

# FIONA Chatbot Application

Said Benaissa

April 29, 2025

# Preventing Hallucinations and Providing Evidence

- ▶ FIONA gets information directly from preloaded documents saved in the backend.
- ▶ It uses embedding search with cosine similarity to find the right document parts.
- ▶ Answers are created by the `flan-t5-large` model and include links to the source documents.
- ▶ Hallucinations are avoided by only using the document data and not guessing.

# CI/CD Pipeline for Document Changes

- ▶ Watches for changes in the code and documents.
- ▶ Automatically updates embeddings when documents are changed by using `embedding_manager.py`.
- ▶ Re-deploys the app to always use the latest content and features.
- ▶ Makes it easy to add new information to the chatbot.

# Deployment on Linux/Debian-based Systems

- ▶ Uses Docker Compose to manage backend and frontend services.
- ▶ Backend runs on `http://localhost:8000`, frontend on `http://localhost:3000`.
- ▶ Can scale from small to large systems.
- ▶ Fully on-premises deployment to keep data private.

# Real-time User Interface

- ▶ Simple and modern design using React.
- ▶ Shows answers immediately in normal human language.
- ▶ Shows document sources to build trust.
- ▶ Responsive design works well on all devices.

Comment structured project de recherche?

We prepared a questionnaire to assist you in setting up your project. Please note that only projects willing to use electronic data capture will be considered. The electronic data capture system is provided as a service of Helse Vest RHF, and therefore, at least one user must be a member of any of the hospitals in the Western Norway health region. The majority of the provisions in this document are in place to comply with requirements for collecting sensitive human subject data.

comment preparer project de recherche

# Integration with Apache2 and PHP 8.x

- ▶ Backend can be connected to Apache2 as a reverse proxy.
- ▶ Frontend is served as static files that work with PHP servers.
- ▶ Fits into existing enterprise systems using PHP 8.x.
- ▶ Supports old systems while still using modern technology.