

Knowledge Assistant

KI-gestuetzter Versicherungsberater mit RAG-Technologie

SYSTEMARCHITEKTUR

PRESENTATION

Streamlit Web UI



APPLICATION

Chat Engine | Schadensmeldung Bot



AI / RAG LAYER

OpenAI GPT-4 | Vector Search



DATA LAYER

ChromaDB | Document Store

Retrieval-Augmented Generation

Dokumente werden in semantische Chunks aufgeteilt und als Vektoren gespeichert. Bei Anfragen werden relevante Chunks dem LLM als Kontext bereitgestellt.

KERNFUNKTIONEN

Ch

Intelligenter Chat

KI-Berater mit Quellenangaben

Bo

Schadensmeldung Bot

Gefuehrter Dialog + Foto-Upload

Do

Wissensbasis

Dokumente automatisch indexiert

Se

Semantic Search

Versteht Bedeutung, nicht nur Keywords

Ad

Admin Dashboard

System-Status und Statistiken

TECHNOLOGY STACK

Frontend

Streamlit

Backend

Python 3.11+

Vector DB

ChromaDB

LLM

OpenAI GPT-4

Embeddings

text-embedding-3-small

PERFORMANCE

Context Window

128K Tokens

Embedding Dim

1536

Chunk Size

1000 Tokens

Top-K Retrieval

10 Chunks

12

Versicherungs-
produkte

30

Knowledge
Chunks

8

Schadens-
typen

5

Wissens-
basen

ANWENDUNGSZWECKE:

1. Kundenberatung

2. Schadenserfassung

3. Self-Service Portal

4. Onboarding & Training