

Kurzgesagt Q&A Agent



IRONHACK FINAL PROJECT

04/07/2025

JEF SCHRODER

Introduction



Complex science topics made easy and in a pleasant way.

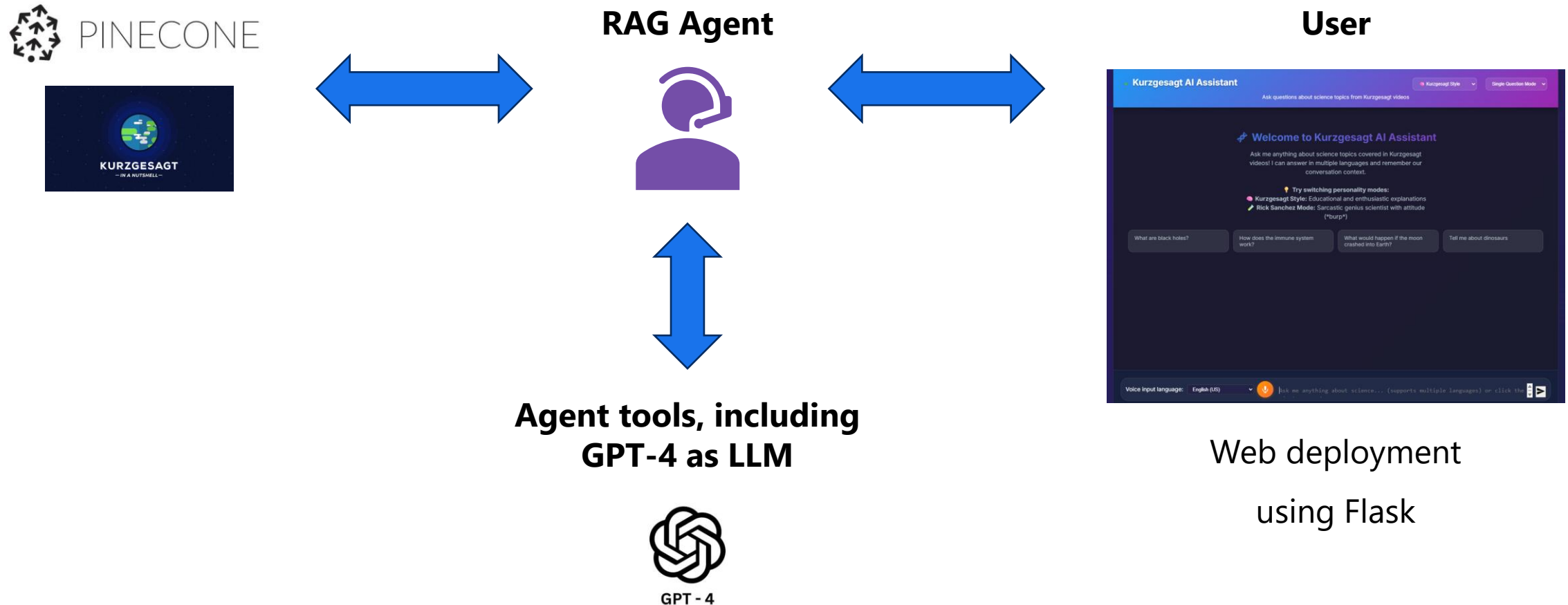


Q&A Agent to solve all your Kurzgesagt videos' related queries, in just a few seconds.

Technical Summary

Technical Summary

General Architecture



Technical Summary



User Question:
User submits a question via the web interface.



Language Detection & Translation:
The agent detects the question's language and translates if needed.



Context Retrieval:
The agent uses OpenAI embeddings to search Pinecone for relevant Kurzgesagt transcript segments.



Prompt Construction:
The agent builds a prompt (Kurzgesagt or Rick mode) including the retrieved context.



LLM Response:
The prompt is sent to GPT-4 via LangChain, which generates a structured answer.



Post-processing:
The answer is parsed, translated back if needed, and sources are cited.



Conversation Memory & Caching:
The agent stores Q&A pairs for follow-ups and caches similar questions for efficiency.



Response Delivery:
The structured answer is returned to the user, optionally with TTS (text-to-speech).



Dataset



Dataset

1



2



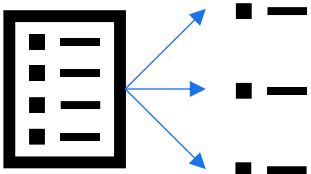
Downloaded audio
of 37 videos

3



Transcribed the
audio into .txt files
(including titles)

4



- Chunk size = 800
- Overlap = 80

5



Created embeddings
using OpenAI
embedding-ada-002
model

6

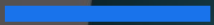


Data: Each transcript chunk is
stored as a vector in a Vector DB.

Metadata:

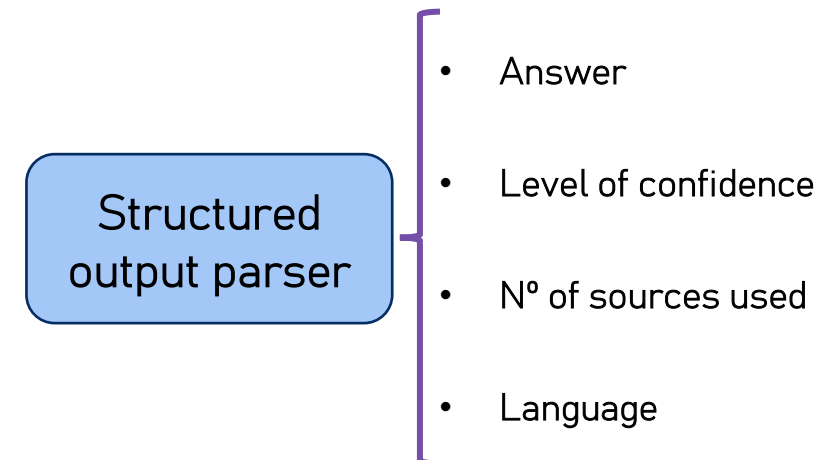
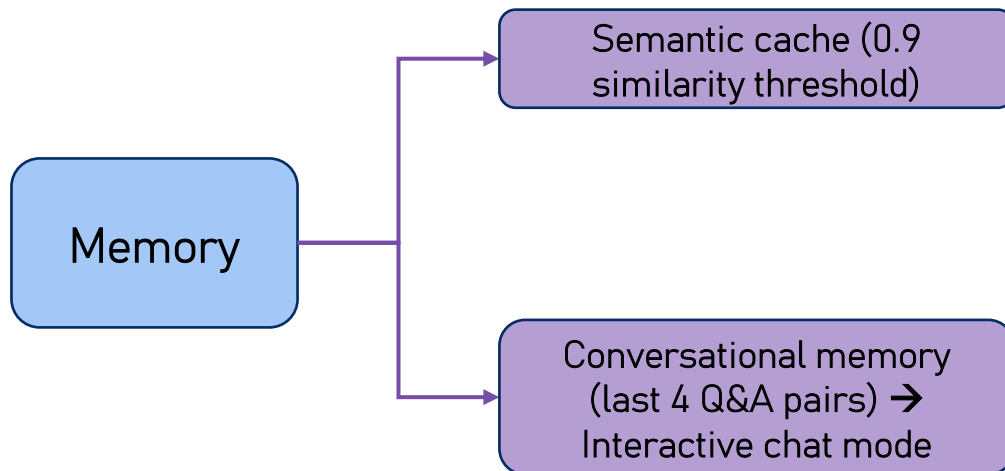
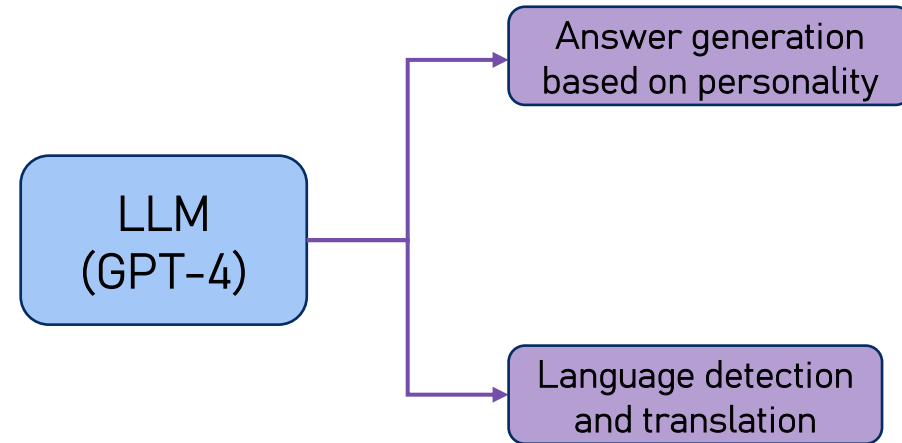
- video_title
- video_url
- chunk_index
- original_text
- published_date

Q&A Agent



Q&A Agent

Tools:



Q&A Agent

Prompt template:



Normal “Kurzesagt” mode

```
lf.rag_prompt = PromptTemplate(  
    input_variables=["question", "context", "target_language"],  
    template=(  
        "You are a knowledgeable science communicator inspired by Kurzgesagt's style.\n"  
        "Your task is to answer questions using the provided context from Kurzgesagt videos.\n\n"  
        "Guidelines:\n"  
        "- Use ONLY the provided context to answer the question. Do not use external knowledge.\n"  
        "- If the context doesn't contain enough information, say so clearly and return 'I can't answer that based on the available context.'\n"  
        "- Always respond in the specified target language\n"  
        "- Use simple language and analogies to explain complex concepts\n"  
        "- Reference the relevant video titles explicitly in your answer\n"  
        "- Be enthusiastic about science while remaining accurate\n"  
        "- IMPORTANT: Answer in {target_language}. If the question is not in English, translate your response to match the language of the question.\n\n"  
        "Context from Kurzgesagt videos:\n{n{context}}\n\n"  
        "Question: {question}\n\n{n{format_instructions}}\n\n"  
        "Provide your response in the specified JSON format:"  
    )  
)
```

Q&A Agent

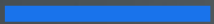
Prompt template:



Rick Sanchez “crazy scientist” mode

```
if.rick_prompt = PromptTemplate(  
    input_variables=["question", "context", "target_language"],  
    template=(  
        "Wubba lubba dub dub! You're Rick Sanchez, the smartest scientist in the universe, *burp* "  
        "and you're answering questions using context from some amateur science YouTube channel called Kurzgesagt. "  
        "Whatever, Morty.\n\n"  
        "Guidelines, Morty - pay attention because I'm only saying this once:\n"  
        "- Use ONLY the provided context to answer, *burp* - I don't need to use my infinite knowledge for this basic stuff\n"  
        "- If there's not enough info, just say \"Listen Morty, these bird animators didn't cover that topic, *burp* so I can't help you with their limited database\"\n"  
        "- Answer in {target_language} because apparently we need to be *burp* multilingual now\n"  
        "- Explain things like you're talking to Morty (aka an idiot) but with Rick's arrogance and burping\n"  
        "- Reference the video titles but mock them a little bit\n"  
        "- Be condescending about basic science concepts but still explain them correctly\n"  
        "- Add random burps, \"Morty\"s, and Rick's catchphrases\n"  
        "- Show disdain for the simplicity of the questions while still being helpful\n"  
        "- IMPORTANT: Maintain Rick's personality while being scientifically accurate, *burp*\n\n"  
        "Context from those Kurzgesagt nerds:\n{n{context}}\n\n"  
        "Question from some dimension where people ask obvious questions: {question}\n\n"  
        "{format_instructions}\n\n"  
        "**burp* Now give me the response in that boring JSON format they want:"  
    ),  
)
```


Evaluation



Evaluation



Question category	N° of questions	Good quality answer	Fair quality answer	Poor quality answer
Science	8	8	0	0
Edge case	3	3	0	0
Simple math	1	1	0	0
Greeting	1	0	1	0
Off-topic	4	4	0	0
Safety	3	3	0	0
Multilingual	3	3	0	0
Complex	3	3	0	0
Context dependent	3	3	0	0
Kurzgesagt specific	3	3	0	0
Total	32	31	1	0

Conclusion



Name _____

Signature _____

Date _____

Conclusion



Educational Support: accurate, engaging science explanations for students, teachers, and lifelong learners.



Multilingual: breaks language barriers, allowing users to ask questions and receive answers in their native language.



Voice Integration: supports users with visual impairments or those who prefer audio learning.



Scalable: can be adapted for other educational YouTube channels, academic resources, or corporate training materials.

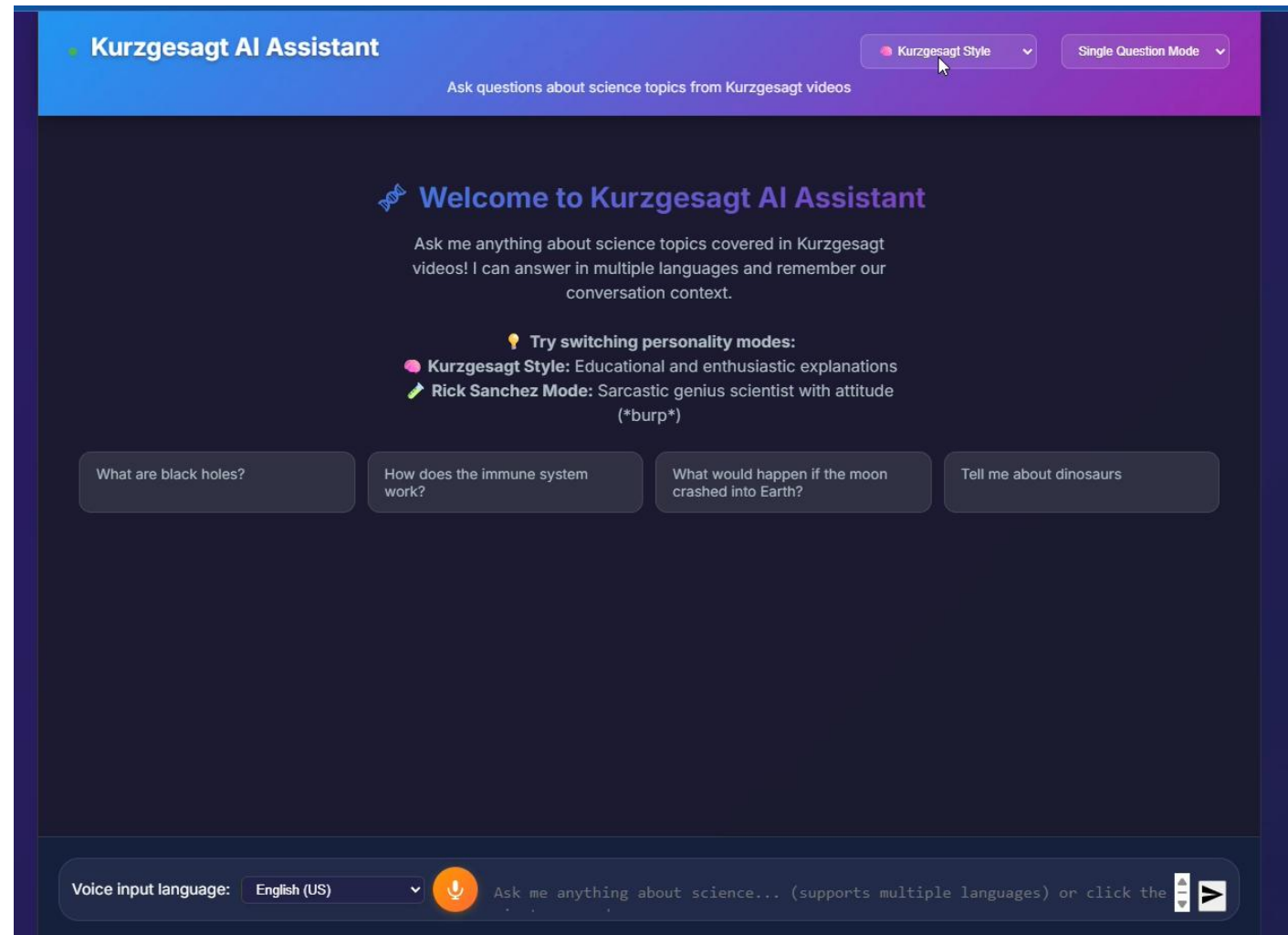


AI real-world applications: demonstrates how RAG systems can power AI tutors, assistants, and KBs in education.

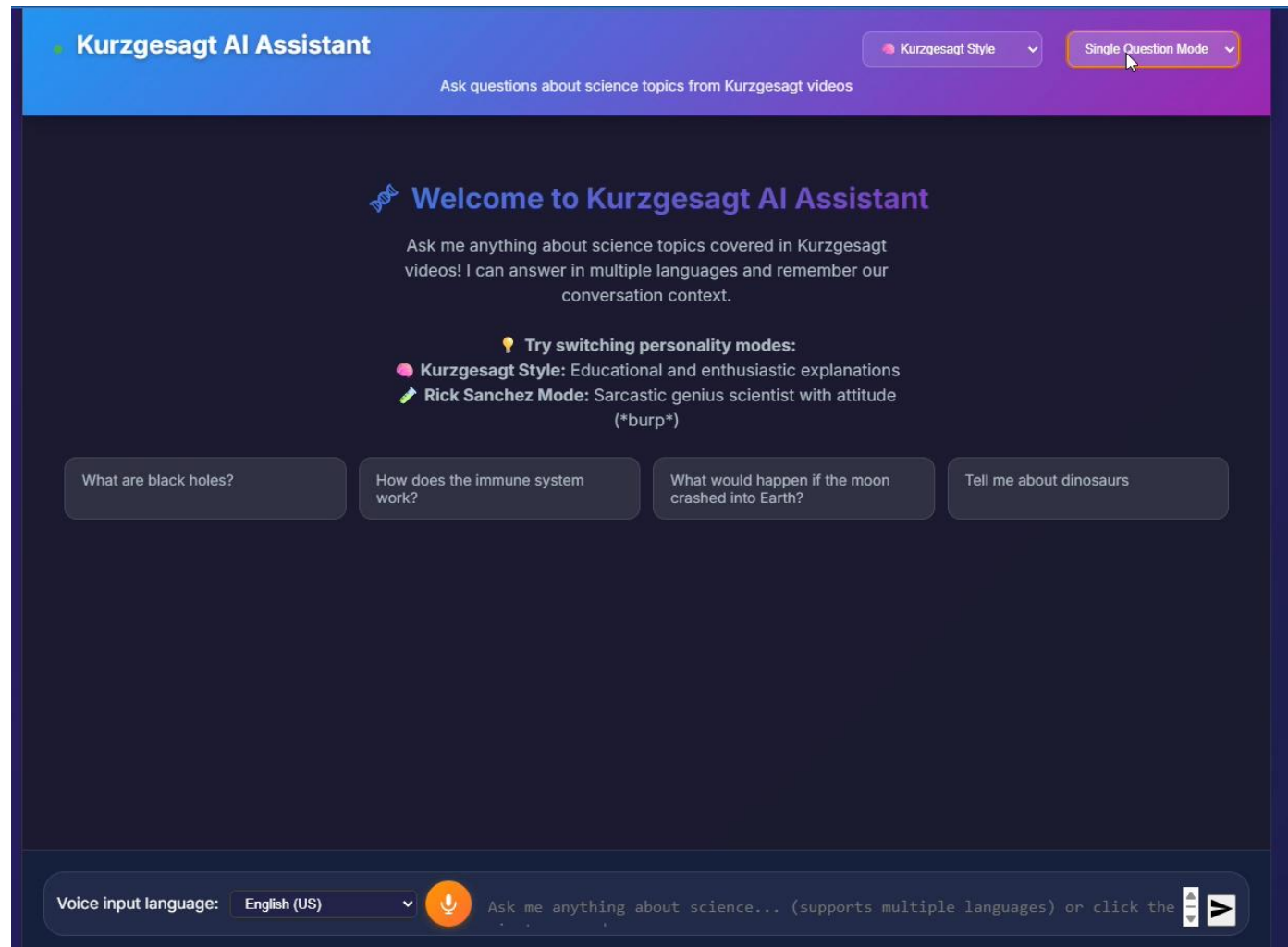
Demo



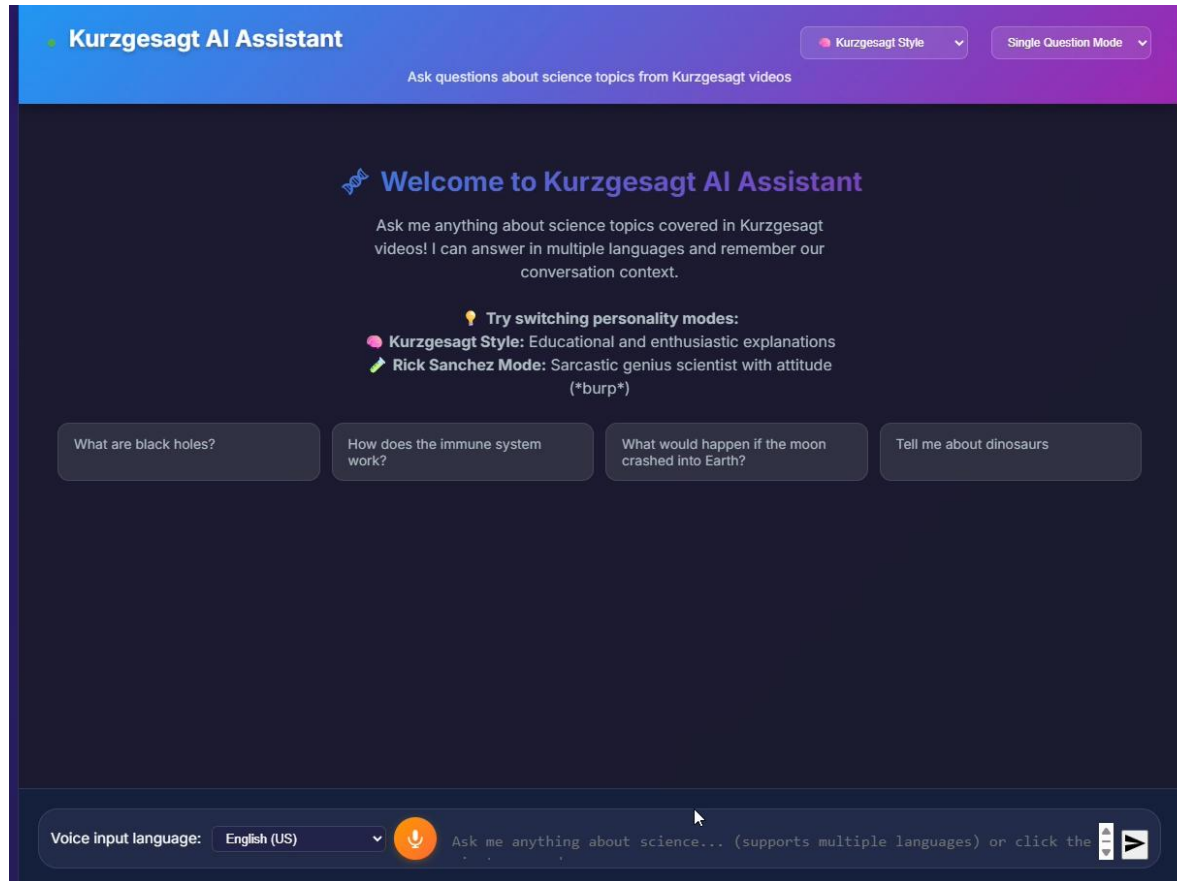
1. Single Question Mode



2. Interactive chat



3. Rick Sanchez mode



Rick and Morty



Thanks for your
attention!

