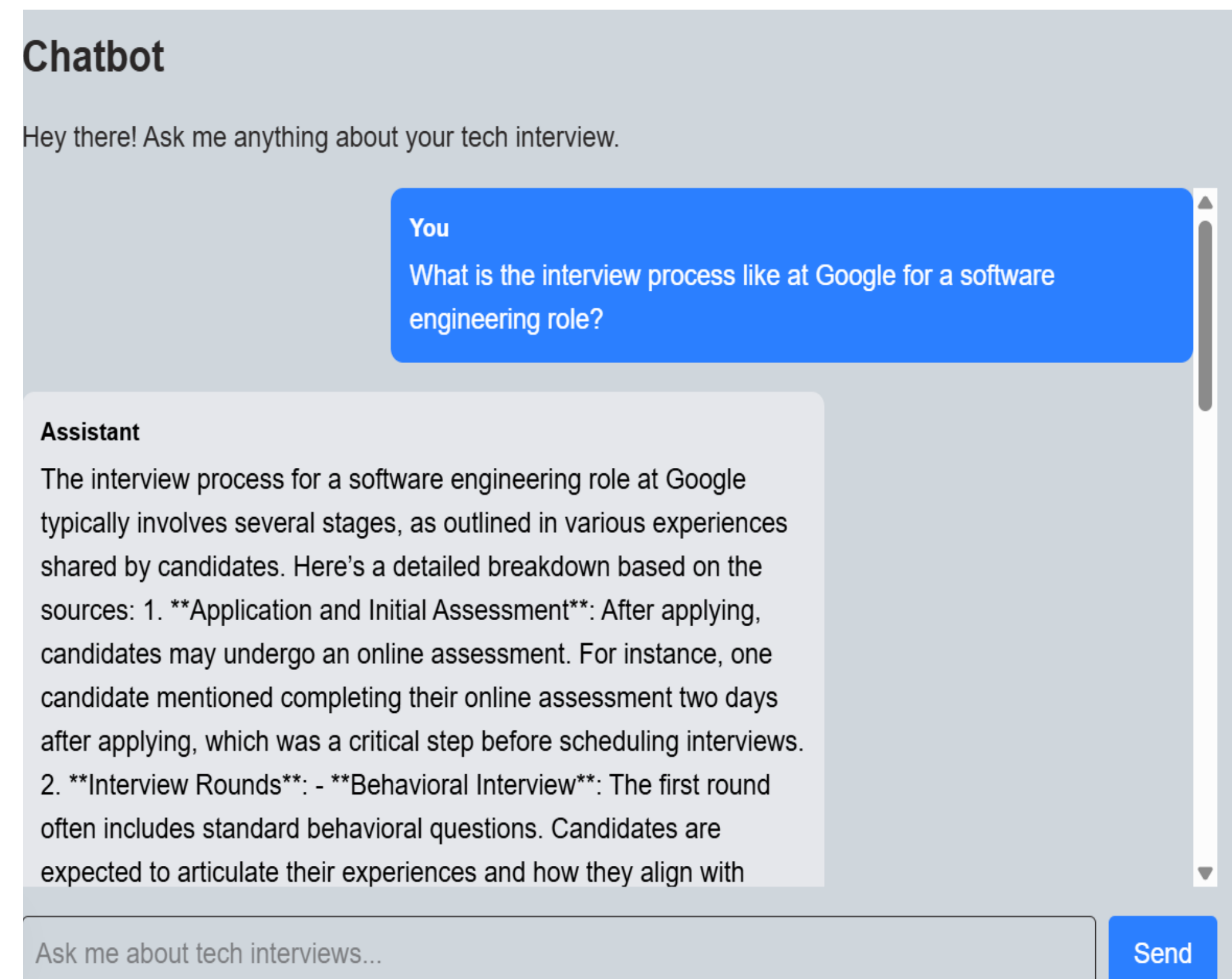


# JobTutor: A RAG-Based Chatbot and ILP-Inspired Question Recommender for Technical Interview Preparation

Mastewal A. Berhe | Computer Science | Advised by Sofia Visa

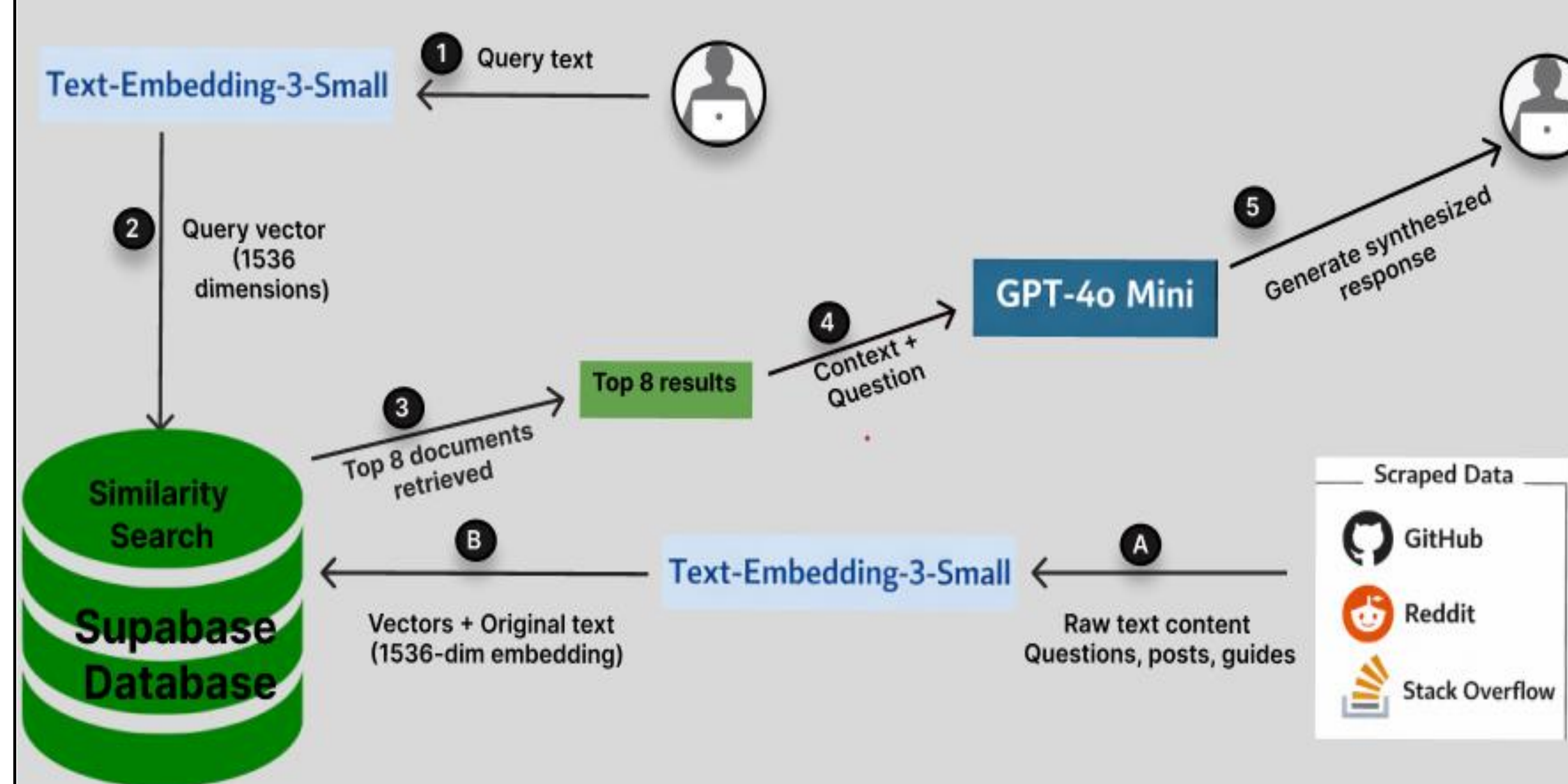
## JobTutor Overview

JobTutor is a web-based technical interview preparation platform built with Next.js. It combines two AI-driven features accessible through a single interface:



The RAG chatbot answers open-ended interview questions by retrieving relevant content from the knowledge base.

## RAG PIPELINE ARCHITECTURE



### RAG Configuration:

- Similarity threshold: 0.3
- Top 5 documents sent to GPT-4o-mini
- System role: *"expert technical interview preparation assistant"*
- Out-of-scope queries are declined

## ILP-INSPIRED GREEDY QUESTION RECOMMENDER

The recommender selects the optimal set of practice questions per session using a greedy scoring function:

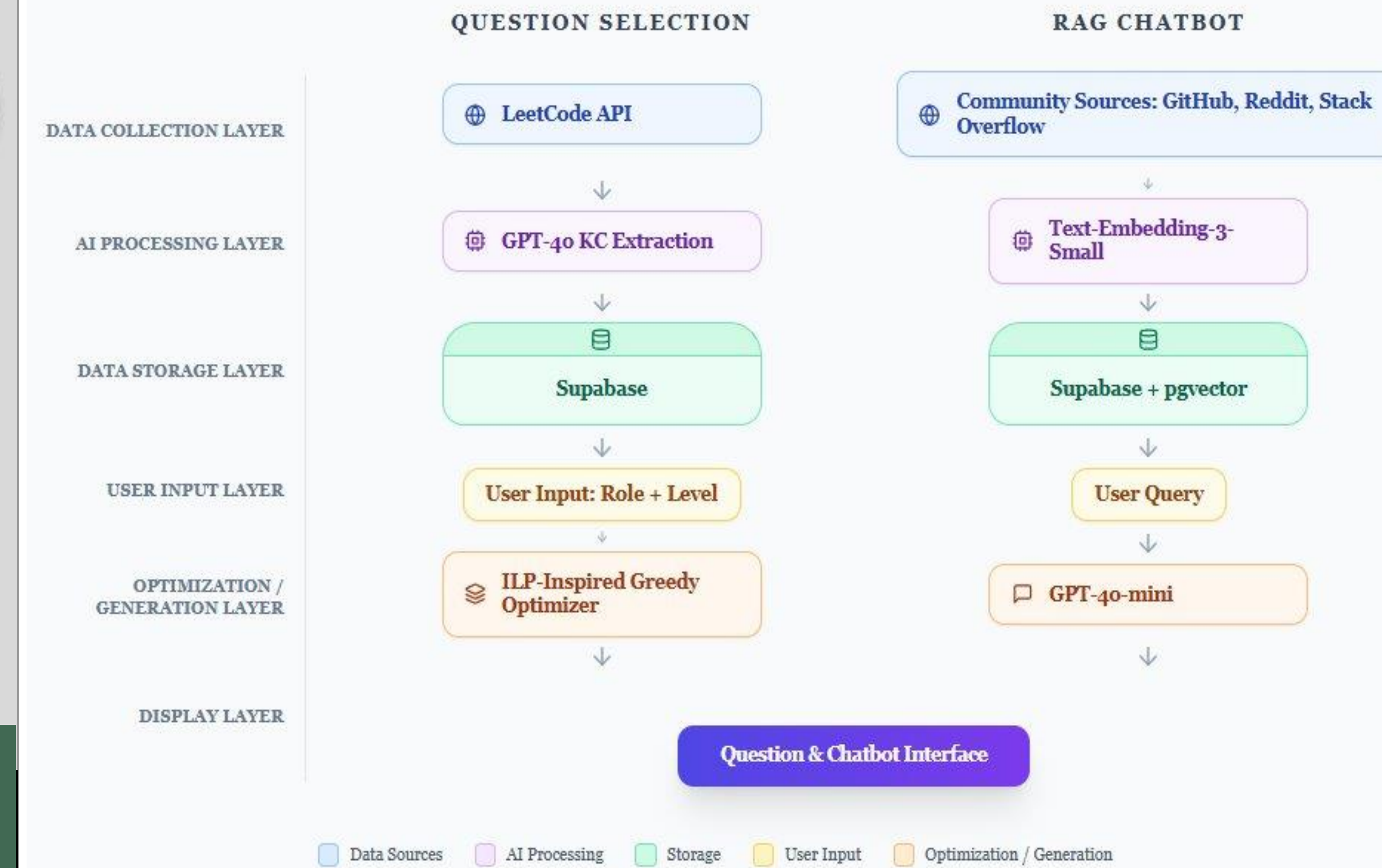
$$\text{Score}(q) = 5 \cdot f_1(q) + 2 \cdot f_2(q) + 1 \cdot f_3(q)$$

- $f_1(q)$  **Concept Coverage** (weight: 5): Counts new KCs not yet covered
- $f_2(q)$  **Depth** (weight: 2): Sums shared KC weights with already selected questions
- $f_3(q)$  **Breadth** (weight: 1): Counts selected questions NOT connected to candidate

### Constraints

- C1** Session must contain exactly 15 questions
- C2** **Difficulty Balance:** Max 5 hard questions and at least 2 easy questions

## System Architecture



## Data Sources

| Source         | Purpose                       | Result        |
|----------------|-------------------------------|---------------|
| LeetCode       | Question bank + KC extraction | 137 questions |
| Stack Overflow | RAG knowledge base            | 165 Q&A pairs |
| Reddit         | RAG knowledge base            | 118 posts     |
| GitHub         | RAG knowledge base            | 402 sections  |

## RAGAS Evaluation

| Metric            | Score | What it means                              |
|-------------------|-------|--|
| Answer Relevancy  | 0.89  | Responses directly address the question    |
| Faithfulness      | 0.88  | Answers grounded in retrieved sources      |
| Context Recall    | 0.83  | Most necessary info successfully retrieved |
| Context Precision | 0.65  | Some retrieved docs less relevant          |

**Conclusions:** JobTutor demonstrates that ILP-inspired greedy optimization and RAG can be combined into a practical, real-time interview prep platform.

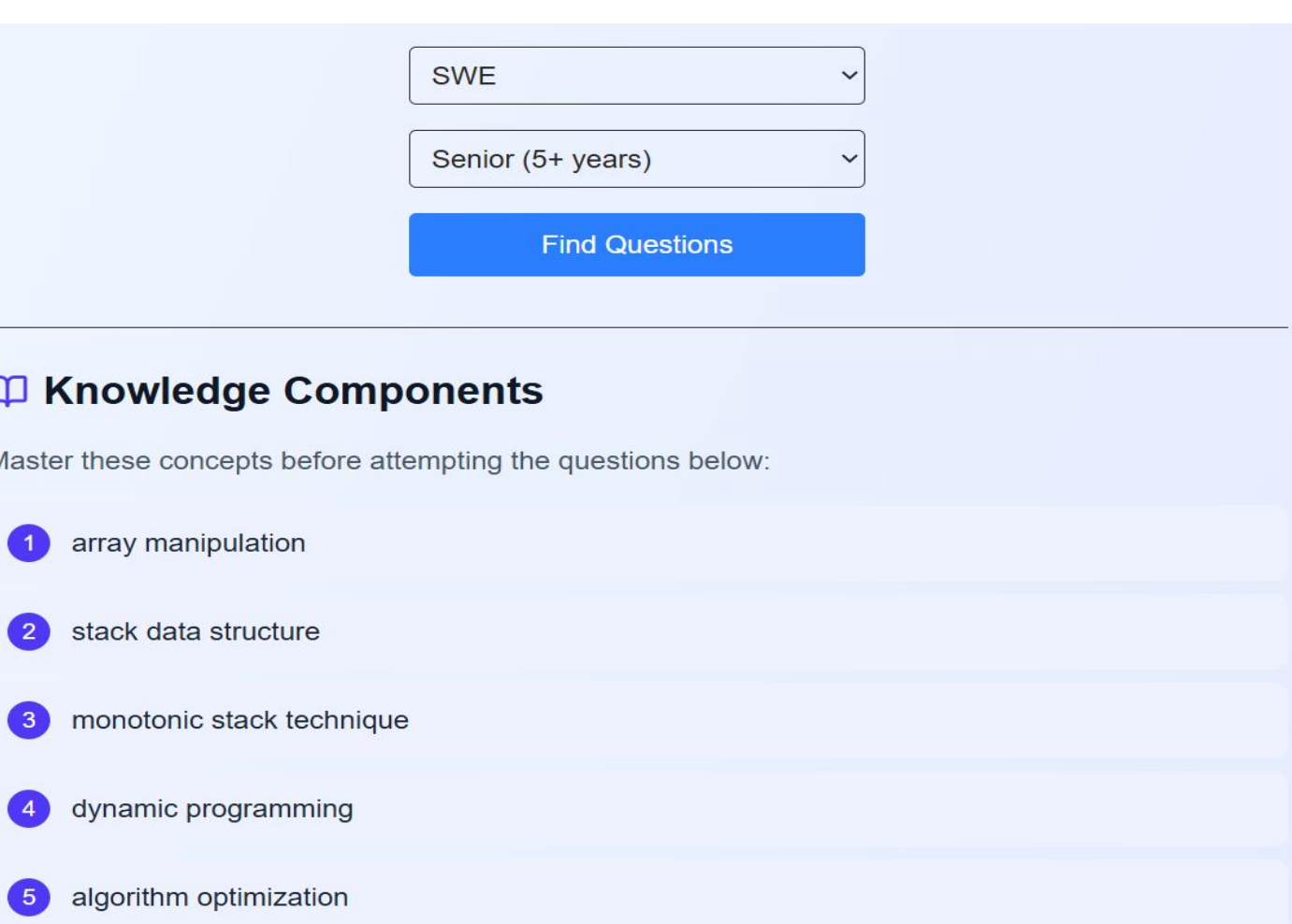
**Future Work:** Reranking for retrieval precision · Conversation memory · Dynamic knowledge base updates · Full KT mastery integration

Responses include cited sources (Reddit posts and GitHub resources) showing transparency in retrieval.

(<https://www.pramp.com/>) and (<https://interviewing.io/>) can be beneficial. This overview is based on shared experiences from candidates on Reddit and insights from GitHub resources.

**Sources:**  
[reddit — I got a SWE job at Google with outstanding interview performance and wanna h...](#)  
[reddit — Google Interview Experience SWE](#)  
[reddit — Google L6 System Design Interview](#)  
[github\\_resources — coding-interview-university - Interview Process & General Interview ...](#)  
[github\\_resources — coding-interview-university - Have questions for the interviewer](#)

Users select their role (e.g., SWE) and experience level. The system shows the top 5 Knowledge Components (KCs) to master.



### Recommended Questions (5)

Practice your skills with these curated questions:

- Q1** Given an array of integers heights representing the histogram's bar height where the width of ea...  
 Difficulty: **hard**  
 Skills Needed: [array manipulation](#) [stack data structure](#) [monotonic stack technique](#) [dynamic programming](#) [algorithm optimization](#)  
[View Full Question](#)
- Q2** You are given a string s and an array of strings words. All the strings of words are of the same len...  
 Difficulty: **hard**  
 Skills Needed: [string manipulation](#) [hash table usage](#) [sliding window technique](#) [permutation generation](#) [index tracking](#)  
[View Full Question](#)

The ILP-inspired recommender returns a curated question set with difficulty labels and KC tags tailored to the user's profile.