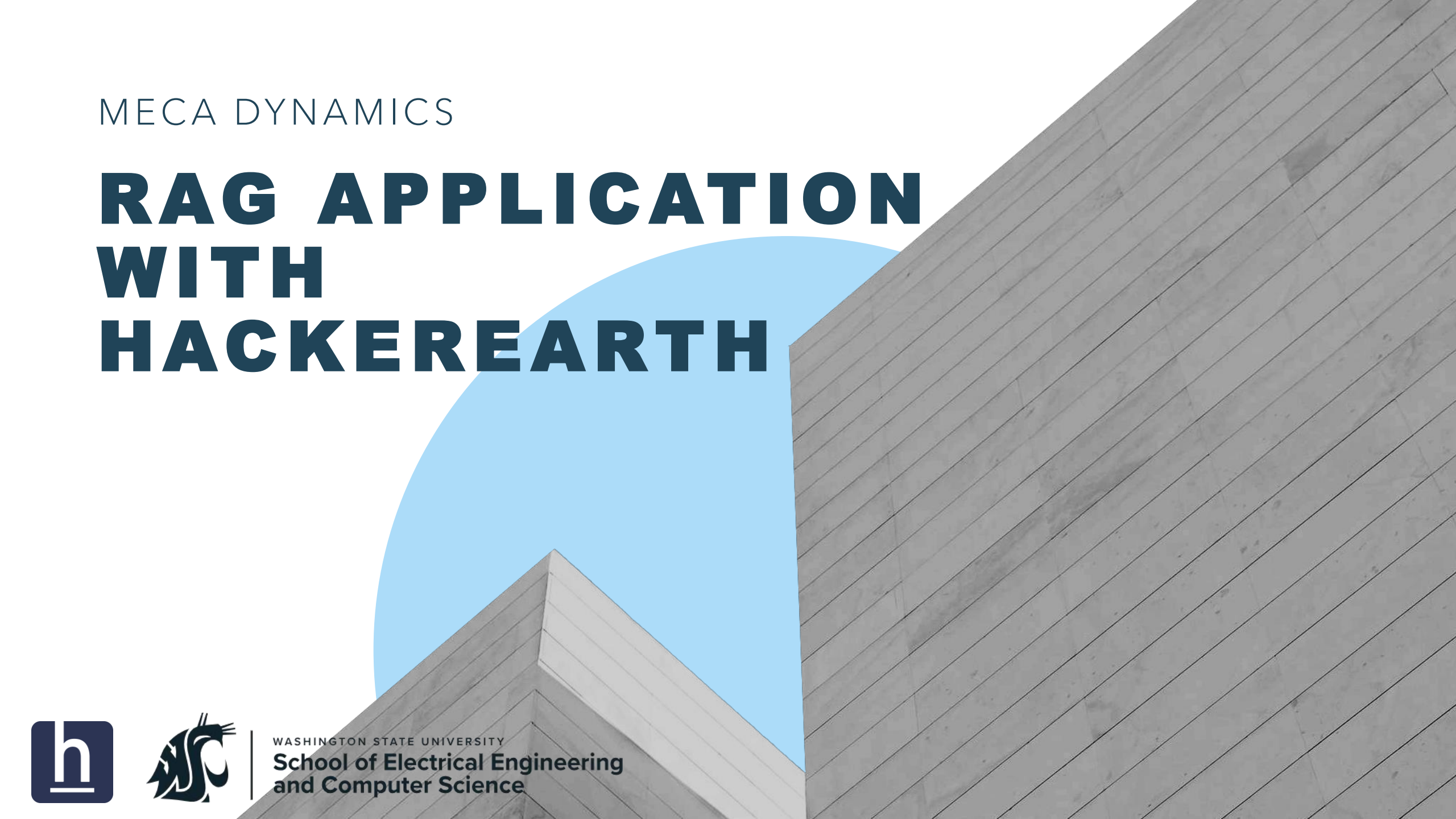
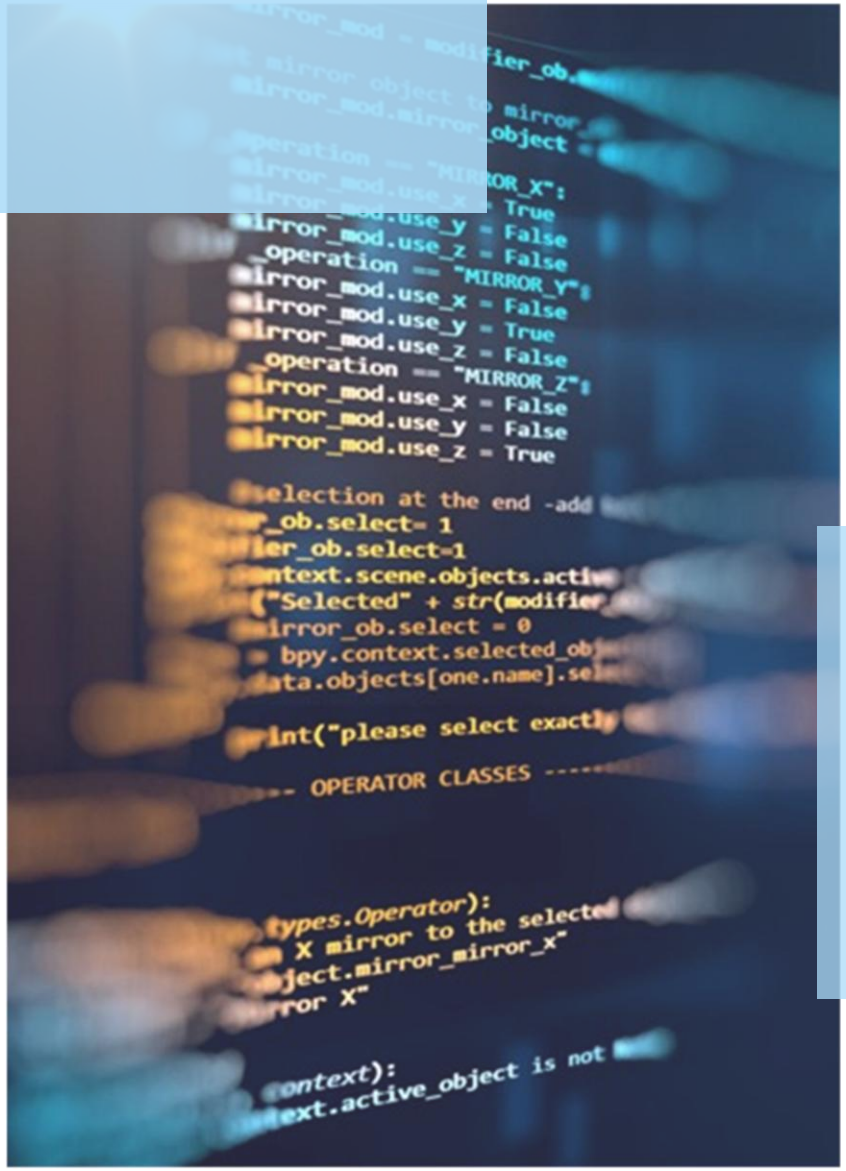
MECA DYNAMICS

RAG APPLICATION WITH HACKEREARTH



WASHINGTON STATE UNIVERSITY
School of Electrical Engineering and Computer Science



AGENDA

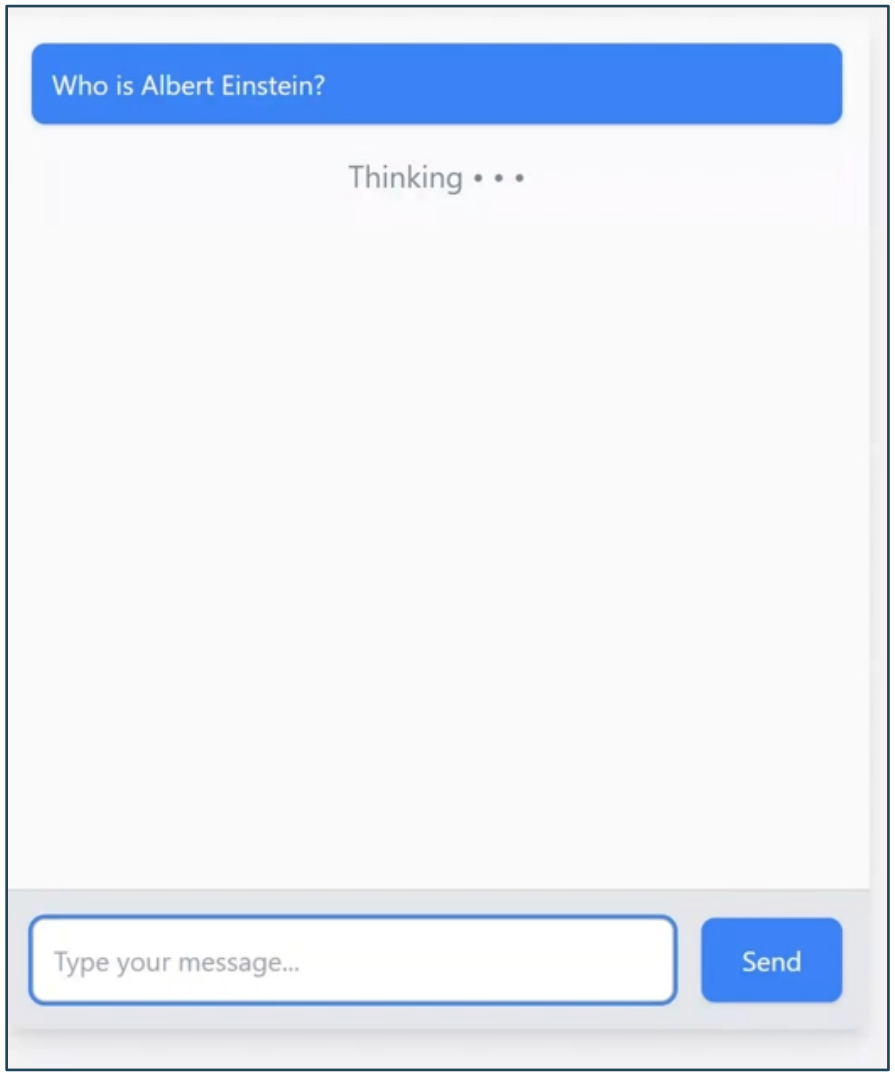
Project Overview

App Demo

Our Solution

Challenges

Future Improvements



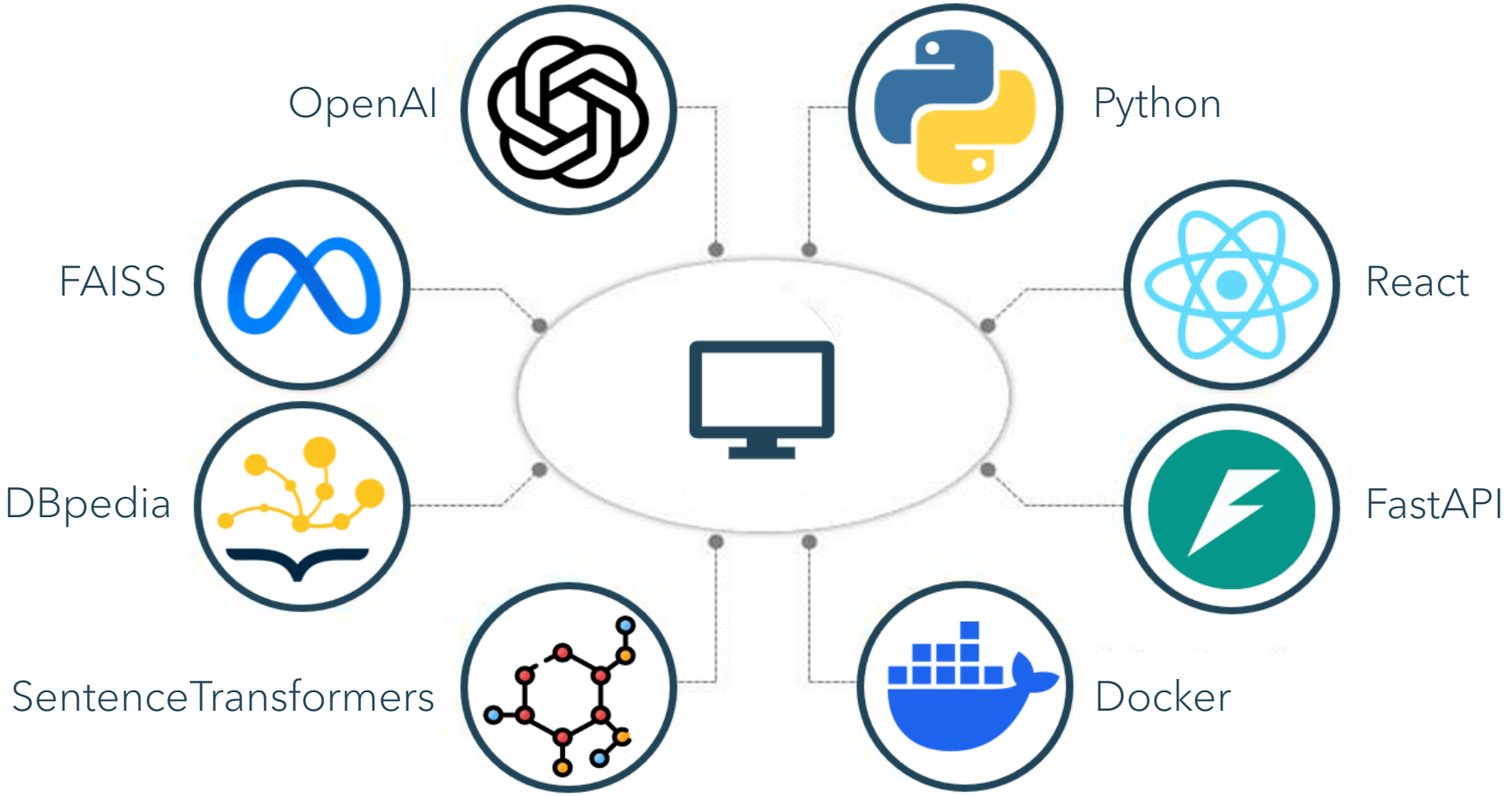
PROJECT OVERVIEW

- RAG application using Wikipedia as a knowledge base
- Combines Knowledge Graph + Vector Search
- Chatbox UI for user questions and answers
- Custom dataset: class notes integration
- Deployed using Docker

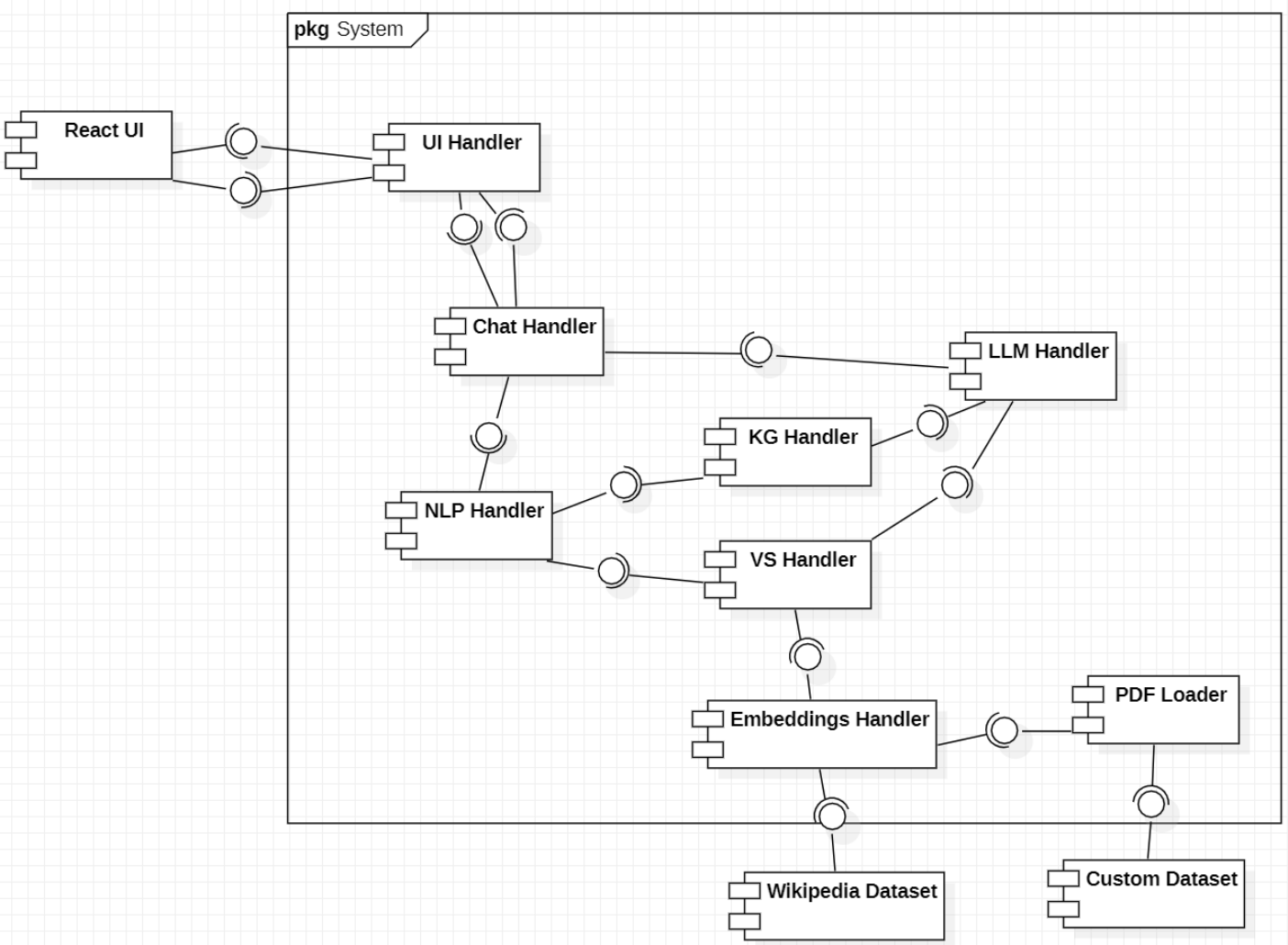
APP DEMO

OUR SOLUTION

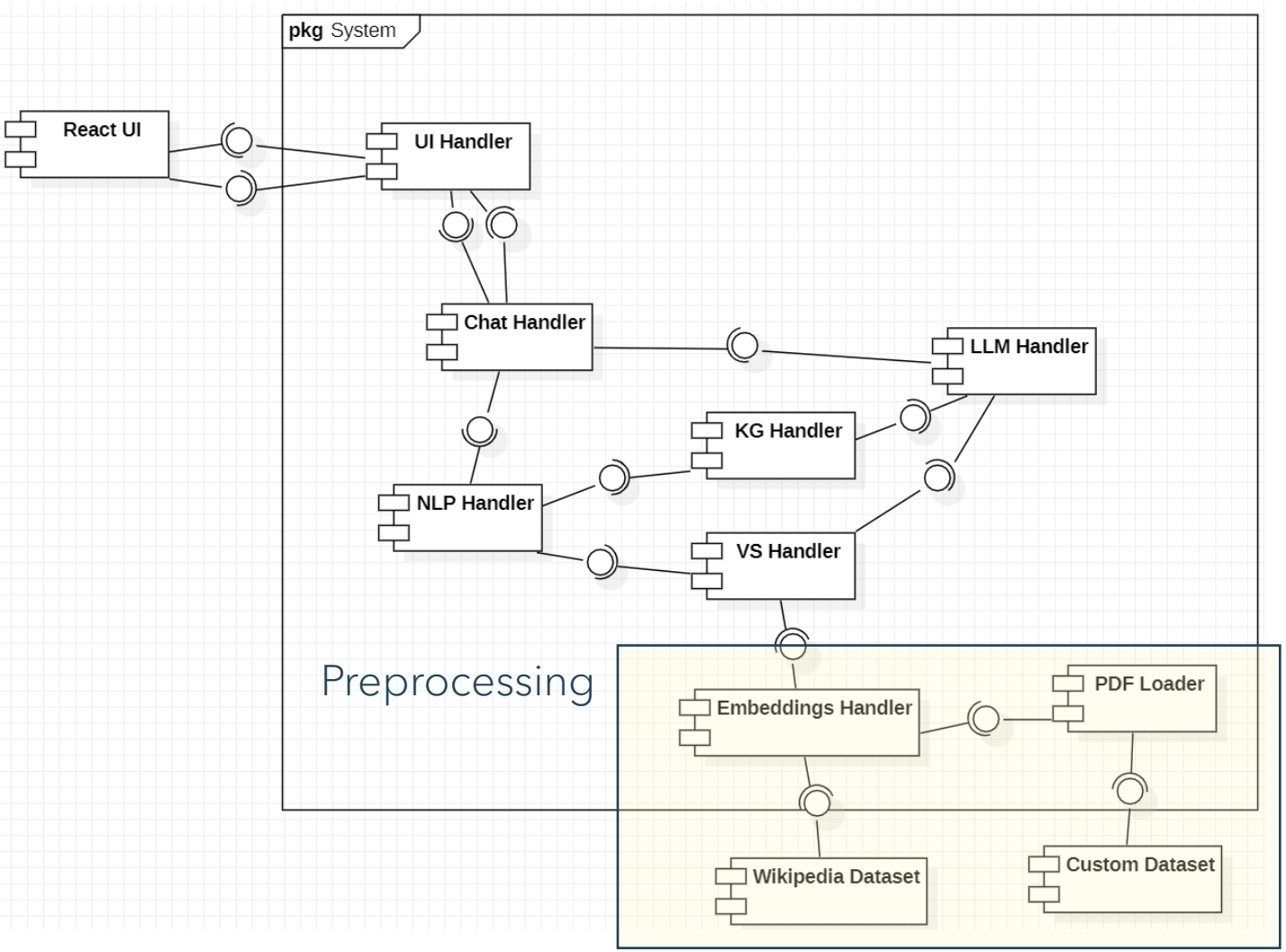
TECH STACK


OpenAI
Python
FAISS
React
DBpedia
FastAPI
SentenceTransformers
Docker

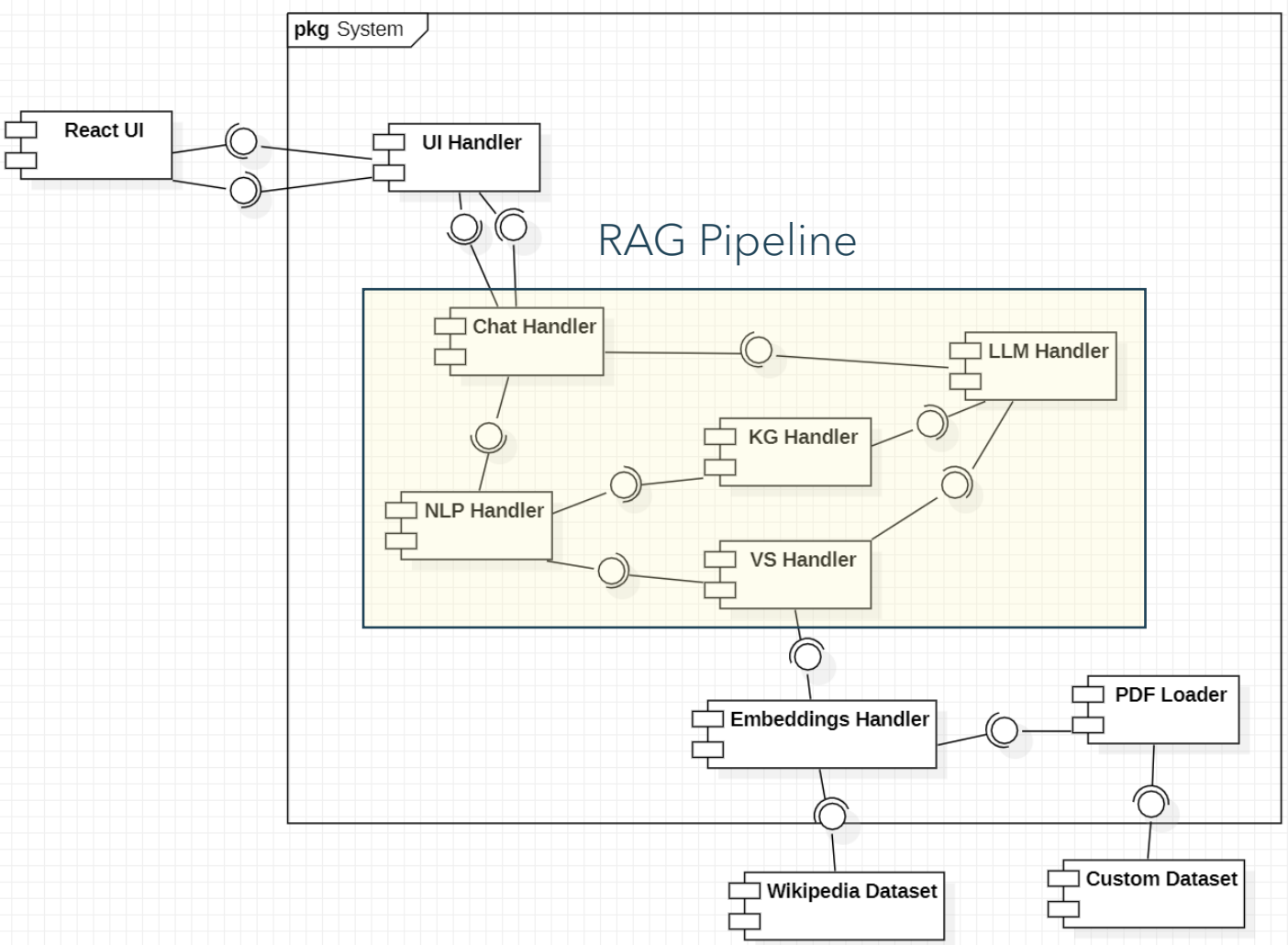
HOW IT WORKS


pkg System
React UI
UI Handler
Chat Handler
LLM Handler
KG Handler
NLP Handler
VS Handler
Embeddings Handler
PDF Loader
Wikipedia Dataset
Custom Dataset

HOW IT WORKS

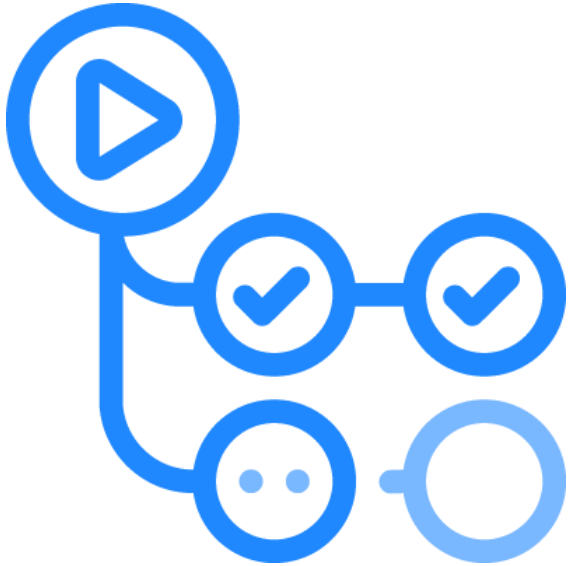

pkg System
React UI
UI Handler
Chat Handler
LLM Handler
KG Handler
NLP Handler
VS Handler
Preprocessing
Embeddings Handler
PDF Loader
Wikipedia Dataset
Custom Dataset

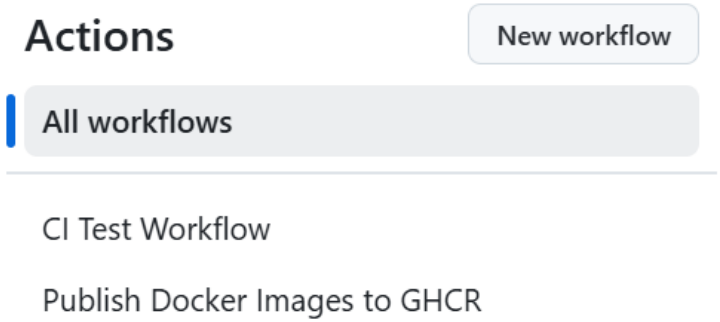
HOW IT WORKS


pkg System
React UI
UI Handler
RAG Pipeline
Chat Handler
LLM Handler
KG Handler
NLP Handler
VS Handler
Embeddings Handler
PDF Loader
Wikipedia Dataset
Custom Dataset

TESTING & PERFORMANCE

- GitHub Actions for testing + deployment
- 16 unit tests across key components
- Avg. response time: ~8 seconds





CHALLENGES

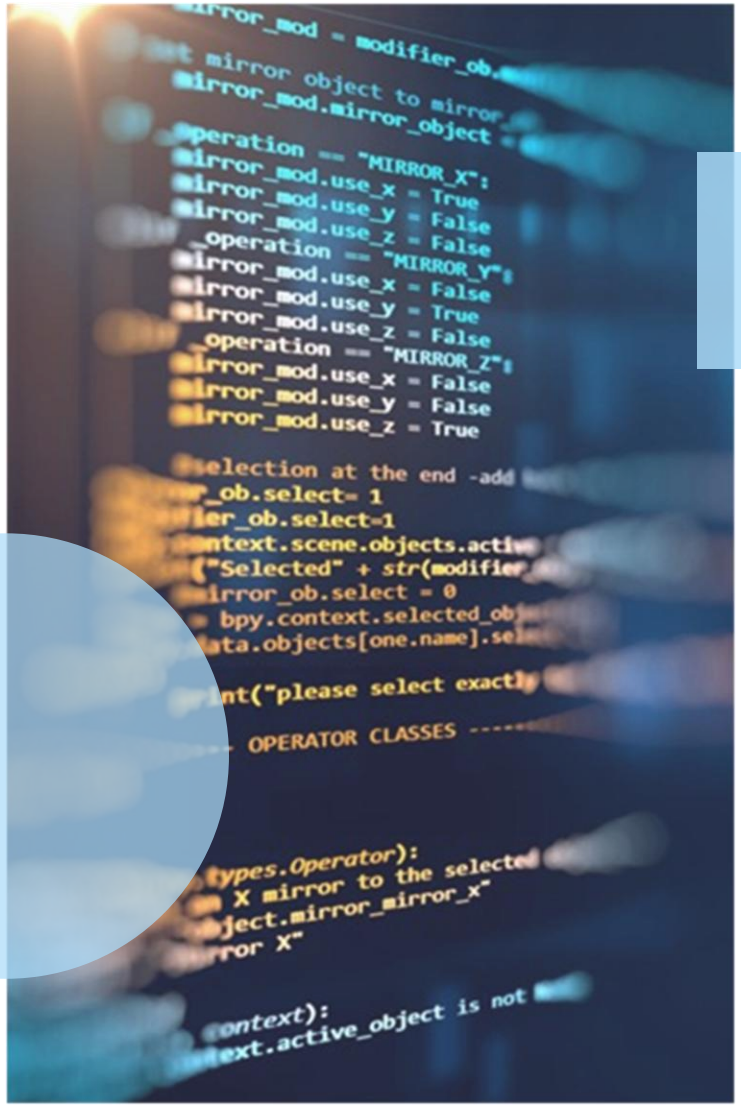
VECTOR SEARCH & EMBEDDINGS

Vector Search

- Worked with small data
- Struggled with large embeddings
- Switched to SentenceTransformers + sentence-based splitting

Embedding Generation

- 2-3 GB file
- Required high computing power
- Regenerated multiple times during tuning



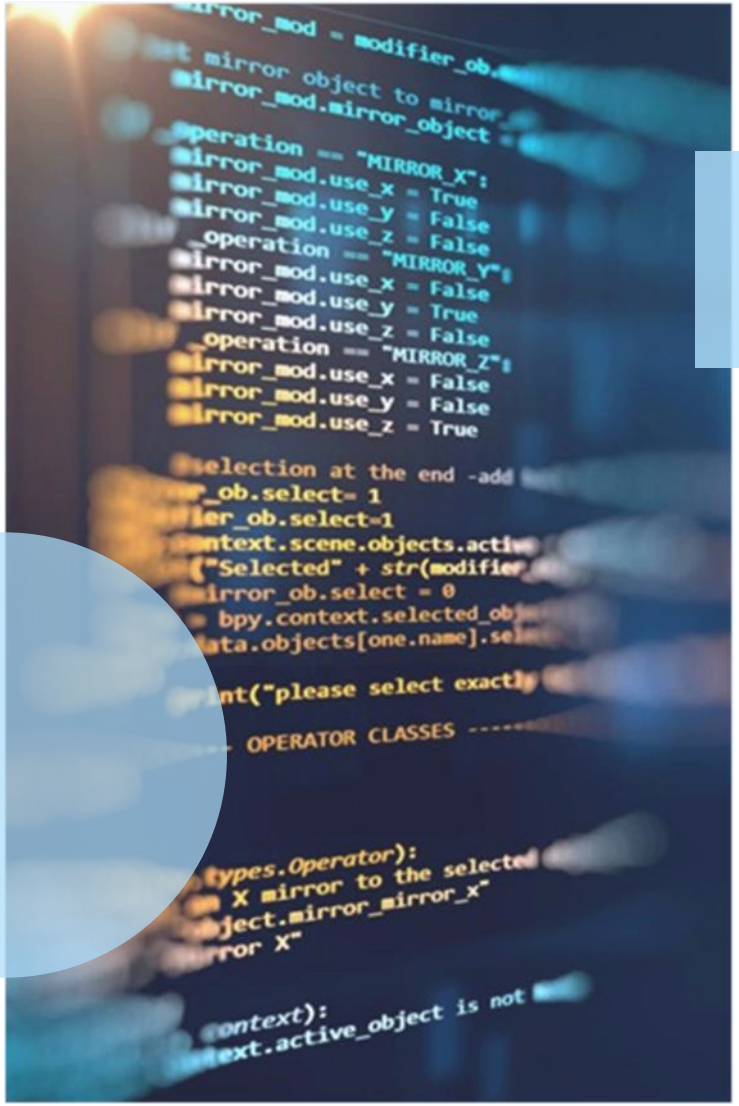
LLM
& QUERY HANDLING

LLM Performance

- DistilBERT: fast but weak
- LLaMA 2.7B: accurate but slow
- OpenAI: fast + accurate

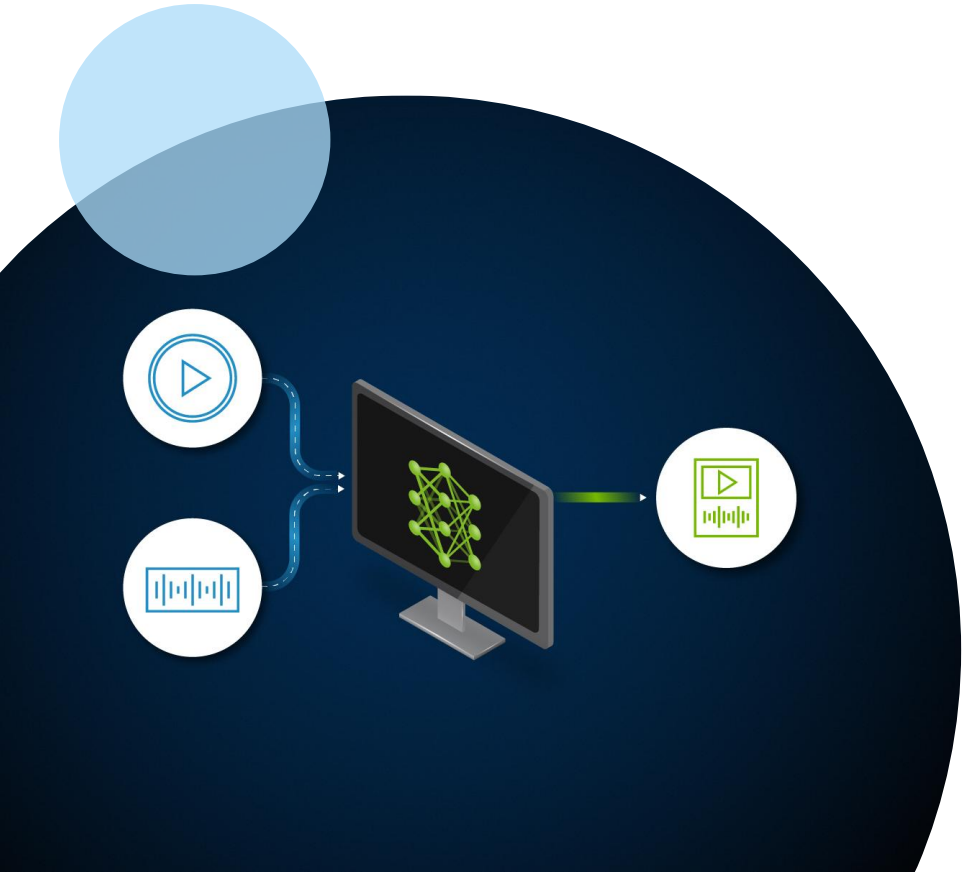
Query Handling

- Recovers from missing KG or VS results
- Rejects harmful or unsafe questions



FUTURE IMPROVEMENTS

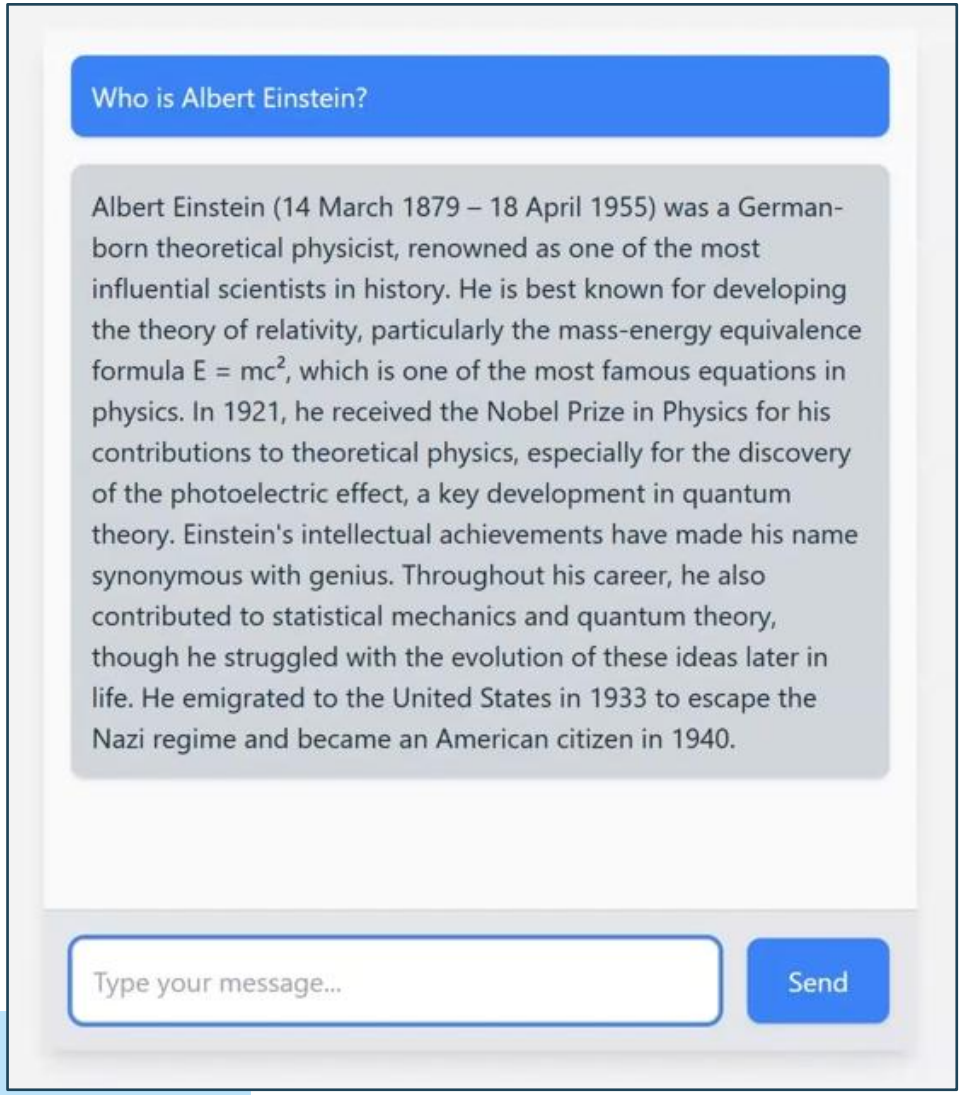
FUTURE WORK



- Improve NLP for complex queries + moderation
- Add more domain-specific datasets (e.g., health, finance)
- Support multimodal input (text, images, diagrams)
- Enable saved chats, user login, and memory

CONCLUSION

- Built a full RAG application using knowledge graph and vector search
- Supports large-scale Wikipedia data and custom class note datasets
- Deployed with Docker and tested with GitHub Actions
- Reliable performance with clear paths for future expansion



THANK YOU!

Any questions?