

AI Community Meeting

May 2026



Today's Topics



Recent updates and changes



Upcoming events

Recent updates

- Released Chat AI v0.9.4
 - Not all changes visible yet – waiting for infrastructure upgrade
 - Error messages are shown properly and permanently
 - Docx files can now be processed
- New models added
 - Gemma 4 31B Instruct
 - Qwen 3.6 35B A3B
 - OpenAI GPT-5.5 (External)
- Benchmarks complete
 - Result: unfortunately, GLM-5.1 **not sustainable** to provide with current hardware

Recap

- Our priorities are:
 - Long-term support, stability for researchers and users
 - Performance and reliability, increasing scalability of infrastructure and models
- This means we cannot always provide the best models
- We still try to provide alternative solutions for user demands

Upcoming Changes

- Upgrading infrastructure
 - Improve scalability and performance
- Clear error messages when model is out of capacity
- Token-based accounting and institution-wide budgets



Hi ↻ ✎

api_error : No instance available for service 'mock-llm-0'. Try again later.

↻ Try Again

User Profile Settings ✕

Ali Doost Hosseini
GWDG

adoosth
adoosth@gwdg.de

Logout

Monthly Usage ▬ €0.00 / €444.44

Organization Budget Usage ▬ 0.01%

Upcoming Changes

- Retiring older models to free up resources
 - Qwen 3.5 27B
 - Qwen 3.5 35B A3B
 - Gemma 3 27B Instruct
 - Possibly Llama 3.1 8B or Llama 3.3 70B
 - OpenAI GPT-5.3 Chat (Microsoft retires it on June 9th)
- Next model voting round soon

Summary and Overview

- Monthly community meetings, usually last Wednesday at 14:00
 - Event page: <https://gwdg.de/hpc/events/chatai/>
- Model voting takes place in the weeks before community meetings
- Results, decisions, and goals are announced in community meetings
- Join our channels and mailing lists:
 - <https://matrix.to/#/#AIComm:uni-goettingen.de>
 - <https://listserv.gwdg.de/mailman/listinfo/ai-saia-users>
 - <https://listserv.gwdg.de/mailman/listinfo/ai-chat-users>
- GöAID - <https://gwdg.de/hpc/events/goeaid/>

Thank you for your attention