**ChatGPT:**

Na klar, Manfred 😄 – hier kommt die **komplette Installationsübersicht**, damit du **das ganze Projekt inkl. Ollama (LLaMA), Chroma, FAISS & LangChain** auch sauber auf deinem neuen Laptop zum Laufen bekommst.

**✅ 1. Grundvoraussetzungen**

**🔹 Python (mind. 3.9 empfohlen)**

Installiere Python, falls noch nicht vorhanden:  
👉 <https://www.python.org/downloads/>

**✅ 2. Ollama installieren (für LLaMA lokal)**

**🔹 Schritt 1: Ollama herunterladen**

👉 https://ollama.com/download  
Installiere es und starte Ollama danach einmal.

**🔹 Schritt 2: Ein LLM-Modell laden**

Z. B. das schnelle mistral-Modell:

bash

KopierenBearbeiten

ollama run mistral

Oder das fortgeschrittene llama3-Modell:

bash

KopierenBearbeiten

ollama pull llama3

**✅ 3. Python-Umgebung einrichten**

Du kannst es direkt installieren oder ein venv (virtuelle Umgebung) nutzen.

**🔹 Wichtigste Python-Pakete installieren:**

bash

KopierenBearbeiten

pip install chromadb faiss-cpu langchain openai langchain-community

Optional (je nach Bedarf):

bash

KopierenBearbeiten

pip install python-dotenv pandas

**✅ 4. Chat-Projekt vorbereiten**

**Deine lokalen Dateien brauchst du ebenfalls:**

* Das Skript chat\_agent\_v\_4\_7.py
* Deine conversations.json
* Deine chat\_infos.xlsx
* Die MySQL-Datenbankverbindung (Zugangsdaten)

Falls du eine .env-Datei nutzt, leg diese mit deinen Zugangsdaten an:

dotenv

KopierenBearbeiten

MYSQL\_HOST=localhost

MYSQL\_USER=deinuser

MYSQL\_PASSWORD=deinpass

MYSQL\_DATABASE=deinedb

**✅ 5. Speicherort für Embeddings**

In deinem Fall:

bash

KopierenBearbeiten

D:\Users\doman\Documents\OneDrive\Dokumente\Programmierung\Projekte\AiAgents\chats

Stelle sicher, dass der Ordner existiert.

**✅ 6. Starttest**

Teste z. B. ob du Embeddings erzeugen kannst (wird im nächsten Schritt mit mir gemacht):

python

KopierenBearbeiten

from langchain\_community.embeddings import OllamaEmbeddings

embeddings = OllamaEmbeddings(model="llama3")

vector = embeddings.embed\_query("Hallo Welt")

print(vector[:5])

Wenn das klappt – läuft alles 💡

**❗Zusammenfassung der wichtigsten Tools**

| **Tool** | **Zweck** |
| --- | --- |
| **Ollama** | Lokaler LLM-Server für Mistral/LLaMA |
| **LangChain** | Framework zum Verwalten von Embeddings etc. |
| **Chroma/FAISS** | Vektor-Datenbanken für Ähnlichkeitssuche |
| **Python + Pip** | Grundgerüst zur Ausführung & Automatisierung |
| **MySQL** | Deine bestehende Datenbank |

pip install langchain chromadb faiss-cpu

pip install openai

pip install langchain-community

pip install ollama

ollama serve

ollama run mistral