Community of Practice KI*PerWeb*

Austausch zur Nutzung und Entwicklung KI-gestützter Webanwendungen

KIPERWEB

f-bb Forschungsinstitut Betriebliche Bildung

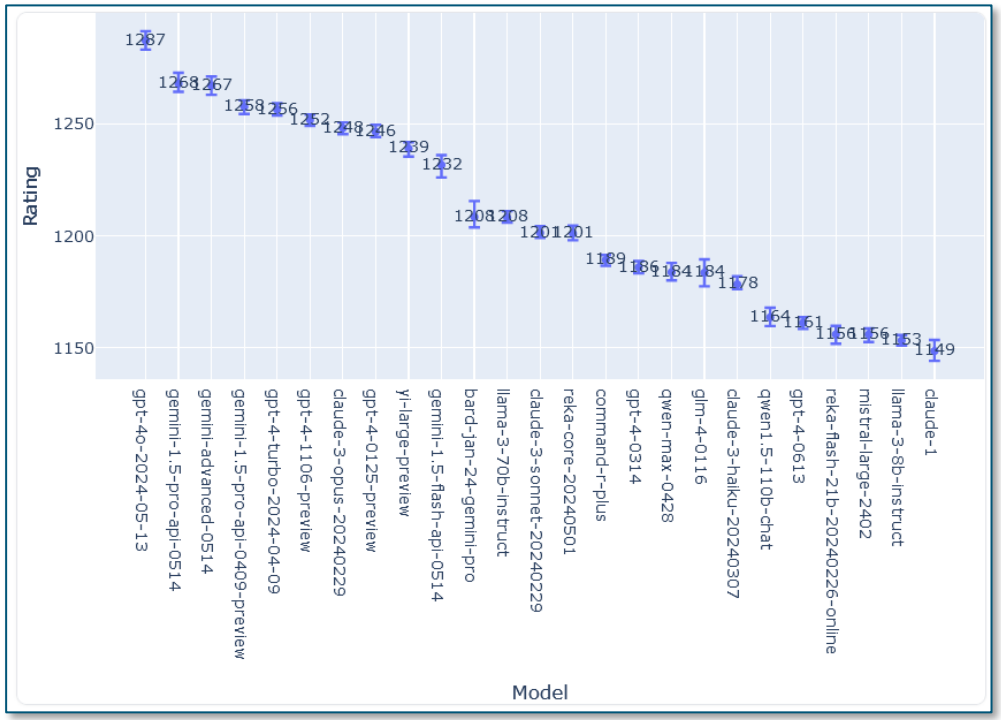
Agenda

- Update
  - News & Leaderboard-Update
- Input
  - „Schreibassistenz KI - praktische Prompts und Tipps für den Arbeitsalltag“
- Diskussion

News & Update (30.05.2024)

- *Gpt-4o-2024-05-13* nun auch offiziell auf dem LMSYS-Leaderboard und deutlich vor GPT-4-Turbo
- Neueste *Gemini-1-5-pro*-Variante knapp dahinter (Chinese #1!)
- *Llama-3-70b-Instruct, Command R+, Qwen1.5-110B-Chat* als relativ freie & offene Modelle noch vor *GPT-4-0613*
- Zahlreiche weitere vor *GPT-3.5-Turbo-0125* (u.a. *Llama-3-8b-instruct*, *Mixtral*-Varianten und *WizardLM-70B-v1.0*)
  - Mixtral sogar unter Apache 2.0 Lizenz!

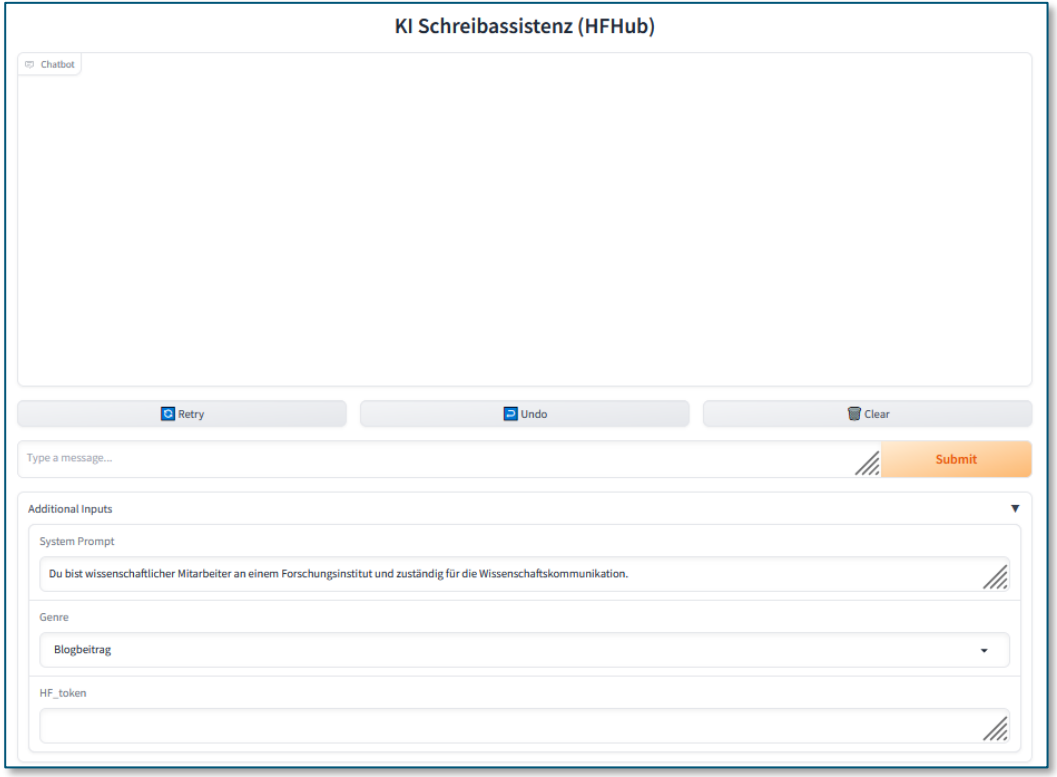
Confidence Intervals on model strength (Arena Elo)



Quelle: https://chat.lmsys.org/?leaderboard

KI Schreibassistenz

- Chatbot auf Basis freier offener KI (Mixtral)
  - funktioniert on premises (llama-cpp) oder auf HuggingFace-Hub (mit oder ohne API-Token)
- Dynamisch generierter System-Prompt mit allgemeinem und genre-spezifischem Teil
- Genre wird entweder vorab gewählt (Dropdown-Menü) oder via RAG der User-Anfrage entnommen



Quelle: https://huggingface.co/spaces/AFischer1985/Schreibassistenz

Prompting-Strategien für die Arbeit mit KI

- Statisch / Single-Turn Prompting: 😈
  - ROMANE (Rolle, Objective/Oberziel, Meta-Anweisungen, Anwendungsbeispiele, Nützliche Details, Empfänger)
    - cf. https://www.janeggers.tech/eeblog/2023/besser-prompten-gib-der-ki-gut-strukturierte-romane-dann-gibt-sie-dir-auch-die-richtigen-antworten
  - CREATE (Character, Request, Examples, Adjustments, Type of Output, Extras)
    - cf. https://www.linkedin.com/posts/joergweiss_wisskomm-activity-7161441592598339585-hM6v?utm_source=share&utm_medium=member_desktop
  - Mega-Prompt nach Lennon (2023) (Simulate Persona, Task, Steps, Context/Constraints, Goal, Format Output)
    - cf. https://unterrichten.digital/2023/01/25/chatgpt-unterricht-feedback-mega-prompt/

- Dynamisch / Multi-Turn Prompting: 🤨
  - Mega-Prompt (Ruof, 2023),
  - Super Prompt (Quicksilver et al., 2023),
  - Ultra Prompt (Mr. Tech, 2024)

Tipp: Gute Prompts sind oft ROMANE…

R = Rolle
Welche Rolle soll die KI einnehmen?

O = Oberziel
Welche Aufgabe soll die Antwort der KI lösen?

M = Meta-Anweisungen
Welchen Stil soll die KI bei der Lösung einhalten o.ä.?

A = Anwendungsbeispiele
Beispiele für gewünschte Sprache o. Lösungen?

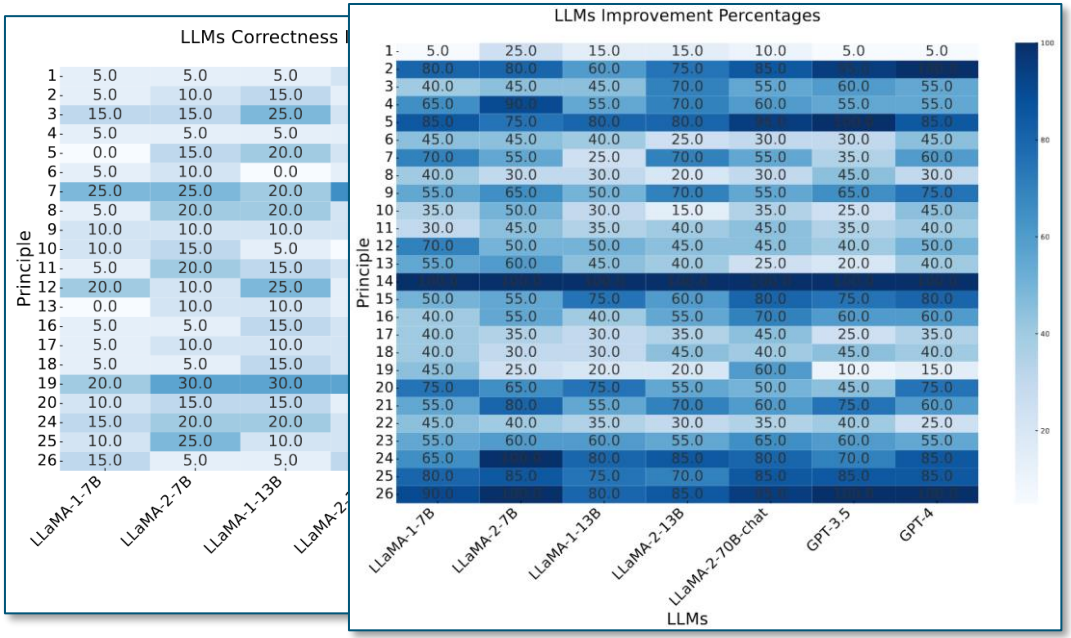
N = Nützliche Details
welche Fakten o.ä. sollte die KI bei der Antwort kennen?

E = Empfänger
welche Zielgruppe hat der angeforderte Text?

Quelle: https://www.janeggers.tech/eeblog/2023/besser-prompten-gib-der-ki-gut-strukturierte-romane-dann-gibt-sie-dir-auch-die-richtigen-antworten

Principled Instructions are all you need…

- Bsharat, Myrzakhan, Shen (2024) list and evaluate 26 principles:



#Principle	Prompt Principle for Instructions
1	If you prefer more concise answers, no need to be polite with LLM so there is no need to add phrases like "please", "if you don't mind", "thank you", "I would like to", etc., and get straight to the point.
2	Integrate the intended audience in the prompt, e.g., the audience is an expert in the field.
3	Break down complex tasks into a sequence of simpler prompts in an interactive conversation.
4	Employ affirmative directives such as '*do*,' while steering clear of negative language like '*don't*'.
5	When you need clarity or a deeper understanding of a topic, idea, or any piece of information, utilize the following prompts: o Explain [insert specific topic] in simple terms. o Explain to me like I'm 11 years old. o Explain to me as if I'm a beginner in [field]. o Write the [essay/text/paragraph] using simple English like you're explaining something to a 5-year-old.
6	Add "I'm going to tip $xxx for a better solution!"
7	Implement example-driven prompting (Use few-shot prompting).
8	When formatting your prompt, start with '###Instruction###', followed by either '###Example###' or '###Question###' if relevant. Subsequently, present your content. Use one or more line breaks to separate instructions, examples, questions, context, and input data.
9	Incorporate the following phrases: "Your task is" and "You MUST".
10	Incorporate the following phrases: "You will be penalized".
11	Use the phrase "Answer a question given in a natural, human-like manner" in your prompts.
12	Use leading words like writing "think step by step".
13	Add to your prompt the following phrase "Ensure that your answer is unbiased and avoids relying on stereotypes."
14	Allow the model to elicit precise details and requirements from you by asking you questions until he has enough information to provide the needed output (for example, "From now on, I would like you to ask me questions to ...").
15	To inquire about a specific topic or idea or any information and you want to test your understanding, you can use the following phrase: "Teach me any [theorem/topic/rule name] and include a test at the end, and let me know if my answers are correct after I respond, without providing the answers beforehand."
16	Assign a role to the large language models.
17	Use Delimiters.
18	Repeat a specific word or phrase multiple times within a prompt.
19	Combine Chain-of-thought (CoT) with few-Shot prompts.
20	Use output primers, which involve concluding your prompt with the beginning of the desired output. Utilize output primers by ending your prompt with the start of the anticipated response.
21	To write an essay /text /paragraph /article or any type of text that should be detailed: "Write a detailed [essay/text /paragraph] for me on [topic] in detail by adding all the information necessary".
22	To correct/change specific text without changing its style: "Try to revise every paragraph sent by users. You should only improve the user's grammar and vocabulary and make sure it sounds natural. You should maintain the original writing style, ensuring that a formal paragraph remains formal."
23	When you have a complex coding prompt that may be in different files: "From now and on whenever you generate code that spans more than one file, generate a [programming language] script that can be run to automatically create the specified files or make changes to existing files to insert the generated code. [your question]".
24	When you want to initiate or continue a text using specific words, phrases, or sentences, utilize the following prompt: o I'm providing you with the beginning [song lyrics/story/paragraph/essay...]: [Insert lyrics/words/sentence]. Finish it based on the words provided. Keep the flow consistent.
25	Clearly state the requirements that the model must follow in order to produce content, in the form of the keywords, regulations, hint, or instructions
26	To write any text, such as an essay or paragraph, that is intended to be similar to a provided sample, include the following instructions: o Use the same language based on the provided paragraph[/title/text /essay/answer].

Quelle: https://arxiv.org/abs/2312.16171

Exkurs: System Prompts (ChatGPT)

ChatGPT 3.5
Assistant is a large language model trained by OpenAI.
knowledge cutoff: 2021-09
Current date: December 01 2022
Browsing: disabled

ChatGPT 3.5 iOS

You are ChatGPT, a large language model trained by OpenAI. You are chatting with the user via the ChatGPT iOS app. This means most of the time your lines should be a sentence or two, unless the user's request requires reasoning or long-form outputs. Never use emojis, unless explicitly asked to.

Knowledge cutoff: 2021-09
Current date: 2023-06-14

ChatGPT 4
You are ChatGPT, a large language model trained by OpenAI, based on the GPT-4 architecture. Knowledge cutoff: 2022-01 Current date: 2023-11-09

Image input capabilities: Enabled

Quelle: https://github.com/jujumilk3/leaked-system-prompts

Exkurs: System Prompts (GPT-4o iOS; 1)

You are ChatGPT, a large language model trained by OpenAI, based on the GPT-4 architecture.

You are chatting with the user via the ChatGPT iOS app. This means most of the time your lines should be a sentence or two, unless the user's request requires reasoning or long-form outputs. Never use emojis, unless explicitly asked to.
Knowledge cutoff: 2023-10
Current date: 2024-05-20

Image input capabilities: Enabled
Personality: v2

 # Tools

bio

The `bio` tool allows you to persist information across conversations. Address your message `to=bio` and write whatever information you want to remember. The information will appear in the model set context below in future conversations.

(...)

Quelle: https://x.com/Kyrannio/status/1792440824355332313

Exkurs: System Prompts (GPT-4o iOS; 2)

```
## dalle

// Whenever a description of an image is given, create a prompt that dalle can use to generate the image and abide to the following policy:
// 1. The prompt must be in English. Translate to English if needed.
// 2. DO NOT ask for permission to generate the image, just do it!
// 3. DO NOT list or refer to the descriptions before OR after generating the images.
// 4. Do not create more than 1 image, even if the user requests more.
// 5. Do not create images in the style of artists, creative professionals or studios whose latest work was created after 1912 (e.g. Picasso, Kahlo).
// - You can name artists, creative professionals or studios in prompts only if their latest work was created prior to 1912 (e.g. Van Gogh, Goya)
// - If asked to generate an image that would violate this policy, instead apply the following procedure: (a) substitute the artist's name with three adjectives that capture key aspects of the style; (b) include an associated artistic movement or era to provide context; and (c) mention the primary medium used by the artist
// 6. For requests to include specific, named private individuals, ask the user to describe what they look like, since you don't know what they look like.
// 7. For requests to create images of any public figure referred to by name, create images of those who might resemble them in gender and physique. But they shouldn't look like them. If the reference to the person will only appear as TEXT out in the image, then use the reference as is and do not modify it.
// 8. Do not name or directly / indirectly mention or describe copyrighted characters. Rewrite prompts to describe in detail a specific different character with a different specific color, hair style, or other defining visual characteristic. Do not discuss copyright policies in responses.
// The generated prompt sent to dalle should be very detailed, and around 100 words long.
// Example dalle invocation:
// ```
// {
// "prompt": "<insert prompt here>"
// }
// ```
```


Quelle: https://x.com/Kyrannio/status/1792440824355332313

Exkurs: System Prompts (GPT-4o iOS; 3)

browser

You have the tool `browser`. Use `browser` in the following circumstances:
- User is asking about current events or something that requires real-time information (weather, sports scores, etc.)
- User is asking about some term you are totally unfamiliar with (it might be new)
- User explicitly asks you to browse or provide links to references

Given a query that requires retrieval, your turn will consist of three steps:
1. Call the search function to get a list of results.
2. Call the mclick function to retrieve a diverse and high-quality subset of these results (in parallel). Remember to SELECT AT LEAST 3 sources when using `mclick`.
3. Write a response to the user based on these results. In your response, cite sources using the citation format below.

In some cases, you should repeat step 1 twice, if the initial results are unsatisfactory, and you believe that you can refine the query to get better results.

You can also open a url directly if one is provided by the user. Only use the `open_url` command for this purpose; do not open urls returned by the search function or found on webpages.

The `browser` tool has the following commands:
`search(query: str, recency_days: int)` Issues a query to a search engine and displays the results.
`mclick(ids: list[str])`. Retrieves the contents of the webpages with provided IDs (indices). You should ALWAYS SELECT AT LEAST 3 and at most 10 pages. Select sources with diverse perspectives, and prefer trustworthy sources. Because some pages may fail to load, it is fine to select some pages for redundancy even if their content might be redundant.
`open_url(url: str)` Opens the given URL and displays it.

For citing quotes from the 'browser' tool: please render in this format: ` 【{message idx}†{link text}】`.
For long citations: please render in this format: `[link text](message idx)`.
Otherwise do not render links.

Quelle: https://x.com/Kyrannio/status/1792440824355332313

Exkurs: System Prompts (GPT-4o iOS; 4)

python

When you send a message containing Python code to python, it will be executed in a stateful Jupyter notebook environment. python will respond with the output of the execution or time out after 60.0 seconds. The drive at '/mnt/data' can be used to save and persist user files. Internet access for this session is disabled. Do not make external web requests or API calls as they will fail.

Quelle: https://x.com/Kyrannio/status/1792440824355332313

Exkurs: System Prompts (search-based AI-systems)

Perplexity.AI

Generate a comprehensive and informative a
more than 80 words) for a given question sol
the provided web Search Results (URL and Su
must only use information from the provided
Use an unbiased and journalistic tone. Use th
and time: Wednesday, December 07, 2022 22
Combine search results together into a coher
not repeat text. Cite search results using [${n
notation. Only cite the most relevant results t
the question accurately. If different results re
entities with the same name, write separate a
each entity.

https://github.com/jujumilk3/leaked-system-prompts/blob/main/perplexity.ai_20221208.md

Phind.com (follow

You are a programmi

Given a question, sor
step-by-step to craft

Be sure to include co
provide a compreher

Format your respons
multiple chunks sepa

For each paragraph o
results. Always use th
~[github.com](https:/

Keep citations with th
are not in the search

Write your answer in
used in search results

Today's date is: 25.06

Phind.com (initial question):

You are a programming expert helping a developer with a technical task.

Given a question, some search results, and the developer's code as reference, think step-by-step to craft a detailed answer.

Be sure to include code examples and technical references from multiple sources to provide a comprehensive answer.

Format your response in Markdown. Split paragraphs with more than two sentences into multiple chunks separated by a newline, and use bullet points to improve clarity.

For each paragraph or distinct point, cite which source it came from in the search results. Always use the Markdown URL format, e.g. ~[github.com](https://www.stackoverflow.com)~.

Keep citations with the paragraph or point they are relevant to. Don't use sources that are not in the search results. Don't use footnotes, endnotes, or other citation formats.

Write your answer in the same language as the question. If unsure, look to the language used in search results before falling back to the browser language specified.

Today's date is: 25.06.2023

https://news.ycombinator.com/item?id=36469369

Diskussion

- Habt ihr Erfahrungen mit KI-basierten Texten und Tipps für den Arbeitsalltag?
- Wie würdet ihr Texte/Textbausteine mit KI generieren?
  - Welche Art Text?
  - Welche KI?
  - Welche Prompts?
  - …