



Institute for Automation and
Applied Informatics (IAI)

HiWi-Tätigkeit

Unterstützung bei dem Aufbau eines Large Language Model (LLM) Servers für die Forschung

Wir sind auf der Suche nach engagierten und technikbegeisterten Studierenden, die aktiv zur Forschungsentwicklung beitragen wollen. Als HiWi unterstützen Sie maßgeblich den Aufbau eines Large Language Models (LLM) Servers, wodurch Sie nicht nur Einblicke in innovative Forschungsprojekte erhalten, sondern auch Ihre technischen Fähigkeiten selbständig weiterentwickeln können. Unser institutsinterner Server ist ausgestattet mit zwei NVIDIA RTX A6000 Grafikkarten, was die Anwendung der neusten und leistungsfähigsten LLMs ermöglicht.

Das Ziel ist es, in den nächsten Monaten ein Chatmodell aufzubauen, welches Datenbanken und das Internet verwendet, um wissenschaftliche Fragen zu beantworten und die Arbeit am Institut zu erleichtern. Damit die Mitglieder unseres Instituts unser Chatmodell verwenden können, soll zusätzlich eine Chat-GPT artige Oberfläche bereitgestellt werden. Diese Aufgabe ermöglicht das Sammeln gefragter Kenntnisse in Software-Engineering, LLMs und Serverentwicklung sowie die Realisierung eigener Ideen auf einem Hochleistungsrechner.

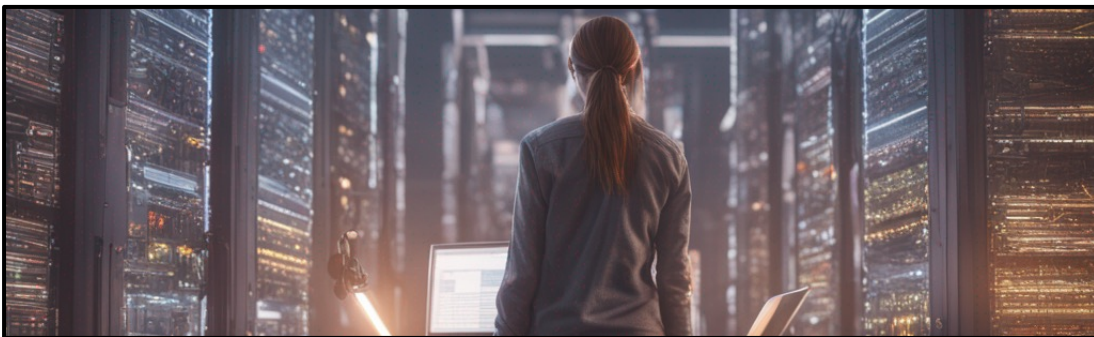


Bild generiert auf einem IAI-Rechner mit [Stable Diffusion](#)

Aufgaben:

- Recherche und Einarbeitung in neuste [LLM Frameworks](#)
- Vergleich von [Modellen](#)
- Generierung von Datensätzen und Training von LLMs
- Frontend Entwicklung einer Chat Oberfläche
- [Prompt Engineering](#)
- Programmierung in Python

Fähigkeiten:

- Starke Affinität zum Programmieren
- Grundlegende Erfahrung mit Python
- Eigenständige Arbeitsweise
- Spaß daran neues zu lernen
- Optional: Erste praktische Deep-Learning-Projekte