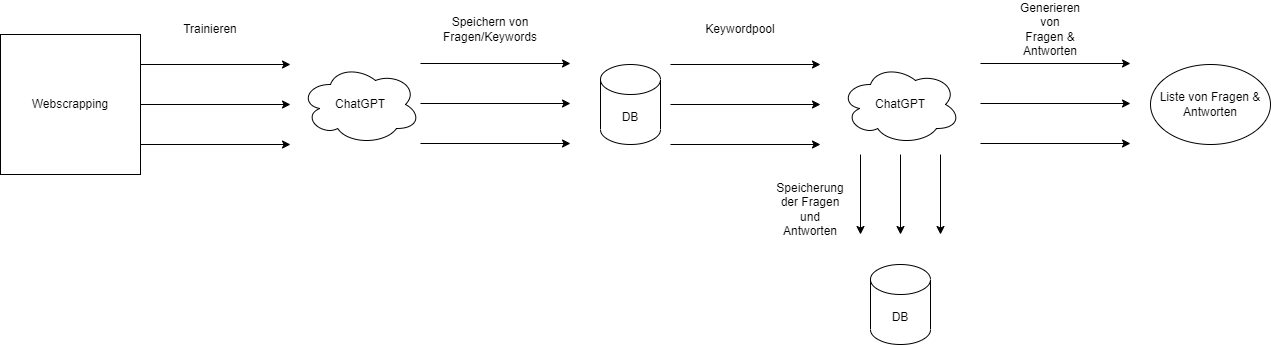
FAQ-Projekt

Abschlussprojekt

# Schautbild



# Ablauf

## Webscrapping

* Webscrapping aufgewählter Domänen unter „svlfg.de“
* vorkommende PDF’s mitziehen
* gescrapte Daten in PDF-Format bringen

Scrapping:

Aufgrund des hohen Aufwandes werden nur eine ausgewählte Anzahl von Domänen gescrapt um im Rahmen der Projektzeit zu bleiben.

## ChatGPT anlernen

* PDF’s über Plugin in ChatGPT laden und Session damit „trainieren“
* Plugin nicht möglich, Session „trainieren“ durch gescrapte Daten mit angepasstem Prompt

Angepasste Prompt:

Prompt gestalten, sodass ChatGPT nach dem „Einleitungsprompt“, nur Keywords als Antwort von Fragen zurückgibt und eine Antwort auf die Frage in den „gelernten“ gescrapten Daten“sucht

## Speicherung der Daten

* Ausgaben von ChatGPT in Datenbank speichern
* Datenbankkonzept:

Conversation\_table:

conversation\_ID(PK), session\_ID(FK), Fragen\_ID(FK), Antwort\_ID(FK)

Session\_table:

Session\_ID(PK), Session\_Name

Antwort\_table:

Antwort\_ID(PK), Antworttext, Fragen\_ID(FK)

Fragen\_table:

Fragen\_ID(PK), Frage

Keyword\_table:

Keyword\_ID(PK), Keyword

Fragen\_Keyword\_table:

Fragen\_Keyword\_ID(PK), Fragen\_ID(FK), Keyword\_ID(FK)

Datenbank:

Tokenausgabe von ChatGPT reduzieren aufgrund der Kosten und als mögliche Schnittstelle, falls Daten anders genutzt werden können.

Conversation\_table soll Nachrichten enthalten, die vom User(Fragen) und ChatGPT(Keywords/Antworten) enthalten.

Session\_table soll als Chatroombezeichnung gelten, kann vernachlässigt werden falls alles in einem Chat möglich ist.

## Fragengenerierung

* Erstellen von Fragen durch ChatGPT anhand des Keywordpools

Fragengenerieren:

Fragen sollen anhand des Keywordpools erstellt werden.

Generierung kann über bisher gespeicherte Keywords in der Datenbank geschehen oder falls alles in einer Session möglich ist, aus bisheriger Session.

## Antwortgenerierung

* Erstellen von Antworten durch vorher erstellte Fragen durch ChatGPT

Antwortgenerieren:

Antworten werden basierend auf den „trainierten“ Daten erstellt, falls keine Antwort zum Thema möglich ist, wird „N/A“ in Datenbank eingetragen.