammenfassen (Gliederung, Schlagwörter, Abstracts etc.)
- Kurze Gespräche führen, Ideen entwickeln, Konzepte erläutern
- (Oberflächliche) Recherchen durchführen
- Programmcode erzeugen und erklären
- Daten aufbereiten, analysieren, erklären und visualisieren

# Was kann Jason nicht?

---

- Einem Gespräch über lange Strecken folgen (Kontextverlust)
- Lange Texte schreiben (Kontextverlust)
- Fiktion von Wirklichkeit unterscheiden (Brain in a vat & Chinese room argument)
- Dreisätze lösen
- Vorhersagen 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: Verwenden Sie klare Trennzeichen “““Antwort hier”“” 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 Richtigkeit.

# Gruppen bilden

---

[Mentimeter](#)

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





# Prompting Guides

---

- <https://www.promptingguide.ai/>
- <https://github.com/brexhq/prompt-engineering>
- <https://platform.openai.com/docs/guides/prompt-engineering/strategy-write-clear-instructions>
- <https://docs.anthropic.com/claude/docs/intro-to-prompting>
- <https://support.google.com/a/users/answer/14200040?hl=en>

# Bibliographie

---

- Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, u. a. 2020. «Language Models Are Few-Shot Learners». arXiv. <http://arxiv.org/abs/2005.14165>.
- Chiang, Ted. 2023. «ChatGPT Is a Blurry JPEG of the Web». *The New Yorker*, 9. Februar 2023. <https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web>.