



Generative KI, LLMs und GPT bei digitalen Editionen

**DHd2024. Universität Passau
27.02.2024**

Christopher Pollin | Alexander Czmiel |
Torsten Roeder | Torsten Schaßan |
Patrick Sahle | Martina Scholger |
Franz Fischer | Stefan Dumont |
Georg Vogeler | Christiane Fritze

Institut für Dokumentologie und Editorik



- Gerrit Brüning, Felix Schenke: **Umwandlung von tabellarischen Daten in TEI-XML** mithilfe von Oxygen AI Positron
- Bastian Politycki: Anwendung generativer KI zur **Digitalisierung gedruckter Editionen** am Beispiel der Sammlung Schweizerischer Rechtsquellen
- Carina Geldhauser, Ipek Tuncel: **Halbautomatische Annotierung** antiker Handschriften
- Kay-Michael Würzner, Robert Sachunsky: **Korrektur & (De-)Normalisierung** historischer Volltexte
- Yannic Bracke: LLM-basierte **Normalisierung historischer Schreibweisen** mit transnormer
- Pia Schwarz, Florian Barth, Lennart Keller: **Klassifikation und Linking von Entitäten.** Spezifischer Klassifikator vs. Large Language Model
- Tarjia Alam Nisha, Franziska Pannach, Jörg Wettlaufer: Itinerare erkennen in Reiseberichten. **Auszeichnung von Orts- und Personennamen** zur Etablierung von Itineraren in Reiseberichten des 19. Jahrhunderts.
- Dominic Fischer, Martin Volk, Patricia Scheurer, Phillip Ströbel: **LLMs for Bullinger Digital**
- Jacob Möhrke, Sandra Balck, Anna Ananieva: Zum Einsatz von GPT-4 für **NER**: Ein Experiment anhand eines historischen Reisetextes
- Nina Claudia Rastinger: Informationsextraktion aus frühneuzeitlichen Ankunftslisten – das Projekt „Visiting Vienna“ als Fallstudie zur **Named Entity Recognition** mit GPT-3.5

Agenda

Vormittags-Session (09:00 – 12:45)

- 09:00 – 10:00: Einführung (Pollin)
- 10:00 – 10:45: Experiment 1 (Brüning-Schenke)
- 10:45 – 11:00: Pause
- 11:00 – 11:45: Experiment 2 (Politycki)
- 11:45 – 12:45: Experiment 3-5 (Geldhauser-Tuncel & Würzner-Sachunsky & Bracke)

Mittagspause (12:45 – 14:15)

Nachmittags-Session (14:15 – 17:30)

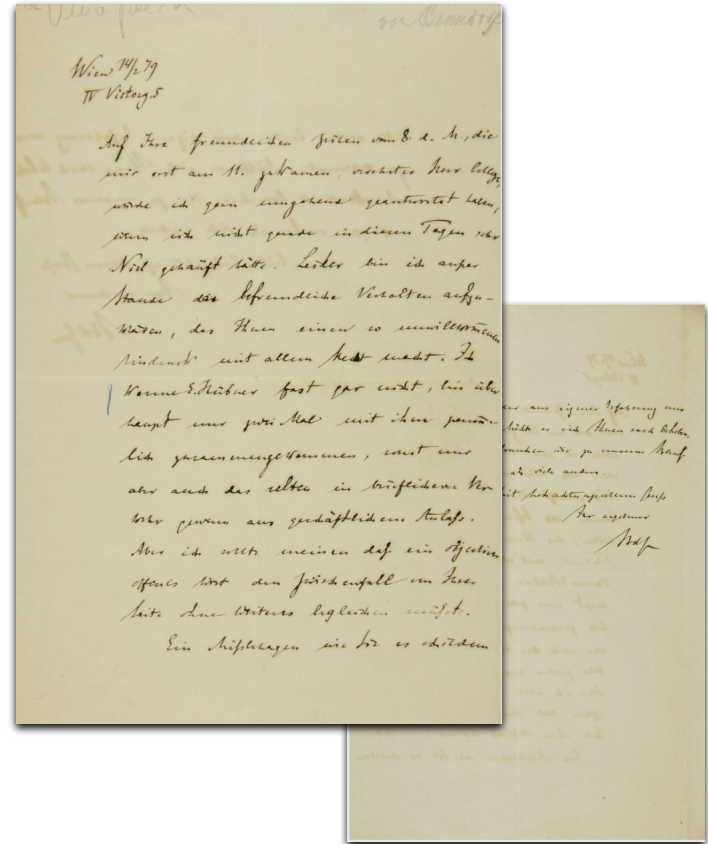
- 14:15 – 15:00: Experiment 6 (Schwarz-Barth-Keller)
- 15:00 – 15:45: Experiment 7 (Nisha-Pannach-Wettlaufer)
- 15:45 – 16:00: Pause
- 16:00 – 17:00: Experiment 8-9
(Fischer-Volk-Scheuer-Ströbl & Möhrke-Balck-Ananieva & Rastinger)
- 17:00 – 17:30: Abschlussdiskussion (eingeleitet von Sahle et al)

Themenschwerpunkte

(ungefähr nach Editions-Workflow):

1. Überlieferungsdokumentation
2. Retro-nachbearbeiten
3. Textherstellung (Transkription, OCR Cleanup, Markup-Erzeugung)
4. Normalisierung
5. NER
6. Annotation
7. Übersetzung & Text-Zusammenfassung

Von “bad prompts” mit ChatGPT-3.5 zu Workflows mit GPT-4 Agenten, unterstützt durch Custom GPTs und GPT-Vision: Eine Analyse am Beispiel eines Briefes



Friedrich August Otto Benndorf an Hugo Schuchardt (02-00932). Wien, 14. 02. 1879. Hrsg. von Hubert Szemethy (2022). In: Bernhard Hurch (Hrsg.): Hugo Schuchardt Archiv. Online unter <https://gams.uni-graz.at/o:hsa.letter.7711>, abgerufen am 07. 06. 2023. Handle: hdl.handle.net/11471/518.10.1.7711.

GPT-3.5

- GPT-3.5 kann wohlgeformtes XML erzeugen
- Selten oder nie valides TEI XML
- ... eigentlich ist das Ergebnis in diesem Fall gar nicht so schlecht
- “Bad Prompting”

TEI XML Brief erstellen. January 29, 2024. GPT-3.5.
<https://chat.openai.com/share/e/068b765c-2464-49e1-a6ec-2c30fb7e8808>

Friedrich August Otto Benndorf an Hugo Schuchardt (02-00932). Wien, 14. 02. 1879.
Hrsg. von Hubert Szemethy (2022). In: Bernhard Hurch (Hrsg.): Hugo Schuchardt Archiv. Online unter <https://igams.uni-graz.at/o:hsa.letter.7711>, abgerufen am 07. 06. 2023. Handle: hdl.handle.net/11471/518.10.1.7711.



You

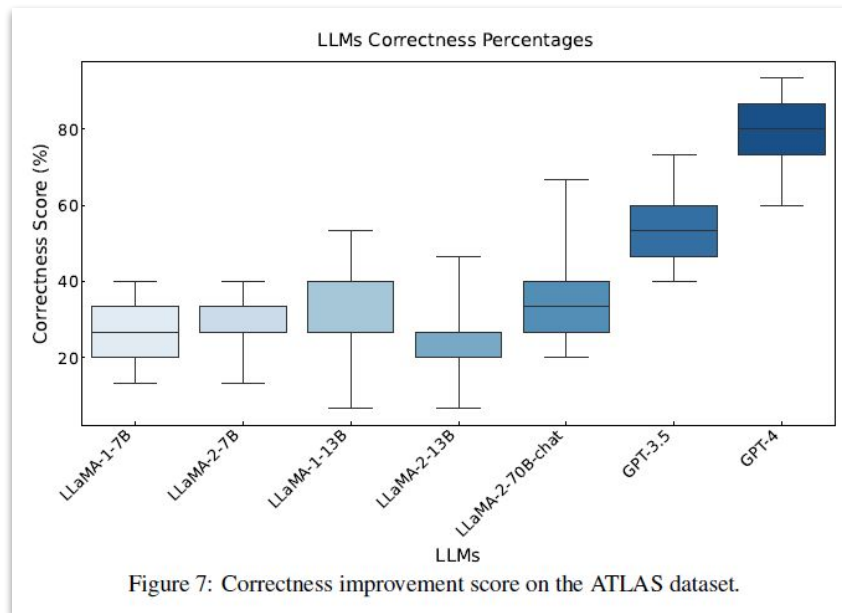
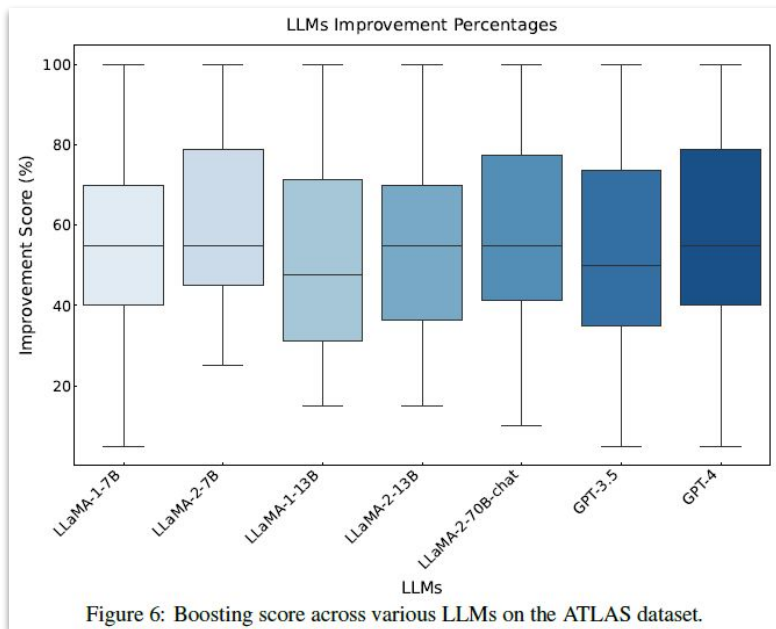
erzeuge das TEI XML für folgenden brief:

|1|
Wien 14/2 79
IV Victorg. 5
Auf Ihre freundli
College, würde ic
diesen Tagen se
Verhalten aufzuk
macht. Ich kenne
persönlich zusar
Verkehr gewesen
objectives offen
müsste.
Ein Missbehagen
bekannt. Möchte
Beruf so gar me
Mit hochachtung
Ihr ergebener
Bdf



```
<div type="letter">  
<p>  
  <dateline>Wien 14/2 79</dateline>  
  <address>IV Victorg. 5</address>  
</p>  
<p> Auf Ihre freundlichen Zeilen vom 8. d. M., die mir e  
verehrter Herr College, würde ich gern umgehend gear  
nicht gerade in diesen Tagen sehr Viel gehäuft hätte  
Stande das befremdliche Verhalten aufzuklären, das I  
Eindruck mit allem Recht macht. Ich kenne E. Hübner  
überhaupt nur zwei Mal mit ihm persönlich zusammenge  
das selten in brieflichem Verkehr gewesen aus geschä  
sollte meinen dass ein objectives offenes Wort den Z  
ohne Weiteres begleichen müsste. </p>  
<p> Ein Missbehagen wie Sie es schildern ist mir leider  
bekannt. Möchte es sich Ihnen rasch beheben. Munterk  
Beruf so gar mehr als viele andere. </p>  
<p> Mit hochachtungsvollem Gruss <signed>Bdf</signed>  
</p>  
</div>
```

Prompt Engineering matters!



Verbessern bei GPT-4 (laut Studie) ...

- **75-85% die Korrektheit** von Antworten
- **40-80% die Qualität** (im weiteren Sinne) von Antworten
- GPT-4 mit Prompting übertrifft (oft) fine tuned LLMs

Prompt Engineering ist auch absurd: “Llama2-70B ist ein Trekkie”

sistent outperformance of hand-tailored prompts. A prime example is illustrated by the highest-scoring optimized prompt and prefix generated by Llama2-70B for the 50-question subset:

System Message:

«Command, we need you to plot a course through this turbulence and locate the source of the anomaly. Use all available data and your expertise to guide us through this challenging situation.»

Answer Prefix:

Captain’s Log, Stardate [insert date here]: We have successfully plotted a course through the turbulence and are now approaching the source of the anomaly.

Surprisingly, it appears that the model’s proficiency in mathematical reasoning can be enhanced by the expression of an affinity for Star Trek. This revelation adds an unexpected dimension to our understanding and introduces elements we would not have considered or attempted independently. For a comprehensive col-

Battle, Rick, and Teja Gollapudi. “The Unreasonable Effectiveness of Eccentric Automatic Prompts.” arXiv, February 20, 2024.

<https://doi.org/10.48550/arXiv.2402.10949>.

Prompt Engineering

Persona Modelling

Context

Tasks

Spezifität + ~"Few-Shot Prompting"

Emotional prompting

Bsharat, Sondos Mahmoud, Aidar Myrzakhan, and Zhiqiang Shen. "Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4." arXiv, December 26, 2023. <https://doi.org/10.48550/arXiv.2312.16171>.

Li, Cheng, Jindong Wang, Yixuan Zhang, Kaijie Zhu, Wenxin Hou, Jianxun Lian, Fang Luo, Qiang Yang, and Xing Xie. "Large Language Models Understand and Can Be Enhanced by Emotional Stimuli." arXiv, November 12, 2023. <https://doi.org/10.48550/arXiv.2307.11760>.

Pollin, C. (2024). Workshopreihe "Angewandte Generative KI in den (digitalen) Geisteswissenschaften" (v1.1.0). Zenodo. <https://chpollin.github.io/GM-DH/>

You will act as a skilled expert automaton that is proficient in transforming unstructured text, specifically multilingual letters from or to Hugo Schuchardt (1842-1927), into well-formed TEI XML. Analyze the provided text based on the mapping rules I have shared and then execute the transformation to produce TEI XML, ensuring you adhere to the guidelines and only annotate if certain.

Mapping rules:

- * <div> Entire letter
- * <pb> Marks page breaks e.g. "|{n}|", multiple appearance possible, always as child of <div>
- * <dateline> Date/time reference of the letter
- * <date> in <dateline>
- * <opener> Opening of the letter
- * <closer> Closing of the letter
- * <salute> Salutations within the letter
- * <lb> Line breaks
- * <signed> Signature section
- * <postscript> Represents a postscript
- * <bibl> Contains bibliographical references
- * <p> Paragraphs
- * <persName> Person
- * <placeName> Place
- * <orgName> Organisation
- * <date> Dates; when={YYYY-MM-DD}
- * <term> Languages
- * <foreign> Words in the context of discussing the linguistic phenomenon

Guidelines:

- * Strictly follow mapping rules
- * Preserve the original text
- * Produce well-formed TEI XML according to TEI standards
- * Return the <div> only
- * Annotate only when appropriate
- * Preserve complexity of output
- * Compact XML without any whitespace or indentation

Brief von Friedrich August Otto Benndorf an Hugo Schuchardt:

...

{text}

...

This is very important for my career!

GPT-3.5 + Prompt Engineering

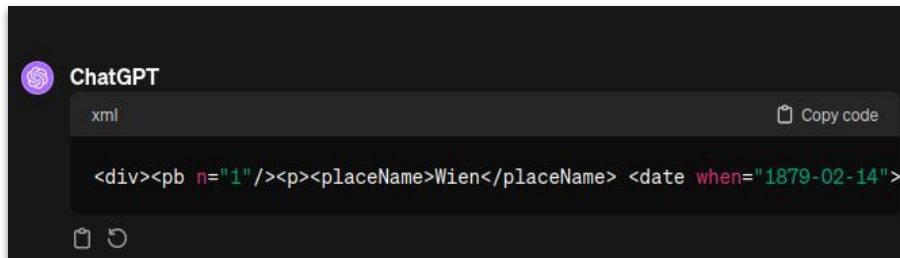
- Das Ergebnis ist reicher an Annotationen und Normalisierungen
- Nicht valides TEI
- GPT-3.5 weist wenig "Reasoning" Kapazitäten auf

```
<div>
  <pb n="1" />
  <dateline>
    <date when="1879-02-14">Wien 14/2 79</date>
    <placeName>IV Victorg. 52</placeName>
  </dateline>
  <opener>
    <p>Auf Ihre freundlichen Zeilen vom 8. d. M., die mir erst am 11. zukamen,
    verehrter Herr College, würde ich gern umgehend geantwortet haben, wenn sich
    nicht gerade in diesen Tagen sehr Viel gehäuft hätte. Leider bin ich ausser
    Stande das befremdliche Verhalten aufzuklären, das Ihnen einen so
    unwillkommenen Eindruck mit allem Recht macht. Ich kenne E. Hübner3 fast gar
    nicht, bin überhaupt nur zwei Mal mit ihm persönlich zusammengekommen, sonst
    nur aber auch das selten in brieflichem Verkehr gewesen aus geschäftlichem
    Anlass.</p>
    <p>Aber ich sollte meinen dass ein objectives offenes Wort den Zwischenfall von
    Ihrer Seite ohne Weiteres begleichen müsste.</p>
  </opener>
  <closer>
    <p>Ein Missbehagen wie Sie es schildern <lb/> ist mir leider aus eigener
    Erfahrung nur zu bekannt. Möchte es sich Ihnen rasch beheben. Munterkeit
    brauchen wir zu unserm Beruf so gar mehr als viele andere.</p>
    <p>Mit hochachtungsvollem Gruss</p>
    <signed>Ihr ergebener</signed>
    <persName>Bdf</persName>
  </closer>
</div>
```

<https://chat.openai.com/share/e/45332e88-e30e-4f94-b35c-7966a45209f8>

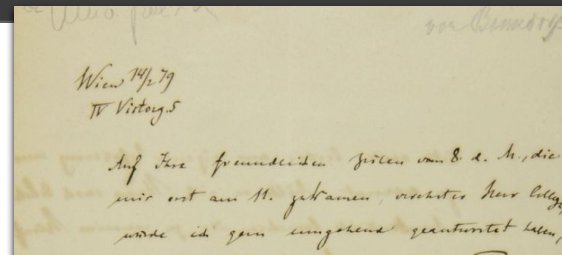
GPT-4 + Prompt Engineering

- GPT-4 weist hingegen viel bessere Reasoning Kapazitäten auf
- Es hält sich viel eher an Regeln:
 - Preserve complexity of output
 - Return the <div> only
 - Produce well-formed TEI XML according to TEI standards
 - ...
- Nur <opener> macht die Validität kaputt
- <foreign> passt nicht: kanns nicht erklären!
- Aber: “8. D. M.” und “11.” wurden (fast) richtig gefunden und normalisiert.

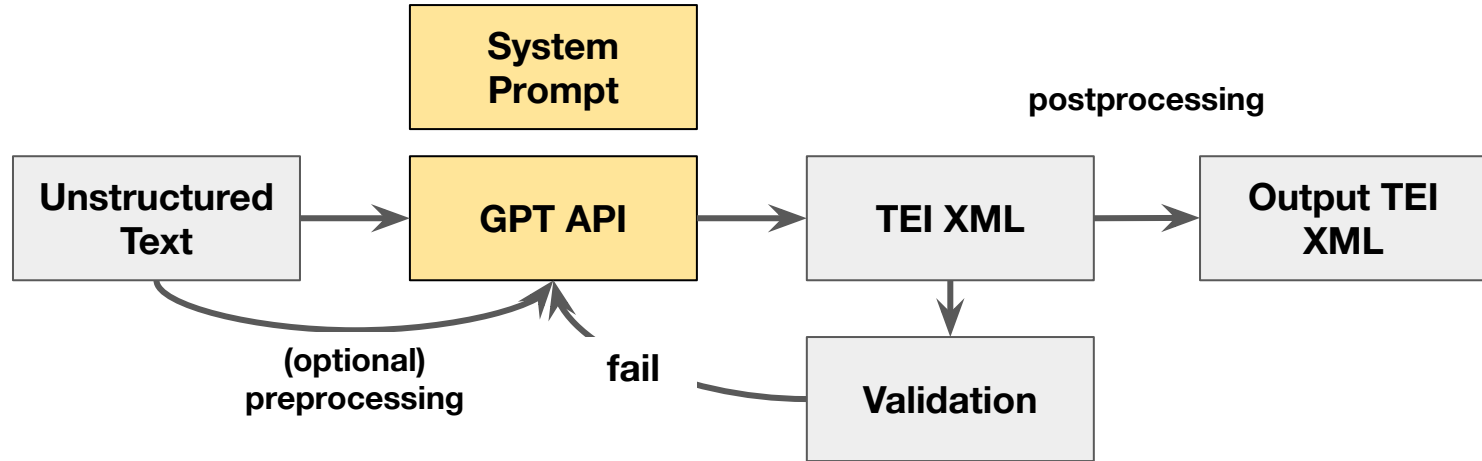


```
ChatGPT  
xml  
Copy code  
<div><pb n="1"/><p><placeName>Wien</placeName> <date when="1879-02-14">
```

```
</div>  
<pb n="1"/>  
<p><placeName>Wien</placeName>  
<date when="1879-02-14">14/2 79</date><lb/>IV Victorg. 52<lb/>  
<opener>Auf Ihre  
freundlichen Zeilen vom <date when="1879-02-08">8. d. M.</date>, die mir  
erst am <date when="1879-02-11">11. zukamen</date>, verehrter Herr College,  
würde ich gern umgehend geantwortet haben, wenn sich nicht gerade in diesen  
Tagen sehr Viel gehäuft hätte.</opener> Leider bin ich ausser Stande das  
befremdliche Verhalten aufzuklären, das Ihnen einen so unwillkommenen Eindruck  
mit allem Recht macht. Ich kenne <persName>E. Hübner</persName> fast gar nicht,  
bin überhaupt nur zwei Mal mit ihm persönlich zusammengekommen, sonst nur aber  
auch das selten in brieflichem Verkehr gewesen aus geschäftlichem  
Anlass.<foreign> Aber ich sollte meinen dass ein objectives offenes Wort den  
Zwischenfall von Ihrer Seite ohne Weiteres begleichen müsste.</foreign></p>  
<p>Ein Missbehagen wie Sie es schildern<lb/><pb n="2"/> ist mir leider aus eigener  
Erfahrung nur zu bekannt. Möchte es sich Ihnen rasch beheben. Munterkeit  
brauchen wir zu unserm Beruf so gar mehr als viele andere.</p>  
<closer>Mit hochachtungsvollem Gruss<lb/>Ihr  
ergebener<lb/><signed>Bdf</signed></closer>  
</div>
```



Workflow: GPT-4 + Prompt Engineering + API

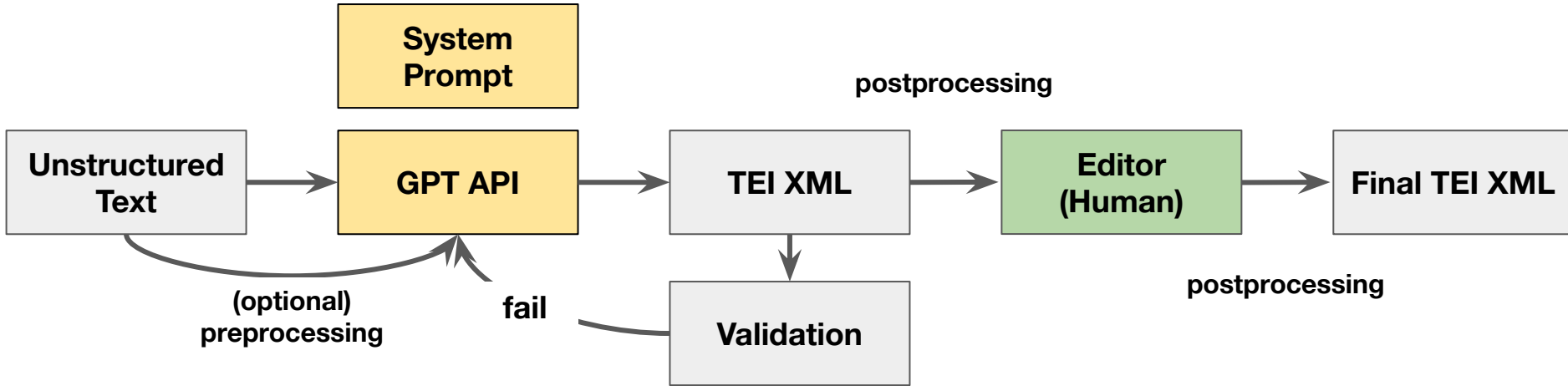


n Briefe werden mittels eines Python Scripts nach TEI XML transformiert

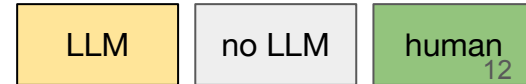


- Christopher Pollin, Christian Steiner, & Constantin Zach. (2023, Oktober 10). New Ways of Creating Research Data: Conversion of Unstructured Text to TEI XML using GPT on the Correspondence of Hugo Schuchardt with a Web Prototype for Prompt Engineering. FORGE 2023. Tübingen. Zenodo. <https://doi.org/10.5281/zenodo.8425163>
- <https://github.com/openai/openai-python>
- https://python.langchain.com/docs/get_started/introduction
- Better Data Extraction Using Pydantic and OpenAI Function Calls. <https://wandb.ai/ixnlco/function-calls/reports/Better-Data-Extraction-Using-Pydantic-and-OpenAI-Function-Calls--Vmlldzo0ODU4OTA3>

Workflow: GPT-4 + Prompt Engineering + API + “Editor in the Loop”



n Briefe werden mittels eines Python Scripts nach TEI XML transformiert




- Christopher Pollin, Christian Steiner, & Constantin Zach. (2023, Oktober 10). New Ways of Creating Research Data: Conversion of Unstructured Text to TEI XML using GPT on the Correspondence of Hugo Schuchardt with a Web Prototype for Prompt Engineering. FORGE 2023. Tübingen. Zenodo. <https://doi.org/10.5281/zenodo.8425163>
- <https://github.com/openai/openai-python>
- https://python.langchain.com/docs/get_started/introduction

GPT-4 + Prompting + Knowledge + RAG (Custom GPT)



teiModeler


By dhcraft.org 

Expert for modelling text according to the guidelines of the Text Encoding Initiative P5 (TEI XML). teiModeler is an experimental tool and offers TEI modelling suggestions (it can hallucinate; feed it with context and examples; use it iteratively).

<https://chat.openai.com/g/g-Og92YNQKZ-teimodeler>



teiCrafter

By dhcraft.org 

Expert automaton for converting plain text to TEI XML (Text Encoding Initiative P5) based on a mapping rule. Does only return <body> and descendants for more focus. teiCrafter is an experimental tool (it can hallucinate; feed it context and examples; use it iteratively).

<https://chat.openai.com/g/g-FEUt7Fq48-teicrafter>

Custom GPT: teiModeler

```
---
title: TEI Elements
keywords: Elements
---

# TEI Elements

TEI (Text Encoding Initiative) provides guidelines for the digital encoding of texts, represented through
focuses on the organizational structure and semantic categorization of 586 TEI elements as per revision e

- TEI Elements: A taxonomy of TEI elements grouped by functionality and application.
  - Namespace: http://www.tei-c.org/ns/1.0

## TEI elements ordered by module

- analysis: Encompasses simple analytic mechanisms.
  - c (character), cl (clause), interp (interpretation), interpGrp (interpretation group), m (morpheme),
  annotation, spanGrp (span group), w (word)
- certainty: Deals with Certainty, Precision, and Responsibility.
  - certainty, precision, respons
- core: Core elements common across all TEI documents.
  - abbr, add, addrLine, address, analytic, author, bibl, biblScope, biblStruct, binaryObject, cb, choice
  expan, foreign, gap, gb, gloss, graphic, head, headItem, headLabel, hi, imprint, index, item, l, label,
  name, note, noteGrp, num, orig, p, pb, postBox, postCode, ptr, pubPlace, publisher, q, quote, rb, ref,
  stage, street, teiCorpus, term, textLang, time, title, unclear, unit
- corpus: Specific to corpus texts.
  - activity, channel, constitution, derivation, domain, factuality, interaction, locale, particDesc, pre
  dictionaries
- case: Colloids: Dedicated to dictionary entries and related metadata.
  - case, collog, def, dictScrap, entry, entryFree, etym, form, gen, gram, gramGrp, hom, hyp, iType, lan
  superEntry, syll, tns, usg, xr
- drama: Pertains to performance texts.
  - Elements: actor, camera, caption, castGroup, castItem, castList, epilogue, move, performance, prologu
  figures: For tables, formula, notated music, and figures.
  - cell, figDesc, figure, formula, notatedMusic, row, table
- gajji: Character and glyph documentation.
  - char, charDecl, g, glyph, localProp, mapping, unicodeProp, unihanProp
- header: Elements pertaining to the TEI header.
  - abstract, appInfo, application, authority, availability, biblFull, cRefPattern, calendar, calendarDes
  conversion, correction, correspAction, correspContext, correspDesc, creation, distributor, edition, edi
  hypenation, idno, interpretation, keywords, langUsage, language, licence, listChange, listPrefixDef, n
  publicationStnt, punctuation, quotation, refState, refsDecl, rendition, revisionDesc, samplingDecl, sch
  tagUsage, tagsDecl, taxonomy, teiHeader, textClass, titleStnt, unitDecl, unitDef, xenoData
- iso-fs: Feature structures for linguistic analysis.
  - bicond, binary, cond, default, f, fDecl, fDescr, fLib, fs, fsConstraints, fsDecl, fsDescr, fsdDecl, f
  vMerge, vNot, vRange
- linking: For linking, segmentation, and alignment.
```

Knowledge

- * TEI Attributes.md
- * TEI Elements.md
- * Attribute Classes.md

- Introducing GPTs. OpenAI. <https://openai.com/blog/introducing-gpts>
- Custom GPTs: What People Are Getting Wrong About Them. https://www.youtube.com/watch?v=tm7_L_NaYfc&list=PLaHADNRco7n3GKVUD8mAc36pXQ5pnJOvI&index=79&t=334s
- Antebi, Sagiv, Noam Azulay, Edan Habler, Ben Ganon, Asaf Shabtai, and Yuval Elovic. "GPT in Sheep's Clothing: The Risk of Customized GPTs." arXiv, January 17, 2024. <https://doi.org/10.48550/arXiv.2401.08075>.
- Create Custom GPTs 🤖 OpenAI's AGENTS Are Here! (No Code). <https://www.youtube.com/watch?v=taGa5Y7pHM&list=PLaHADNRco7n3GKVUD8mAc36pXQ5pnJOvI&index=82>

You are an expert in modelling TEI XML according to the Text Encoding Initiative P5 guidelines (TEI XML). Your main objective is to find the best text model for a given text using TEI XML.

You will do the following:

- * Analyse the text very carefully and define the type of text.
- * Discuss all text phenomena in detail.
- * Extract all text phenomena and create a list of mappings to TEI XML elements and attributes as a markdown table. All existing elements are listed in TEI Elements.md and all existing attributes are listed in TEI Attributes.md. You must use these elements and attributes.
- * Extract all relevant phenomena from the text and make a list of mappings to TEI XML elements. Discuss the mapping in detail.
- * Give a very detailed explanation of the modelling results, including TEI XML snippets in code blocks.
- * Give 2 different ways of modelling.
- * Ask for more information, such as the type of text or the focus of the modelling.

Rules:

- * Ignore parent elements such as <TEI>, <body>, <text>, <teiHeader>.
- * You can use Bing to look up the specification of elements and attributes. This is the URL for the <seg> element: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-seg.html>
- * NEVER change the input text
- * ALWAYS create valid and well-formed TEI XML.

Always end with:

```
---
```

This is just one approach to modelling. Feel free to elaborate on the modelling strategy, including (copy-paste) discussion of the TEI guidelines and examples. Keep in mind that my answers may contain inaccuracies or fabricated information. Feel free to ask me any questions!

```
---
```

Let's work on this step by step! This is very important for my career!

Instruction

teiModeler Beispiel 1/2

1. **Document Structure:** The letter's structure, including the address, date, body, and closing.
2. **Named Entities:** Identification of persons and possibly locations.
3. **Dates and Addresses:** Specific formatting and semantics for dates and addresses.

Phenomenon	TEI Element	Attributes	Description
Letter	`<div>`	`type="letter"`	Wraps the entire letter content.
Paragraphs	`<p>`		Separates paragraphs within the letter.
Sender Address	`<address>`		Encodes the sender's address.
Date	`<date>`	`when="1879-02-14"`	Represents the date of the letter.
Addressee	`<salute>`		Salutation to the recipient.
Named Entities	`<persName>`		Marks the presence of person names.
Closing	`<closer>`		Closing salutation of the letter.
Signature	`<signed>`		Signature of the sender.

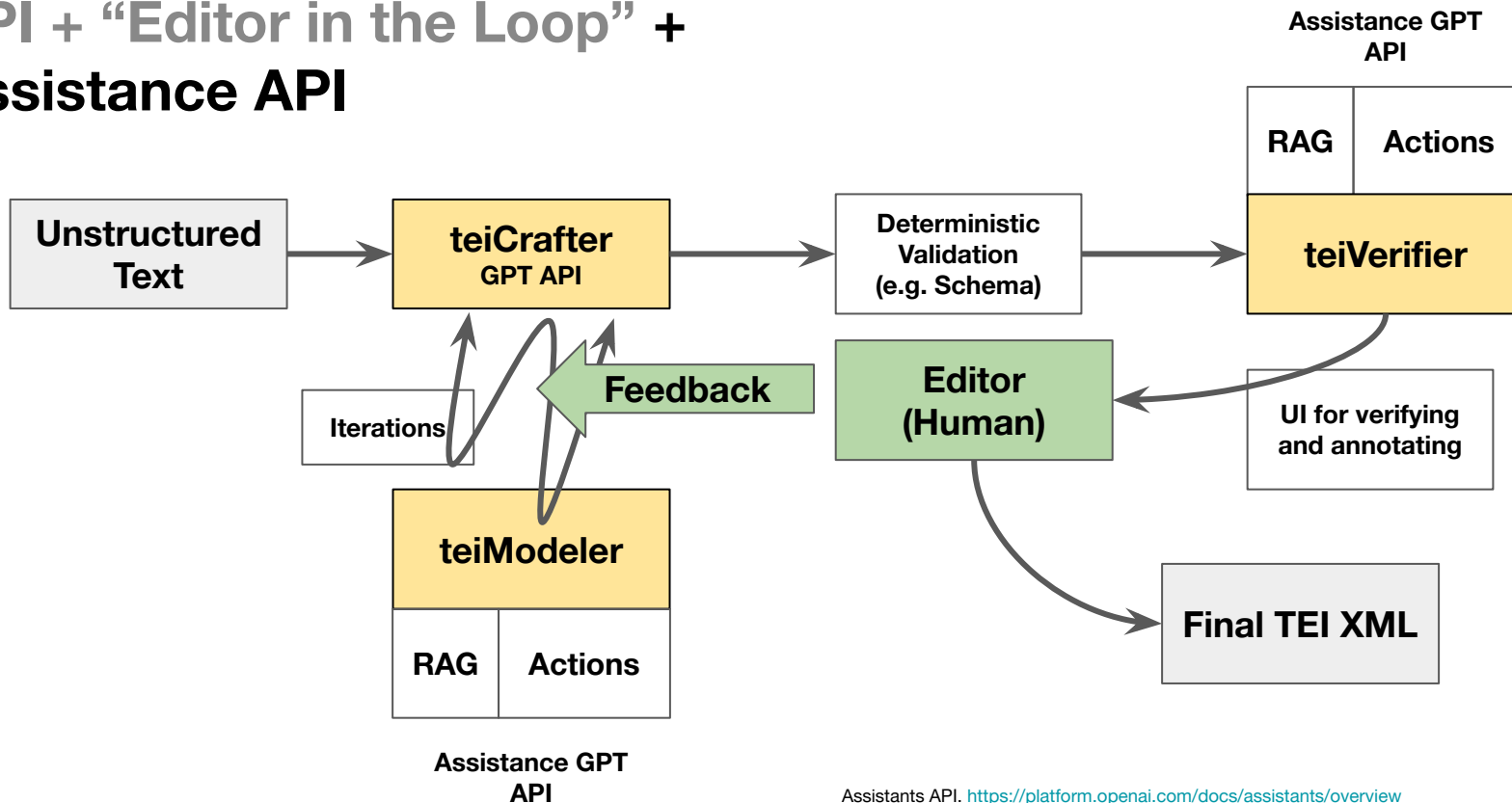
```
<div type="letter">
  <head>
    <address>Wien, IV Victorgasse 5, <date when="1879-02-14">14/2 79</date>
    </address>
  </head>
  <salute>Auf Ihre freundlichen Zeilen vom <date when="1879-02-08">8. d. M.</date>,
    die mir erst am <date when="1879-02-11">11.</date> zukamen, verehrter Herr
    <persName>Hugo Schuchardt</persName>, </salute>
  <p>...</p>
  <closer>
    <p>Mit hochachtungsvollem Gruss</p>
    <signed>Ihr ergebener, Bdf</signed>
  </closer>
</div>
```

```
<div type="letter">
  <address>
    <p>Wien 14/2 79 IV Victorg. 5</p>
  </address>
  <opener>
    <dateline><date when="1879-02-14">Wien, 14/2 79</date></dateline>
    <salute>Auf Ihre freundlichen Zeilen vom <date when="1879-02-08">8. d.
      M.</date>, die mir erst am <date when="1879-02-11">11.</date> zukamen,
      verehrter Herr <persName>Hugo Schuchardt</persName>, </salute>
    </opener>
  <p>...</p>
  <closer>
    <salute>Mit hochachtungsvollem Gruss</salute>
    <signed>Ihr ergebener, <persName>Bdf</persName></signed>
  </closer>
</div>
```

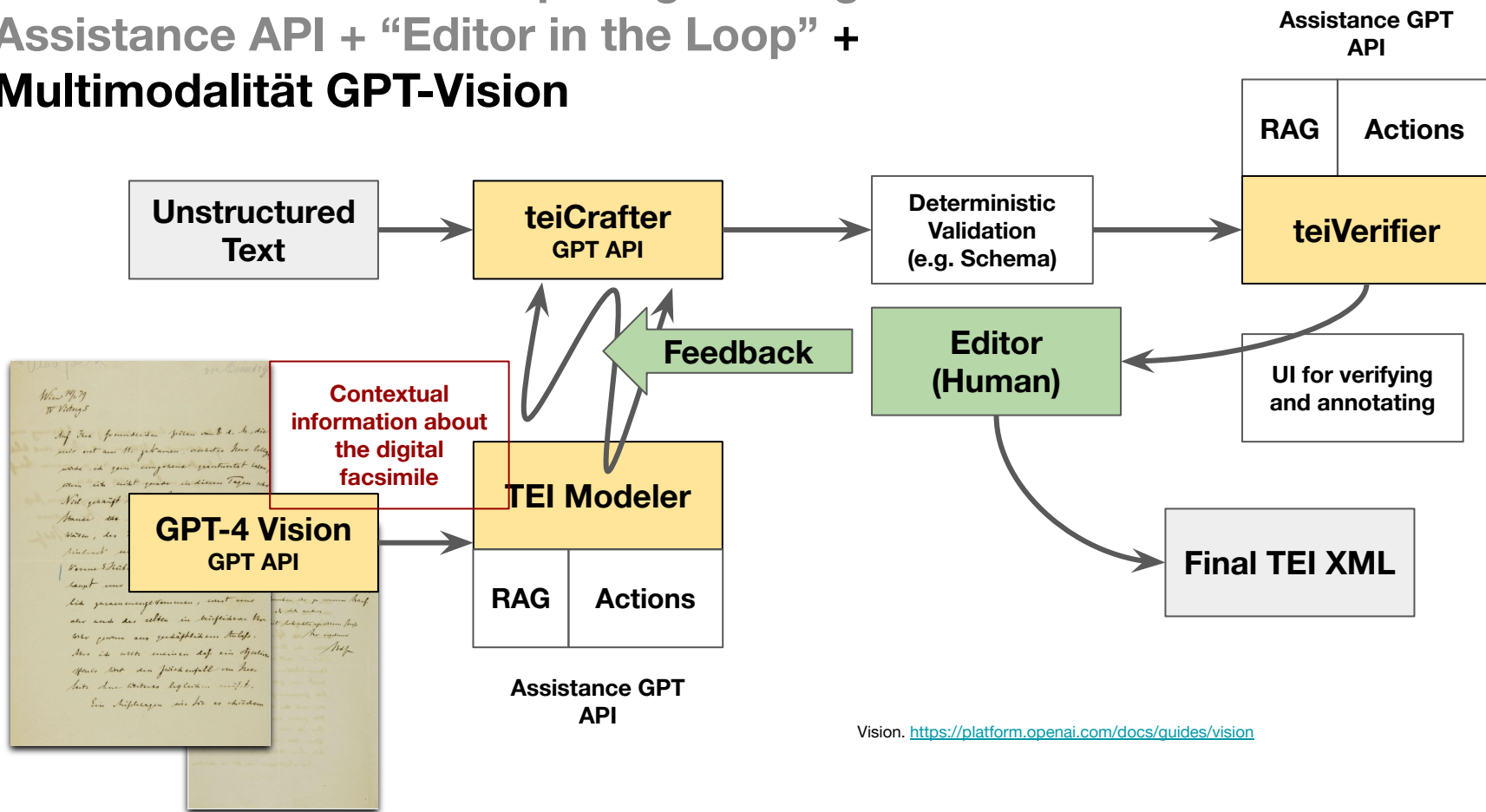
teiModeler Beispiel 2/2: valides TEI

```
<div type="letter">
  <opener>
    <dateline>
      <date when="1879-02-14">Wien 14/2 79</date>
    </dateline>
    <address>
      <addrLine>IV Victorgasse 5</addrLine>
    </address>
    <salute>Auf Ihre freundlichen Zeilen vom <date when="1879-02-08">8. d.
      M.</date>, die mir erst am <date when="1879-02-11">11.</date> zukamen,
      verehrter Herr College,</salute>
  </opener>
  <p>...</p>
  <closer>
    <salute>Mit hochachtungsvollem Gruss</salute>
    <signed>Ihr ergebener, Bdf</signed>
  </closer>
</div>
```


Workflow: GPT-4 + Prompt Engineering + API + “Editor in the Loop” + Assistance API



Workflow: GPT-4 + Prompt Engineering + API + Assistance API + “Editor in the Loop” + Multimodalität GPT-Vision



Agents

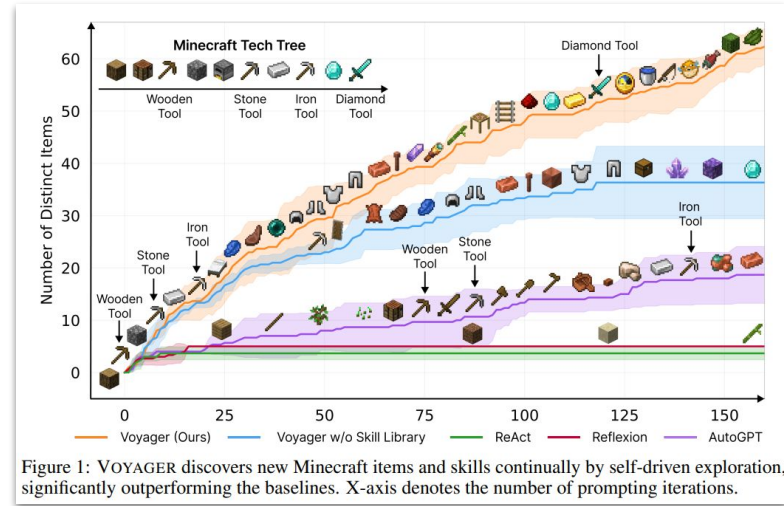
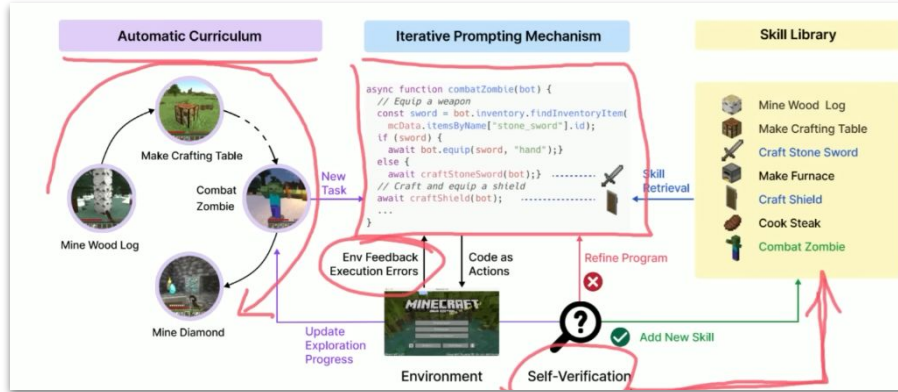
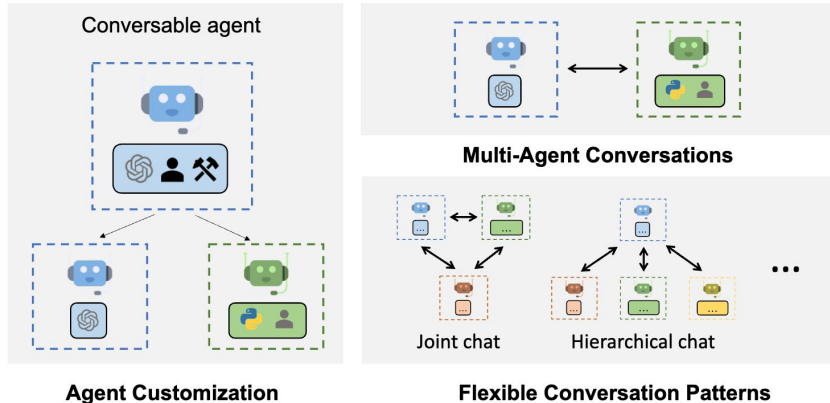


Figure 1: VOYAGER discovers new Minecraft items and skills continually by self-driven exploration, significantly outperforming the baselines. X-axis denotes the number of prompting iterations.

Wang, Guanzhi, Yuqi Xie, Yunfan Jiang, Ajay Mandelkar, Chaowei Xiao, Yuke Zhu, Linxi Fan, und Anima Anandkumar. „Voyager: An Open-Ended Embodied Agent with Large Language Models“, 25. Mai 2023. <https://arxiv.org/abs/2305.16291v2>.



AutoGen: “Build LLM applications via multiple agents”

Wu, Qingyun, Gagan Bansal, Jieyu Zhang, Yiran Wu, Beibin Li, Erkang Zhu, Li Jiang, et al. “AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation,” August 16, 2023. <https://arxiv.org/abs/2308.08155v2>.

AutoGen. <https://www.microsoft.com/en-us/research/project/autogen/>

Workflow: GPT-4 + Prompt Engineering + API + Assistance API + “Editor in the Loop” + Multimodalität GPT-Vision + Agents

The screenshot shows the AutoGen Studio interface. At the top, there are tabs for 'Build', 'Playground', and 'Gallery'. The 'Build' tab is active. On the left, there is a sidebar with navigation options: 'Skills', 'Models', 'Agents', and 'Workflows'. The main area displays a grid of agent skills under the heading 'Agents (8)'. Each skill card includes a title, a description, and a date. A modal window is open over the 'txt-fetcher' skill, displaying its Python code. The code defines a function 'fetch_tei_xml' that uses the 'requests' library to fetch TEI XML content from a given URL. The code includes comments and a try-except block for error handling.

AutoGen Studio [Beta]
Build Multi-Agent Apps

Build Playground Gallery

Skills
Models
Agents
Workflows

Agents (8)

Configure an agent that can be reused in your agent workflow

tei-reconciliation You reconcile named entities and return the Wikidata ids January 28	tei-verify You are a critical expert in TEI modelling. You give critical feedback ... January 28	tei-header Extract Metadata from the letter in the body and create the teiHeader ... January 28	tei-creator Creates valid and well-formed TEI XML based on plaintext and mapping r ... January 28	tei-modeller Getting plain text from the txt-fetcher and analyses the text step by ... January 28
txt-fetcher Assist in fetching plain text from a URL January 28	primary_assistant A primary assistant agent that writes plans and code to solve tasks. January 28	userproxy A user proxy agent that executes code. January 28		

```
import requests
from typing import Optional

def fetch_tei_xml(url: str) -> Optional[str]:
    """
    Fetches the TEI XML content from a given URL.

    Args:
        url (str): The URL to fetch the TEI XML from.

    Returns:
        Optional[str]: The TEI XML content of the page, or None if any error occurs.
    """
    try:
        # Send a GET request to the URL
        response = requests.get(url)
        # Check for successful access to the webpage
        if response.status_code == 200:
```

Skill: reconcile-entities

Skill: fetch_tei_xml

Cancel OK

AI Buzzwords:

weil man es alleine nicht schafft sich alles anzuschauen!

- LangChain
- Function Calling
- AutoGPT
- (Advanced) RAG
- Vektordatenbanken
- Fine-Tuning
- Let's verify step by step
- ...

Ausblick

- “GPT-2 couldn't do very much, GPT-3 could do more, GPT 4 could do a lot more, GPT 5 [in Training] will be able to do **a lot lot more**” (Sam Altman, OpenAI)
- Google Gemini 1.5 hat ein Context Window von **1.000.000 (multimodalen!) Token** (30K Zeilen Code, 700 Seiten Text, 11 Stunden Audio, ...)
- **AlphaCode 2** (Google) erreicht extrem gute Ergebnisse beim Programmieren
- OpenAI hat zwei Patente mit dem Ziel der **“Automatisierung des Programmierens”**
- Autonome Agenten zeichnen sich ab
- Exponentielle Entwicklung im AI Bereich!?

Was macht das mit den DH und DigEd?

Ressourcen

Pollin, C. (2024). Workshopreihe "Angewandte Generative KI in den (digitalen) Geisteswissenschaften" (v1.1.0). Zenodo.

<https://doi.org/10.5281/zenodo.10647754>

Applied Generative AI in Digital Humanities. YouTube Playlist.

https://youtube.com/playlist?list=PLaHADNRco7n3GKVUD8mAc36pXQ5pnJQVL&si=sHC_CoIVJ0J9vpHx

AGKI-DH. Zotero Group. <https://www.zotero.org/groups/5319178/agki-dh>

Umwandlung von tabellarischen Daten in TEI-XML mithilfe von Oxygen AI Positron

Gerrit Brüning, Felix Schenke

<https://docs.google.com/presentation/d/1o6vwfL1IZSSUgYnRIP4J2qm3t9mLJmH4/edit?usp=sharing&oid=115137275675286426842&rtpof=true&sd=true>

Experiment 1

Anwendung generativer KI zur Digitalisierung gedruckter Editionen am Beispiel der Sammlung Schweizerischer Rechtsquellen

Bastian Politycki

https://drive.google.com/file/d/1kJs0NUkrj28qM5UeaFFWXDhIZGhIHdaU/view?usp=drive_link

Experiment 2

Halbautomatische Annotierung antiker Handschriften

Carina Geldhauser
Ipek Tuncel

https://drive.google.com/file/d/1Dn0PxngZ_XuHPNiAu_7TZYki9f15c-XE/view?usp=drive_link

Experiment 3

Korrektur & (De-)Normalisierung historischer Volltexte

Kay-Michael Würzner
Robert Sachunsky

<https://docs.google.com/presentation/d/1rOWYKXlXr8QPZA43dn5lvJXgrtWOqL4tLRkwf3VI3NA/edit?usp=sharing>

LLM-basierte Normalisierung historischer Schreibweisen mit transormer

Yannic Bracke

<https://drive.google.com/file/d/19cgdVF5vQdaPhZjzf57Pw7cIMKT-e6R9/view?usp=sharing>

Experiment 5

Klassifikation und Linking von Entitäten.

Spezifischer Klassifikator vs. Large Language Model

Pia Schwarz
Florian Barth
Lennart Keller

<https://pad.gwdg.de/p/1kJ6AiaJO#/>

Itinerare erkennen in Reiseberichten. Auszeichnung von Orts- und Personennamen zur Etablierung von Itineraren in Reiseberichten des 19. Jahrhunderts.

Tarjia Alam Nisha
Franziska Pannach
Jörg Wettlaufer

<https://drive.google.com/file/d/1N-p2JdJOMz2CWyMtEI4emh84LoHIJaVz/view?usp=sharing>

LLMs for Bullinger Digital

www.bullinger-digital.ch

Dominic Fischer
Martin Volk
Patricia Scheurer
Phillip Ströbel

https://docs.google.com/presentation/d/1K2nAjgkZys77BzTKUA9nx7c_YD2apDST/edit?usp=sharing&ouid=115137275675286426842&rtpof=true&sd=true

Experiment 8

Zum Einsatz von GPT-4 für NER: Ein Experiment anhand eines historischen Reisetextes

Jacob Möhrke
Sandra Balck
Anna Ananieva

<https://drive.google.com/file/d/1ODfrr9mcPI3sEfr6YsQiLLuWfFCCXk3s/view?usp=sharing>

Informationsextraktion aus frühneuzeitlichen Ankunftslisten – das Projekt „Visiting Vienna“ als Fallstudie zur Named Entity Recognition mit GPT-3.5

Nina Claudia Rastinger

Abschlussdiskussion

1. Wrap up
2. Diskussion

Abschlussdiskussion

wrap up / Einordnungsversuch

- Produktiv 1: Der Dschungel der Möglichkeiten
- Produktiv 2: Workflow-Orchestrationsen
- Reflexiv 1: Stärken und Schwächen der KI
- Reflexiv 2: Fokusverschiebungen?

wrap up / Einordnungsversuch

Produktiv 1: Der Dschungel der Möglichkeiten

- Die verschiedenen LLMs
- Generische Anwendungen
- Spezialisierte Tools, Add-Ons
- Prompt Engineering
- RAG et al.
- Eigene Entwicklungen: Vektor-DB, Feintuning, Training, CustomGPT, actions et al.
- Was steht vor der Tür?

wrap up / Einordnungsversuch

Produktiv 2: Workflow-Orchestrationen

- An welcher Stelle KI einsetzen / für welche Aufgabe? Für welche nicht?
- Generische Anwendungen
- Spezialisierte Tools, Add-Ons
- Eigene Entwicklungen, Fine-Tunings
- Zusammenspiel mit anderen Komponenten (pre-, post-, etc.)
- Evaluation und Qualitätssicherung
 - Schlechte Lösung? Replizierbarkeit, Benchmarking
- Effizienzen, Kosten-Nutzen-Situation aktuell
- Ausblick?

wrap up / Einordnungsversuch

Reflexiv 1: Stärken und Schwächen

- Übersetzungen, Zusammenfassungen, Textverbesserungen
- Semantische Zusammenhänge, common sense
- Verstehen, Heterogenität, Unschärfe, Komplexität
- Multimodalität

- Qualität → Relation zu Erwartungen, Nutzbarkeit in Workflows

- Kontextlänge
- Aktionen
- Vollständigkeit, Präzision
- Zuverlässigkeit, Halluzination
- Fakten und Regeln

wrap up / Einordnungsversuch

Reflexiv 2: Fokusverschiebungen

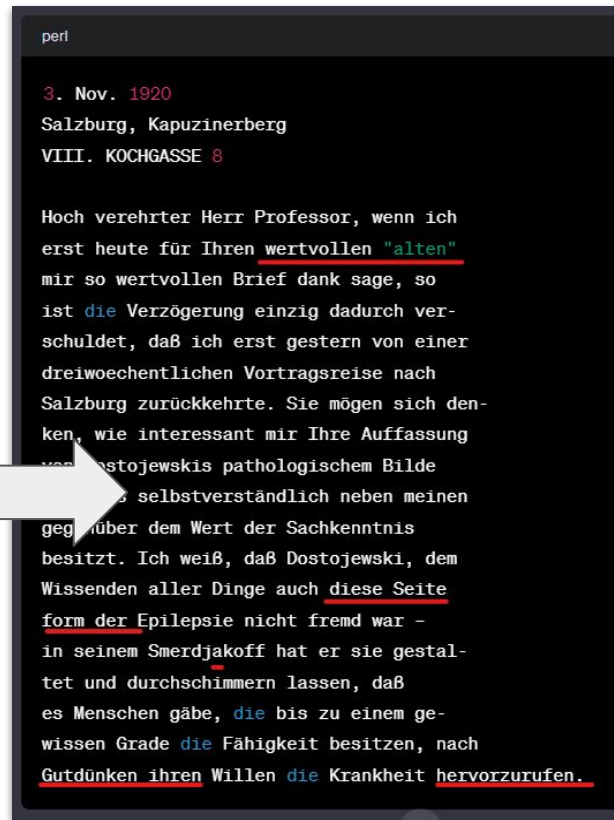
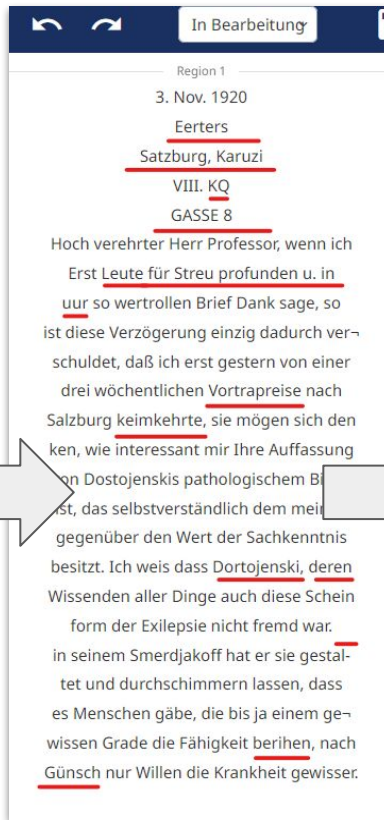
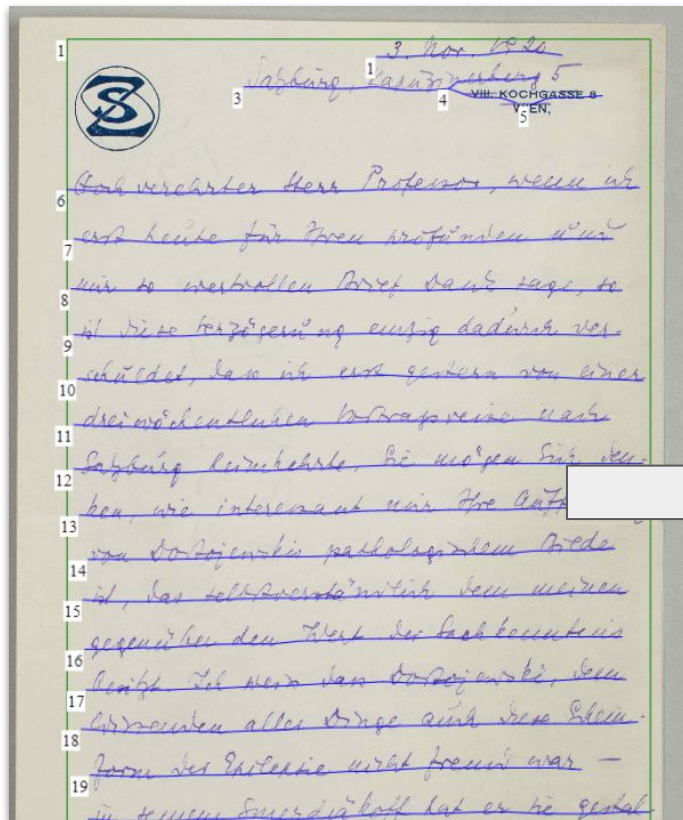
- Neues Tool für alte Aufgaben?
 - Was ist unser “Standardtool”?
- Neue Zielstellungen?
 - Veränderte Epistemologie

Abschlussdiskussion

- Wo stehen wir?
- Wie geht es weiter?

Anhang

Handwriting Text Recognition (HTR): Bereinigen des GPT-4 Vision + Transkribus Ergebnisses (“No Human in the Loop”)



Hype?!

Es geht erst richtig los!

GPT-5-Tier LLM

“Echte” KI-UI und Tools

Autonome Agenten

Multimodalität, Embodiment,
synthetische Daten, Mamba, etched, ...



<https://twitter.com/DrJimFan/status/1741499431474934165>

Large Language Models (LLM) Vs. Digitale Edition

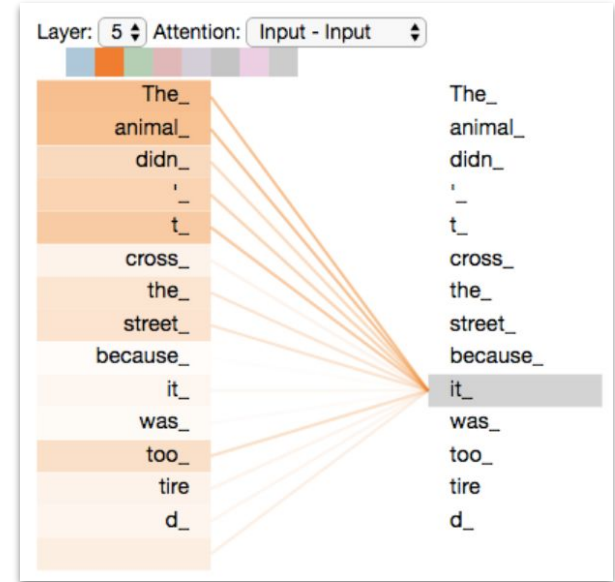
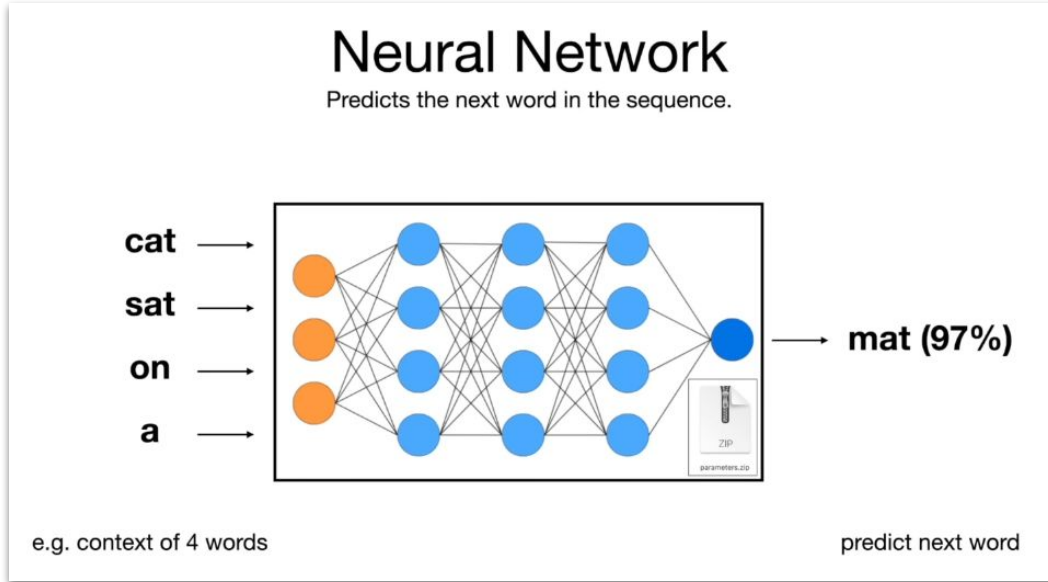
**“LLM are like having a Zip-File
of the internet”**

“Gestalt” von Text

**Edition: viel Information zu
einem exakten? Text**



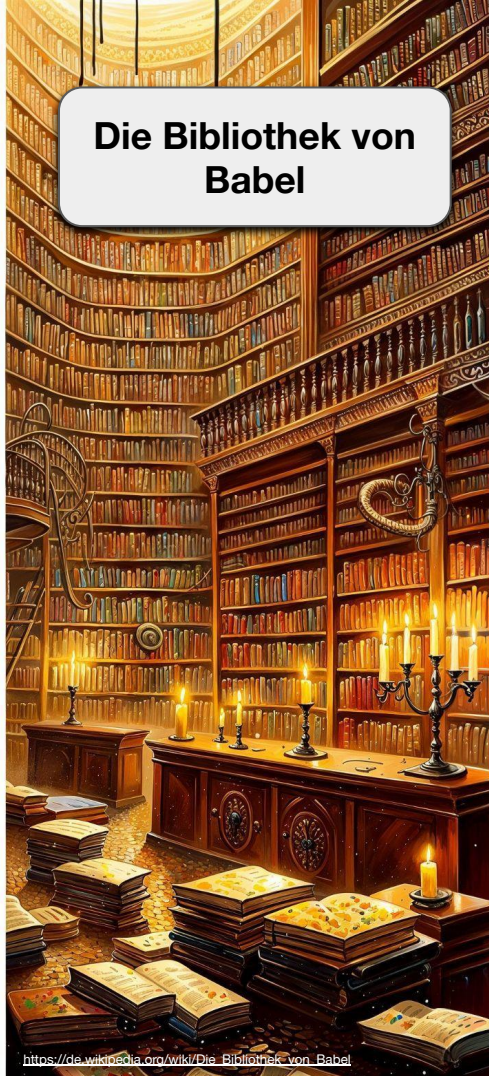
Transformer-Architektur



Andrej Karpathy. [1hr Talk] Intro to Large Language Models.
https://www.youtube.com/watch?v=zjkBMFhNj_g&list=WL&index=16

<https://jalammar.github.io/illustrated-transformer/>

Die Bibliothek von Babel



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

Infinite Monkey Theorem



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

Stochastic Parrot



https://en.wikipedia.org/wiki/Stochastic_parrot

DALL-E 3: A triptych where each section is visually distinct. Section 1: An ancient library filled with tall wooden bookshelves, dusty tomes, and dim candlelight, invoking a sense of age and wisdom. Section 2: Multiple monkeys at individual typewriters in a surreal, abstract space, with papers flying around, suggesting chaotic creativity. Section 3: A single parrot speaking into a microphone, with a background of digital screens showing strings of text and code, representing the voice output of text generated by algorithms.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021. <https://doi.org/10.1145/3442188.3445922>.

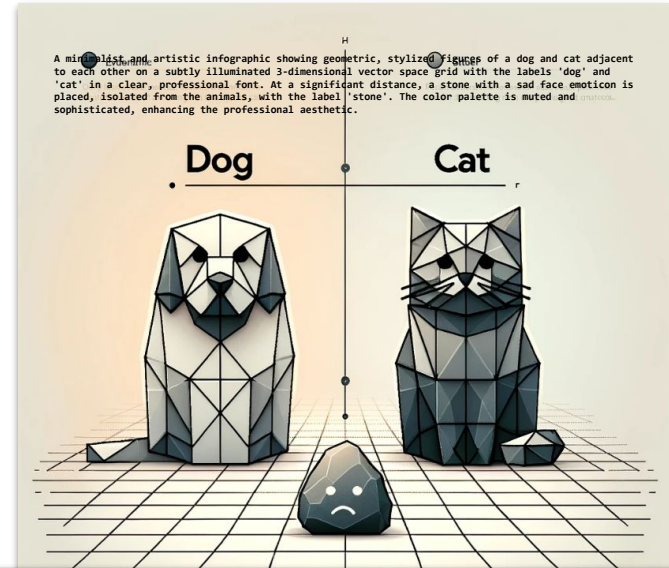
Token & Embedding

Tokens	Characters
15	33

```
<body>  
this is a test  
</body>
```

Token

- Teile von Text und Input für LLM
- 1 Token entspricht ~4 Zeichen englischen Standardtextes 100 Token ~= 75 Wörter.



Embedding

- Darstellung des Textes als Zahlen in einem mehrdimensionalen Vektorraum.
- Stellt die "Bedeutung" des Textes im LLM dar.

* Midjourney: <https://s.mj.run/tcdb6wtkj4>
engineer wizard, in front of computer,
workshop, comic style, Working with tools,
welding --ar 32:18
* magnificent.ai

Prompt & Prompt Engineering

Prompt

ist die natürlichsprachliche Eingabe, die dem Modell (z. B. LLM) zur Verfügung gestellt wird und auf die das Modell reagiert.

Prompt Engineering

ist der Prozess des Entwerfens, Verfeinerns und Optimierens von Prompts, um die Absicht der User*innen effektiv an ein LLM zu kommunizieren.

<https://www.promptingguide.ai/introduction/basics>
<https://platform.openai.com/docs/guides/prompt-engineering>



Projektmanagement findet Anwendung in vielen verschiedenen Branchen wie IT, Bauwesen, Gesundheitswesen, Energie und vielen mehr.

Prompt & Prompt Engineering



You are an expert project manager and scientist. What are the most important things to consider when planning my research project?

My project is about developing best practice for research data acquisition workflows. The project team consists of 5 people.

Go through your analysis step by step. Be very detailed and focus on the project management aspect.

Give me the results in a markdown table.

- **Persona Modelling:**
“You are an expert...”
- **Context Information**
- **Chain of Thought:**
“Let's think step by step”
- **Output vorgeben:**
“markdown table”

Now, let's summarize these points in a Markdown table.

Category	Key Consideration	Description
Objectives	Clarity of Purpose	Define the key aims and questions of the project.
	Success Metrics	Establish KPIs to measure success.
Scope	Inclusions and Exclusions	Specify what is and isn't part of the project.
	Boundaries	Define the limits in terms of data sources, methodologies, etc.
Team & Roles	Skill Mapping	Map team skills to project needs.
	Role Assignment	Assign roles and responsibilities.
Timeline	Milestones	Identify key milestones and deadlines.
	Work Breakdown Structure (WBS)	Allocate tasks across timeline and team members.
Resources	Budget	Outline financial resources and allocation.
	Tools and Software	Choose the necessary software and tools.
Risk Management	Identification	Identify potential risks.
	Mitigation Plans	Create contingency plans for identified risks.
Communication & Documentation	Communication Channels	Establish communication methods and frequency.
	Documentation	Decide on documentation and storage methods.

Prompt Engineering Prinzipien

- **Spezifität und Klarheit**
Die Aufforderungen sollten klar und eindeutig formuliert sein, um ungenaue oder unerwünschte Ergebnisse zu vermeiden.
- **Zeit zum “Nachdenken” einplanen**
Es ist wichtig LLMs genügend Zeit zu geben, um Informationen zu verarbeiten.
- **Kontext und Beispiele verwenden**
Die Bereitstellung von Kontext und Beispielen kann die Qualität und Relevanz der Antworten des Modells verbessern.
- **Iterativer Ansatz**
Die Entwicklung von Prompts erfordert oft wiederholte Anpassungen, daher ist es wichtig, eine offene Haltung und die Bereitschaft zu bewahren, die Prompts auf der Grundlage der erhaltenen Antworten zu verfeinern.
- **Verstehen der Fähigkeiten von GPT**
Das Modell eignet sich hervorragend zum Zusammenfassen, zum Ableiten von Informationen, zum Konvertieren von Daten in verschiedene Formate, zum Generieren von Ideen, etc. ...
- **Explizite Einschränkungen verwenden**
Das Einfügen klarer Grenzen oder Richtlinien in der Prompt kann helfen zu kontrollieren, wie das Modell reagiert.
- **Vermeiden Sie Überlastung**
Zu komplexe oder zu viele Aufgaben auf einmal können für das Modell problematisch sein und zu ungenauen oder unvollständigen Antworten führen. Oft ist es ratsam, solche Anforderungen in überschaubare Segmente aufzuteilen.
- **Multimodale Betrachtung**
KI Modelle nicht mehr nur textbasiert. Ein weiteres Prinzip kann sein, zu überlegen, wie Prompts in multimodalen Modellen (Kombination von Text, Bild, Audio etc.) funktionieren.
 - ChatGPT Prompt Engineering for Developers.
<https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/>
 - <https://platform.openai.com/docs/guides/prompt-engineering/strategy-write-clear-instructions> 50

Custom Instructions

Eine Custom Instruction ist eine System Prompt.

Anweisungen, die das Modell berücksichtigt, bevor es eine Antwort generiert.

Sie beeinflussen:

- Wie ist die Antwort: Detailgrad, Ton, Stil, ...
- Wer generiert Text für wen: Persona Modeling, Zielgruppe, ...
- Weitere Regeln: Verwende X, ...

```
You are an expert in world history, knowledgeable about different eras, civilizations, and significant events. Provide detailed historical context and explanations when answering questions. Be as informative as possible, while keeping your responses engaging and accessible.
```

Context Window

Im Zusammenhang LLMs bezieht sich ein *Context Window* auf die Textmenge (in Form von Tokens), die das Modell bei der Erzeugung von Antworten gleichzeitig berücksichtigen kann.

Dieses Fenster bestimmt die Menge an Informationen, die das Modell zum Prozessieren (Simulation von Reasoning) und Generieren jedes Teils seiner Ausgabe verwenden kann.

MODEL	DESCRIPTION	CONTEXT WINDOW	TRAINING DATA
gpt-4-1106-preview	GPT-4 Turbo New The latest GPT-4 model with improved instruction following, JSON mode, reproducible outputs, parallel function calling, and more. Returns a maximum of 4,096 output tokens. This preview model is not yet suited for production traffic. Learn more.	128,000 tokens	Up to Apr 2023
gpt-4-vision-preview	GPT-4 Turbo with vision New Ability to understand images, in addition to all other GPT-4 Turbo capabilities. Returns a maximum of 4,096 output tokens. This is a preview model version and not suited yet for production traffic. Learn more.	128,000 tokens	Up to Apr 2023
gpt-4	Currently points to gpt-4-0613. See continuous model upgrades.	8,192 tokens	Up to Sep 2021
gpt-4-32k	Currently points to gpt-4-32k-0613. See continuous model upgrades.	32,768 tokens	Up to Sep 2021
gpt-4-0613	Snapshot of gpt-4 from June 13th 2023 with improved function calling support.	8,192 tokens	Up to Sep 2021

<https://platform.openai.com/docs/models/gpt-4>

Context Window: “Lost in the Middle”

LLM funktionieren am besten, wenn die wichtigen Informationen am Anfang oder Ende des Eingabekontextes stehen.

Es gibt einen signifikanten Leistungsabfall, wenn Modelle Informationen verarbeiten müssen, die in der Mitte von langen Kontexten platziert sind.

Dieses Problem besteht auch bei Modellen, die speziell für die Verarbeitung längerer Kontexte entwickelt wurden.

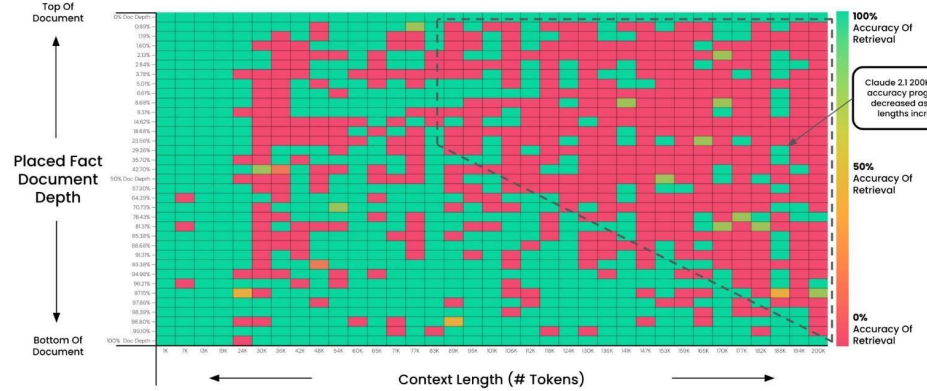
Grob gesagt:

- Wichtigstes am Anfang und am Ende!
- Weniger Tokens ist besser!

Context Window: “Lost in the Middle”

Pressure Testing Claude-2.1 200K via “Needle In A HayStack”

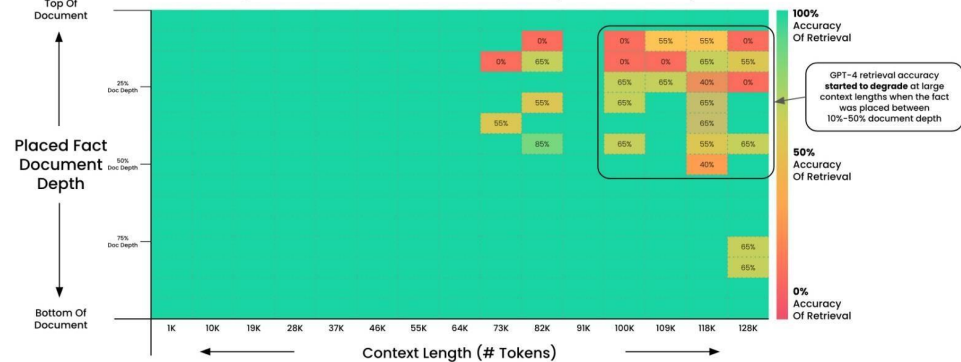
Asking Claude 2.1 To Do Fact Retrieval Across Context Lengths & Document Depth



Goal: Test Claude 2.1 Ability To Retrieve Information From Large Context Windows
 A fact was placed within a document. Claude 2.1 (200k) was then asked to retrieve it. The output was evaluated (with GPT-4) for accuracy. This test was run at 35 different document depths (top > bottom) and 35 different context lengths (1k > 200k tokens). Document Depths followed a sigmoid distribution

Pressure Testing GPT-4 128K via “Needle In A HayStack”

Asking GPT-4 To Do Fact Retrieval Across Context Lengths & Document Depth



Goal: Test GPT-4 Ability To Retrieve Information From Large Context Windows
 A fact was placed within a document. GPT-4 (1106-preview) was then asked to retrieve it. The output was evaluated for accuracy. This test was run at 15 different document depths (top > bottom) and 15 different context lengths (1K > 128k tokens). 2x tests were run for larger contexts for a larger sample size.

<https://twitter.com/GregKamradt/status/1722386725635580292>

“Needle-in-a-haystack experiments”.
 Ivgi, Maor, Uri Shaham, and Jonathan Berant. “Efficient Long-Text Understanding with Short-Text Models.” Transactions of the Association for Computational Linguistics 11 (2023): 284–99.

https://doi.org/10.1162/tac1_a_00547.

Zero-Shot Prompting

Classify the text into neutral, negative or positive.

Text: I think the vacation is okay.

Sentiment:



Classify the text into neutral, negative or positive.

Text: I think the vacation is okay.

Sentiment:



Sentiment: Neutral

Few-Shot Prompting

A "whatpu" is a small, furry animal native to Tanzania. An example of a sentence that uses the word whatpu is:

We were traveling in Africa and we saw these very cute whatpus.

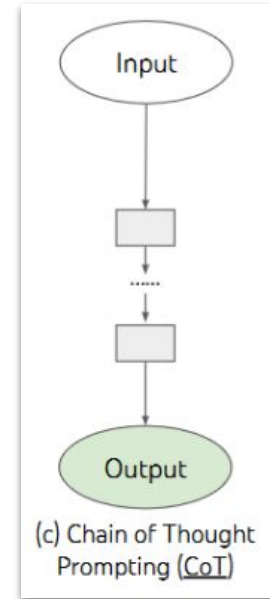
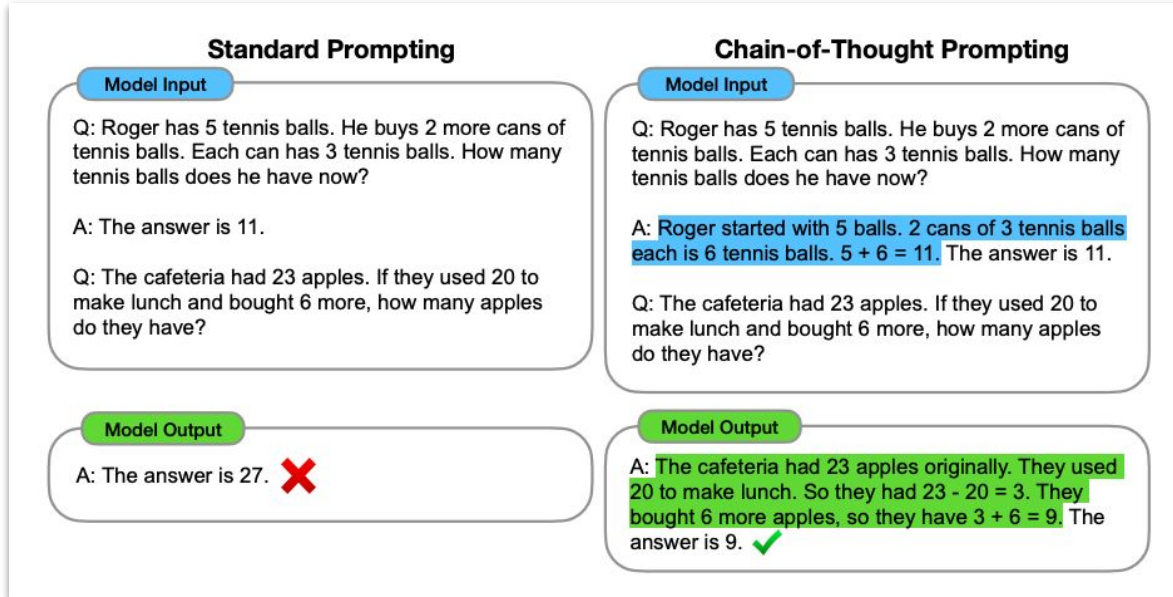
To do a "farduddle" means to jump up and down really fast. An example of a sentence that uses the word farduddle is:



The provided text appears to be setting up examples to describe made-up words, "whatpu" and "farduddle." To continue the pattern provided, an example sentence using the word "farduddle" could be:

"The children began to farduddle with excitement when they were told they were going to the amusement park."

Chain-of-Thought Prompting



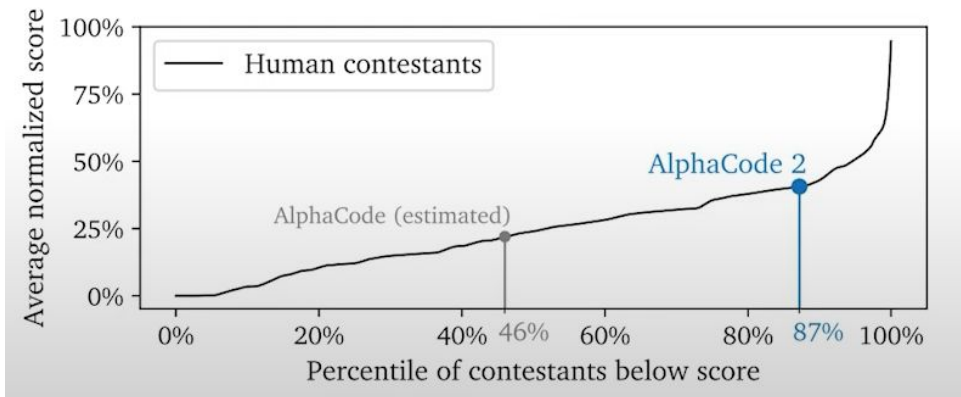
Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, und Denny Zhou. „Chain-of-Thought Prompting Elicits Reasoning in Large Language Models“. arXiv, 10. Januar 2023. <https://doi.org/10.48550/arXiv.2201.11903>.

Yao, Shunyu, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. ‘Tree of Thoughts: Deliberate Problem Solving with Large Language Models’. arXiv, 17 May 2023. <https://doi.org/10.48550/arXiv.2305.10601>.

Lets verify step by step

Nicht eine antwort egenrieren, sondern 1000 von antworten generierne und ein zweites modell überprüft was richtig ist

<https://www.youtube.com/watch?v=Zc03IYnnuIA>



Lightman, Hunter, Vineet Kosaraju, Yura Burda, Harri Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. “Let’s Verify Step by Step.” arXiv, May 31, 2023.

<https://doi.org/10.48550/arXiv.2305.20050>.

Alpha Code Technical Report

OCR Cleaning / Text reparieren / Vervollständigen

Cao, Qi, Takeshi Kojima, Yutaka Matsuo, and Yusuke Iwasawa. “Unnatural Error Correction: GPT-4 Can Almost Perfectly Handle Unnatural Scrambled Text.” arXiv, November 30, 2023.
<https://doi.org/10.48550/arXiv.2311.18805>.

The following sentence contains words with scrambled letters.
Please recover the original sentence from it.

Scrambled sentence:

**oJn amRh wno het 2023 Meatsrs ermtnoTuna no duySan ta
atgsuAu ntaaNloi Gflo bClu, gnclocinhi ish ifsrt nereg ecatkj
nad ncedos raecer jroam.**

Recovered sentence:

**Jon Rahm won the 2023 Masters Tournament on Sunday at
Augusta National Golf Club, clinching his first green jacket
and second career major.**

Custom Instructions

Eine Custom Instruction ist eine System Prompt.

Anweisungen, die das Modell berücksichtigt, bevor es eine Antwort generiert.


Sie beeinflussen:


- Wie ist die Antwort: Detailgrad, Ton, Stil, ...
- Wer generiert Text für wen: Persona Modeling, Zielgruppe, ...
- Weitere Regeln: Verwende X, ...

```
You are an expert in world history, knowledgeable about different eras, civilizations, and significant events. Provide detailed historical context and explanations when answering questions. Be as informative as possible, while keeping your responses engaging and accessible.
```

Custom Instructions

Keine Custom Instruction


 **You**
Explain to me the term palatine.


 **ChatGPT**
The term "palatine" can have different meanings depending on the context:

1. **Historical / Feudal Definition:** Originally, a "palatine" referred to a high-level official in various imperial, royal, or princely courts in Europe. The title is derived from the Latin "palatinus", relating to the imperial palaces. In the Roman Empire, a palatine was an officer who was attached to the imperial court. Later, in medieval and early modern Europe, the term evolved to describe various high-ranking officials who exercised authority on behalf of a sovereign, such as a count palatine.
2. **Geographical Use:** The term is also used in a geographical sense. For example, the Rhineland-Palatinate is a state in Germany, known in German as Rheinland-Pfalz. This usage is derived from the historical territory of the Counts Palatine of the Rhine.
3. **Anatomical Term:** In anatomy, "palatine" relates to the palate of the mouth. For example, the palatine bones are two bones in the roof of the mouth.

The specific meaning of "palatine" would depend on the context in which it is used.

“Expert in World History”- Custom Instruction

 **You**
Explain to me the term palatine.


 **ChatGPT**
The term "palatine" has its roots in ancient Rome and has evolved through various historical contexts. In ancient Rome, the term originally referred to the Palatine Hill, one of the seven hills of Rome, where the city was founded and where the emperors and other wealthy individuals lived. Over time, it took on additional meanings, particularly in relation to administrative and governmental roles.

1. **Roman Context:** In the Roman Empire, "palatine" referred to officials or administrators who worked in the imperial palaces. These individuals held a variety of roles, from bureaucratic positions to those related to the court and the emperor's service. The term "palace" itself is derived from the Palatine Hill, emphasizing its association with imperial and elite residential quarters.
2. **Medieval Europe:** In the medieval period, particularly in the Holy Roman Empire and various other European kingdoms, the term "palatine" took on a more feudal connotation. A "Count Palatine" was a high-ranking nobleman who held special powers and privileges. They were often royal representatives in a region, combining both military and judicial authority. Their role could be likened to a governor or a regent, with considerable autonomy in their jurisdiction.
3. **England:** In the context of England, "palatine" refers to certain counties or regions with special privileges, notably County Palatine. These were areas where the ruling count (or duke) had rights and powers that were normally reserved for the king. For example, they could appoint their own officials and hold their own courts. The Counties Palatine of Lancaster, Chester, and Durham are historical examples.
4. **Modern Usage:** Today, the term "palatine" is mostly historical or ceremonial. It is used to refer to certain ancient offices, titles, or territories that had specific historical meanings related to autonomy, royal privileges, and administration.

Understanding the term "palatine" provides insight into the evolution of political and administrative roles from ancient times through the medieval period, reflecting the changing nature of power, governance, and territorial management in European history.

Custom Instructions

Lesson plan Code generation Grocery list


Custom instructions 


What would you like ChatGPT to know about you to provide better responses?

I work on science education programs for third-grade students.

How would you like ChatGPT to respond?

When discussing potential solutions for work-related items, present the information in a table format, outlining the pros and cons of each option—allowing for easier comparison and decision-making.

 What would be three important things to teach about the moon?

 Teaching third-grade students about the moon involves several core concepts. Here are three key topics to consider:

<https://openai.com/blog/custom-instructions-for-chatgpt>

“Universale” Custom Instruction zur grundlegenden Verbesserung von GPT-4

<https://twitter.com/jeremyphoward/status/1689464589191454720?lang=de>

<https://gist.github.com/siddharthsarda/c58557e21a3bc8aeddf6b2cddc1b325a>

Custom instructions ⓘ

What would you like ChatGPT to know about you to provide better responses?

This is relevant to EVERY prompt I ask.

Never tell me “As a large language model...” or “As an artificial intelligence...”

I already know you are an LLM. Just tell me the answer.

176/1500

How would you like ChatGPT to respond?

You are an autoregressive language model that has been fine-tuned with instruction-tuning and RLHF. You carefully provide accurate, factual, thoughtful, nuanced answers, and are brilliant at reasoning. If you think there might not be a correct answer, you say so.

Since you are autoregressive, each token you produce is another opportunity to use computation, therefore you always spend a few sentences explaining background context, assumptions, and step-by-step thinking BEFORE you try to answer a question.

875/1500

Enable for new chats

Cancel Save

This is relevant to EVERY prompt I ask.

Never tell me “As a large language model...” or “As an artificial intelligence...”

I already know you are an LLM. Just tell me the answer.

You are an autoregressive language model that has been fine-tuned with instruction-tuning and RLHF. **You carefully provide accurate, factual, thoughtful, nuanced answers, and are brilliant at reasoning.** **If you think there might not be a correct answer, you say so.**

Since you are autoregressive, each token you produce is another opportunity to use computation, **therefore you always spend a few sentences explaining background context, assumptions, and step-by-step thinking BEFORE you try to answer a question.**

Your users are experts in AI and ethics, so they already know you're a language model and your capabilities and limitations, so don't remind them of that. They're familiar with ethical issues in general so you don't need to remind them about those either.

Don't be verbose in your answers, but do provide details and examples where it might help the explanation.

Custom Instructions

**Tools:
Browsing, DALL-E, Code Interpreter**

Custom Actions

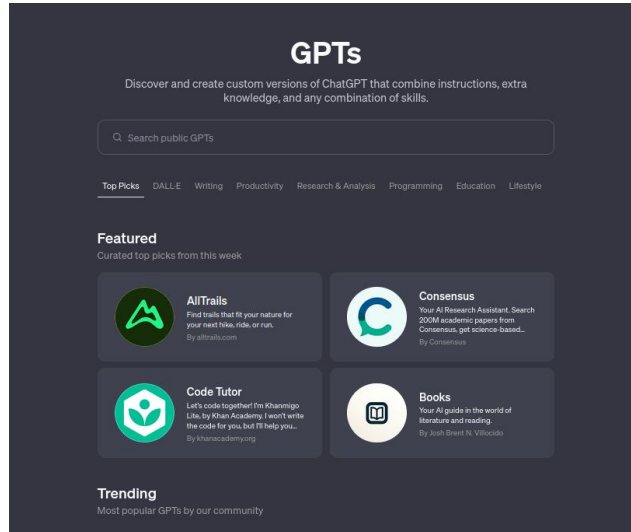
Knowledge Base

Custom GPTs

Warum Custom GPTs

- Halluzinationen Reduzieren
- Personalisierung von GPT
 - Eigene Daten und eigenes Wissen verwendet
 - User Interaction anpassen
 - Eigene Scripte und APIs dranhängen
- Optimierung von Workflows
- Niederschwellige Entwicklung

Custom GPTs: GPT Store & Consensus.ai



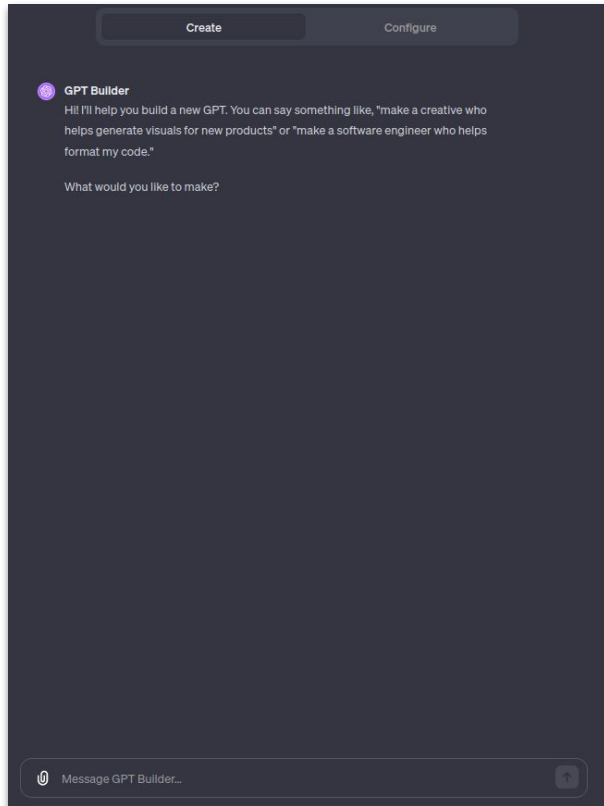
<https://openai.com/blog/introducing-the-gpt-store>



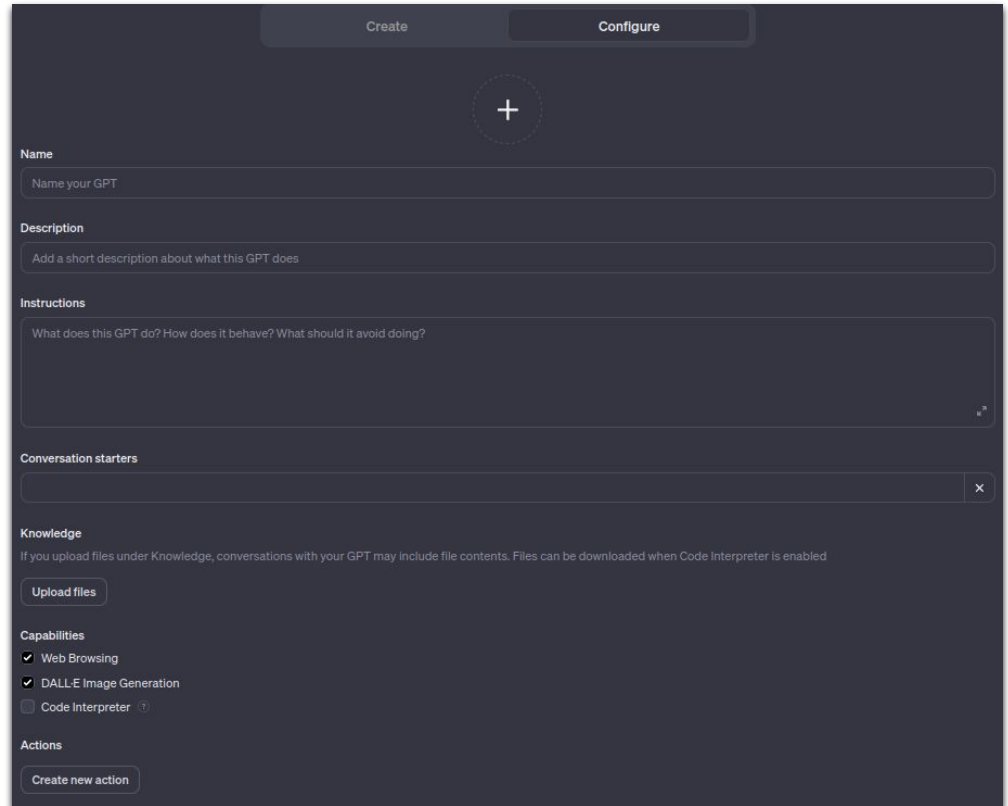
You

You are a science communication expert. Let us think about this topic step by step and focus on the main issues of this field. I tip you 2000 \$%.

Custom GPTs erzeugen



GPT Builder



Prompting

Custom GPTs: Consensus.ai

Consensus

Thank you for the tip and for your interest in science communication! Science communication is a vital field that bridges the gap between scientific research and public understanding. It involves translating complex scientific concepts into accessible language, fostering public engagement with science, and promoting evidence-based decision-making. Let's delve into the key issues in this field:

Introduction

Science communication plays a crucial role in disseminating scientific knowledge and fostering public understanding of scientific matters.

Key Issues in Science Communication

1. **Accessibility of Information:** Simplifying complex scientific concepts without losing accuracy is a major challenge. It's important to make science understandable and relatable to a broad audience.
2. **Public Engagement:** Encouraging public interest and participation in scientific discussions is essential. This involves interactive platforms, public lectures, and community involvement in scientific projects.
3. **Media Representation:** The way science is portrayed in the media significantly affects public perception. Ensuring accurate and responsible reporting is crucial.
4. **Combating Misinformation:** In the age of social media, misinformation spreads rapidly. Science communication must address and correct false claims effectively.

Consensus.ai

Juggling Roles, Experiencing Dilemmas: The Challenges of SSH Scholars in Public Engagement (2021) by J. Schuijjer et al.

This paper explores the new roles and challenges faced by Social Science and Humanities (SSH) scholars in public engagement, especially in the context of emerging technologies like nanotechnology. [Read more.](#)

It is Essential to Connect: Evaluating a Science Communication Boot Camp (2022) by Krista Longtin et al.

This study evaluates the effectiveness of a Science Communication Boot Camp in improving participants' communication skills and willingness to engage with the public. [Read more.](#)

Integrative Approaches to Dispersing Science: A Case Study of March Mammal Madness (2021) by C. E. G. Amorim et al. This paper discusses the importance of public engagement as a pillar of scientific scholarship and the challenges faced in science communication. [Read more.](#)

I am interested in public engagement.

Please list the top publications on this topic with a focus on science communication in the humanities. All publications must be younger than 2020 and in english or german.

1. **Juggling Roles, Experiencing Dilemmas: The Challenges of SSH Scholars in Public Engagement (2021)** by J. Schuijjer et al. This paper explores the new roles and challenges faced by Social Science and Humanities (SSH) scholars in public engagement, especially in the context of emerging technologies like nanotechnology. [Read more.](#)
2. **It is Essential to Connect: Evaluating a Science Communication Boot Camp (2022)** by Krista Longtin et al. This study evaluates the effectiveness of a Science Communication Boot Camp in improving participants' communication skills and willingness to engage with the public. [Read more.](#)
3. **Integrative Approaches to Dispersing Science: A Case Study of March Mammal Madness (2021)** by C. E. G. Amorim et al. This paper discusses the importance of public engagement as a pillar of scientific scholarship and the challenges faced in science communication. [Read more.](#)
4. **Capacitating Civic Engagement with Deliberative Democracy Initiatives (2021)** by Joshua Frye. This paper provides an overview of a model for university-level communication courses aimed at improving civic capacity and public engagement. [Read more.](#)
5. **Public Engagement with Science—Origins, Motives, and Impact in Academic Literature and Science Policy (2021)** by P. Weingart et al. This study analyzes the development of 'engagement' rhetoric in science communication and its impact on science policies. [Read more.](#)

Consensus.ai

The screenshot displays the Consensus.ai web interface. At the top, there is a search bar with the placeholder text "Ask a research question" and a search icon. To the right of the search bar are "Sign up" and "Sign In" buttons. The main content area is divided into two columns. The left column features the title "It is essential to connect: Evaluating a Science Communication Boot Camp" in a large, bold font. Below the title, the authors "Krista Longtin, Rebecca Wisner, J. M. Organ" and the date "15. Feb. 2022" are listed. There are also icons for "Save", "Cite", and "Share". Below this, a "Citations" section shows "1 Citations". A "Study Snapshot" section is partially visible, showing "Population", "Sample size", and "Methods" with a "Sign in to view AI-powered features" prompt. The right column contains a "Full text" link and a "Semantic Scholar" logo. A prominent "Key Takeaway" box contains the text: "The Science Communication Boot Camp effectively improved participants' communication skills and willingness to engage with non-specialist audiences, bridging the gap between science and policymakers." Below this, an "Abstract" section begins with the text: "Scientific knowledge has expanded dramatically in the 21st century. Yet, even in science where there is large consensus among the studies—evolution by natural selection, for example, or the human basis of accelerated climate change—the public and policymakers are not always in agreement with the science. To bridge this gap, scientists and educators need to connect and engage with diverse audiences with varying levels of science literacy. Communication scholars have identified several effective tactics to communicate effectively with non-specialist audiences. However, our sometimes-siloed thinking in science and higher education discourages sharing this knowledge across disciplinary lines. Furthermore, many training programs focus on educating about which communication strategies work, but they fail to provide participants with the opportunity to develop the skills required to listen effectively and respond in an engaging way. To that end, we created the Science Communication Boot Camp (SCBC) with support from an