



# FAIRD

Eine RAG-AI für die HS-KL und darüber hinaus  
*Ein Exposé*

## Was ist FAIRD?

FAIRD ist eine lokale Chat-Anwendung, die entwickelt wurde, um Large Language Models (LLMs) für die Beantwortung von Fragen zu nutzen. Die Beantwortung kann dabei auch die Inhalte aus selbst hochgeladenen PDF-Dokumenten einbinden. Sowohl der Chat als auch die Dokumente werden dabei datenschutzkonform behandelt.

FAIRD vereinfacht mehrere wesentliche Aufgaben: es speichert Chatverläufe, erleichtert das Hochladen und Bearbeiten von Dokumenten und verwaltet einen lokalen Vektorspeicher. Mittels des Vektorspeichers können aus dem Fundus an Dokumenten die zu einer Frage passenden Dokumente effizient gefunden werden. Ziel des Projektes ist es, auf eine Frage möglichst zügig eine präzise, kontextbezogene Antwort zu liefern.

Das eingesetzte Open-Source-LLM wird auf leistungsstarker Infrastruktur der Hochschule betrieben.

# Datenschutz im Fokus

## ▪ DSGVO-konform

Alle Daten, einschließlich des Chats und der hochgeladenen Dokumente, werden lokal auf Ihrem Rechner gespeichert, so dass Sie die volle Kontrolle über Ihre Informationen haben. Die einzige Übertragung der für eine Anfrage relevanten Daten findet mit dem Aufruf des LLMs statt. Die Daten werden auf dem Server allerdings nicht gespeichert und auch nicht für das künftige Trainieren des Modells genutzt.

## ▪ Zugriffsschutz

Die Server-Infrastruktur ist ausschließlich innerhalb des Netzes der Hochschule Kaiserslautern zugänglich, d.h. FAIRD arbeitet NUR innerhalb dieses Netzes. Dies gewährleistet, dass alle Daten innerhalb der vertrauenswürdigen Umgebung verbleiben.

# Features

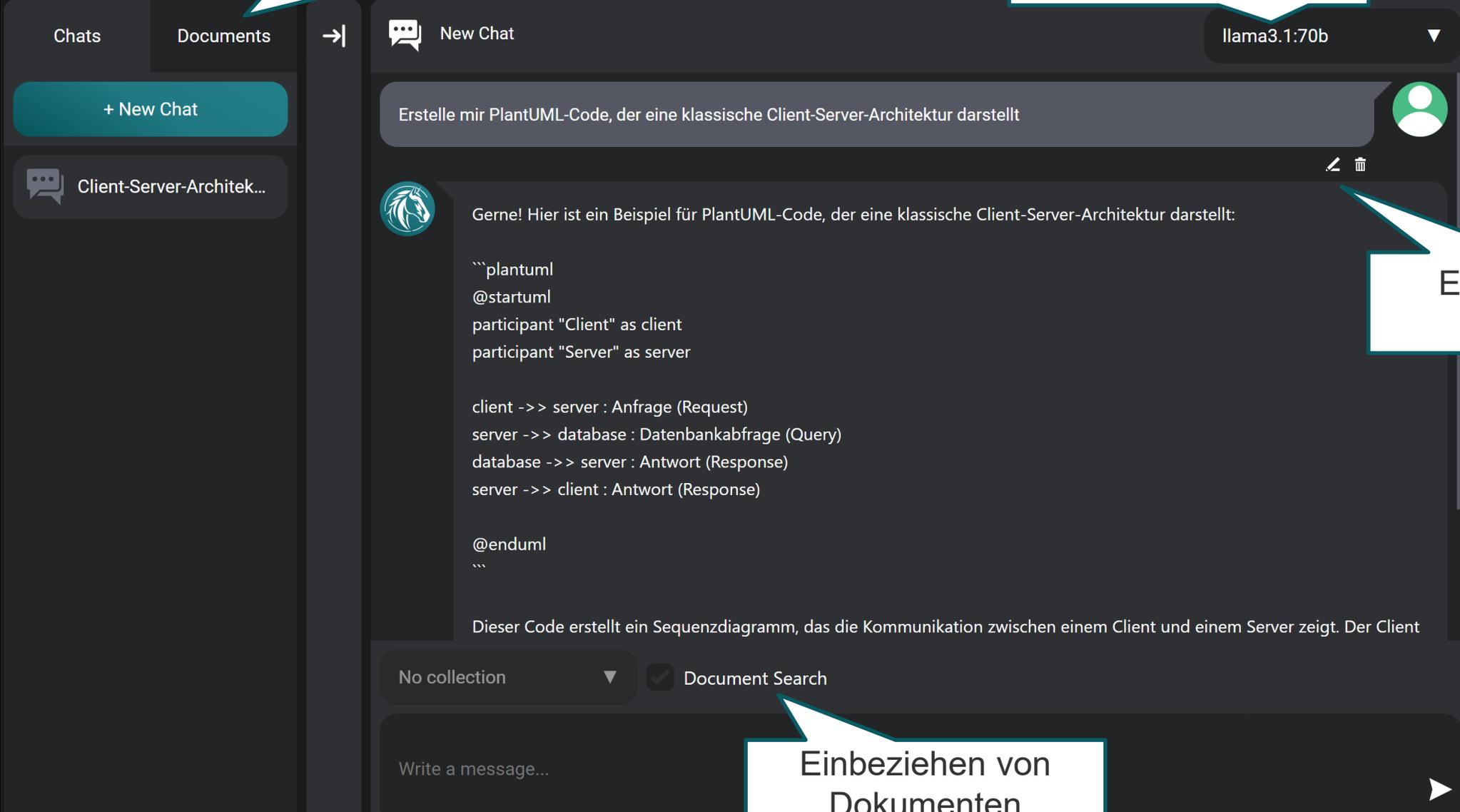
- DSGVO-konform
- Selbst gehostet
- Foundation Modell agnostisch
- Einbindung von PDF-Dokumenten
- Eigenes User-Interface mit der Möglichkeit, Dokumente hochzuladen

## Zukünftige Features

- Headless-Modus zur Integration in Webseiten der Hochschule
- RAG auf Server-Seite mit hochschulspezifischen Inhalten
- Custom FAIRDs – eigene Systemprompts hinterlegen
- Multi-Modalität

Hochladen von Dokumenten

Wählen eines Open-Source-LLMs



Chats Documents → | New Chat llama3.1:70b

+ New Chat

Client-Server-Architek...

Erstelle mir PlantUML-Code, der eine klassische Client-Server-Architektur darstellt

Gerne! Hier ist ein Beispiel für PlantUML-Code, der eine klassische Client-Server-Architektur darstellt:

```
``plantuml
@startuml
participant "Client" as client
participant "Server" as server

client ->> server : Anfrage (Request)
server ->> database : Datenbankabfrage (Query)
database ->> server : Antwort (Response)
server ->> client : Antwort (Response)

@enduml
``
```

Dieser Code erstellt ein Sequenzdiagramm, das die Kommunikation zwischen einem Client und einem Server zeigt. Der Client

No collection Document Search

Write a message...

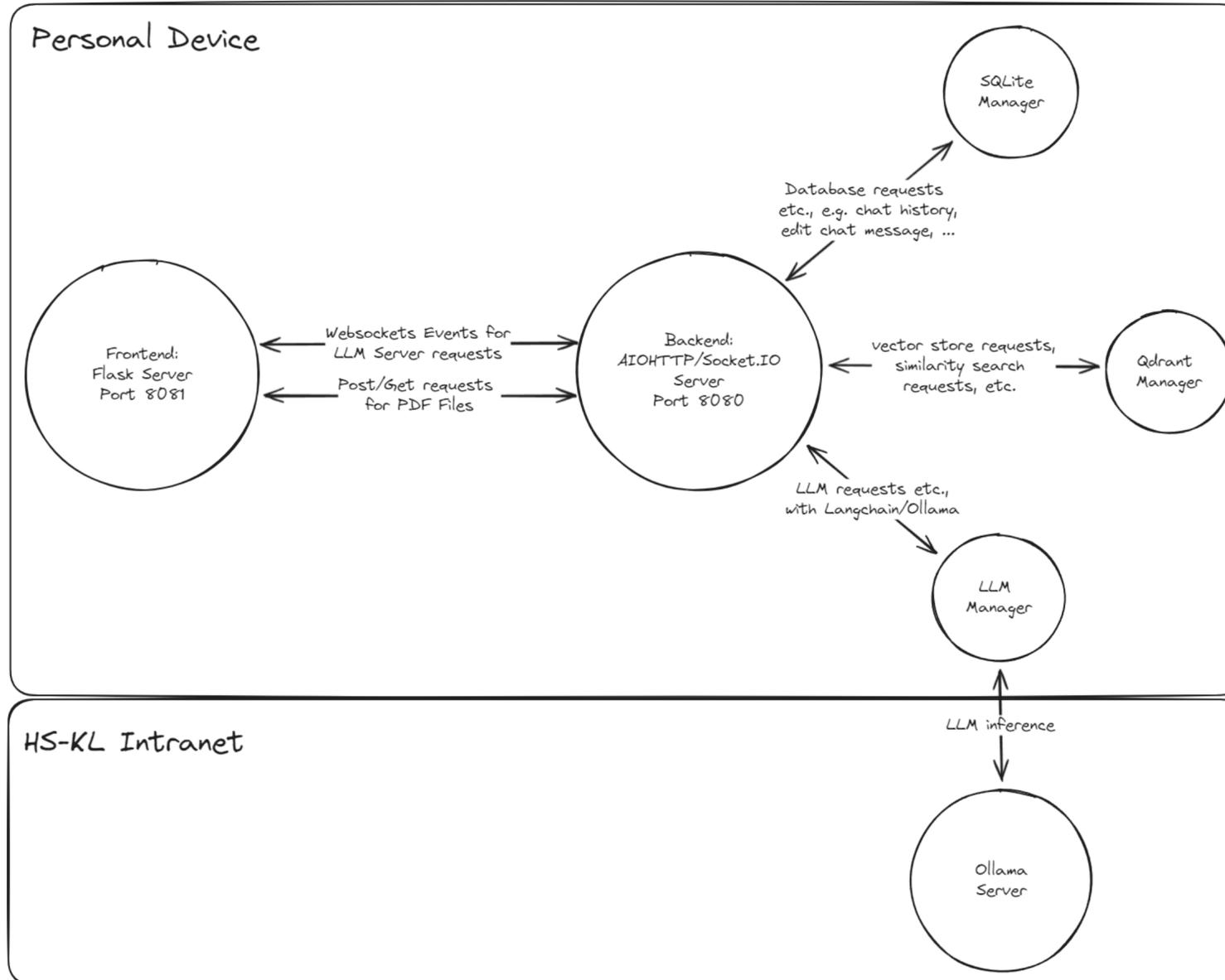
Editieren einer Anfrage

Einbeziehen von Dokumenten

# Eine Auswahl an Use-Cases

- Dozent\*innen
  - Allgemein: GenAI in der Lehre
  - Anfragen zu vergangenen Abschlussarbeiten, ggf. auch mit Sperrvermerk
  - Analyse von Texten
- Studierende
  - Anfragen zu Prüfungsordnungen
  - Feedback zu eigenen Texten
- Potenzielle Bewerber
  - Informationen zu Studiengängen
  - Studiengangsempfehlungen

# Architektur



# Team

- Contributors
  - Eric Gaida (development lead)
  - Nicholas Grund
  - Nicolas Sand
  - Niklas Borresch
  - Jens Müller
- Projektleitung
  - Prof. Dr. Jan Conrad
  - Prof. Dr. Eugen Staab



# FAIRD

Eine RAG-AI für die HS-KL und darüber hinaus  
*Ein Exposé*