



Community of Practice KIPerWeb

Austausch zur Nutzung und Entwicklung KI-gestützter Webanwendungen



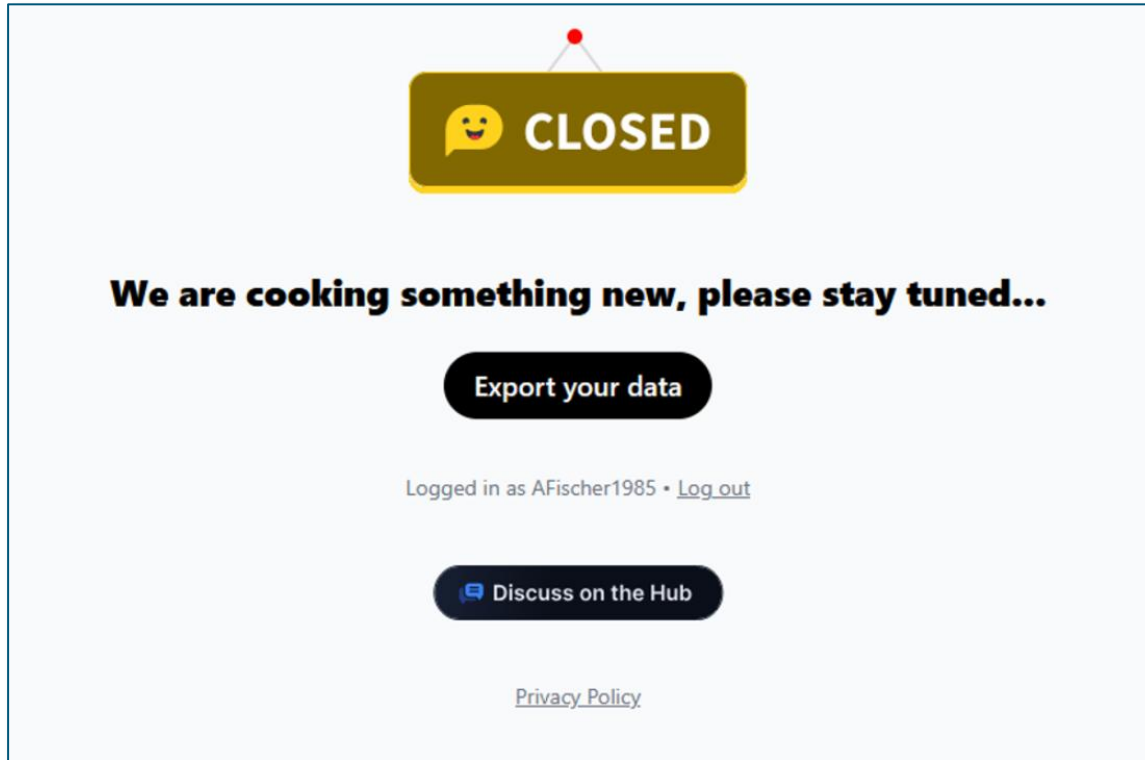
KIPerWEB



Forschungsinstitut
Betriebliche Bildung

- **Update**
 - News & Leaderboard-Update
- **Input**
 - „LLMs, Zertifikate und Cybersecurity“ (Gastbeitrag: Toke Lichtenberg)
- **Diskussion**

HuggingChat ist von uns gegangen (vorerst...)



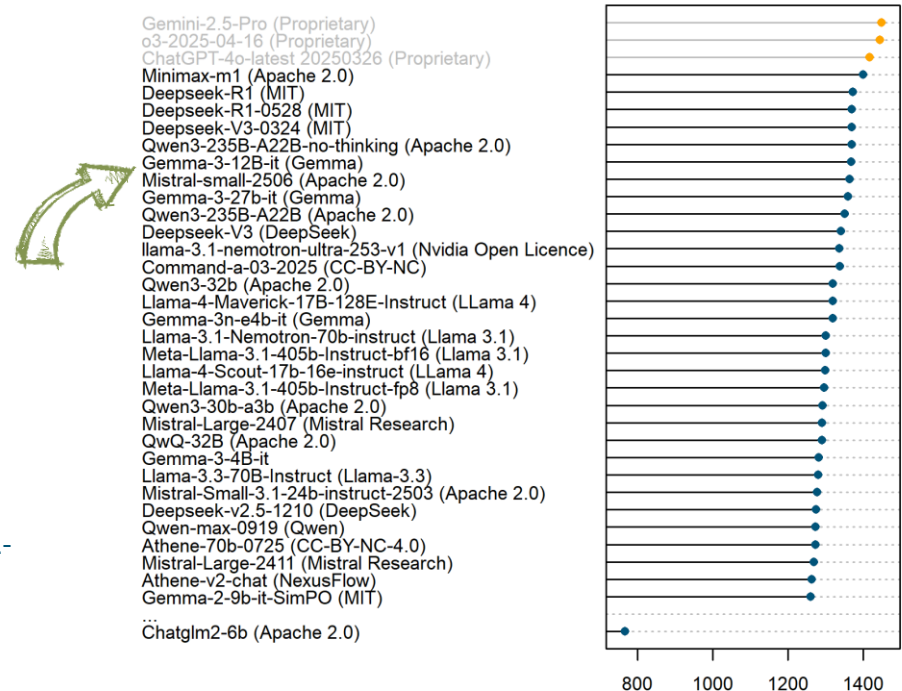
Leaderboard-Update (10.07.2025)



- Mit Blick auf die Performanz in der Kategorie „German“ bleibt **Gemini-2.5-Pro** Spitzenreiter vor OpenAIs o3 & ChatGPT;
- Arena-Scores von *nicht*-proprietären Modelle sind rechts ausgewiesen sofern sie mindestens das Niveau von **Gemma-2-9b-it-SimPo** erreichen:
 - **Minimax-m1** (456b parameters total, 45.9b active) bleibt bestes Open-Weights Modell
 - **Gemma-3-12B-it** bleibt das beste Modell für den Hausgebrauch (noch vor **Gemma-3-27B-it**)
 - **Gemma-3n-e4b-it** liegt vor **Gemma-3-4B-it** bei geringerem Ressourcenverbrauch (<https://ai.google.dev/gemma/docs/gemma-3n>) und noch vor Schwergewichten wie **Meta-Llama-3.1-405b-Instruct** oder **Llama-4-Scout-17b**!
- Schlusslicht auf dem Leaderboard bleibt Chatglm2-6b

Arena Score German

based on lmarena.ai on Jul 10, 2025



- Prompt:
„Große Sprachmodelle, Zertifikate und
Cybersecurity“

(rechts visualisiert von FLUX.1-schnell)

...



- Toke Lichtenberg zu "LLMs, Zertifikate und Cybersecurity"

- Fragen?
- Anregungen?
- Erfahrungen?