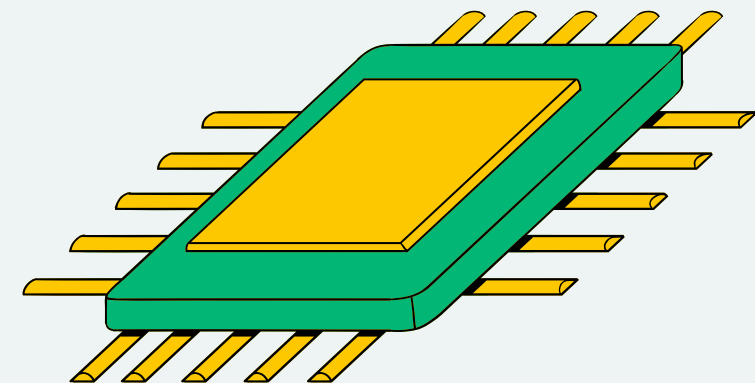


YOUTUBE VIDEO ASSISTANT: AI-POWERED TRANSCRIPTION & QUERYING

PRESENTED BY:

DANA ALOMARI





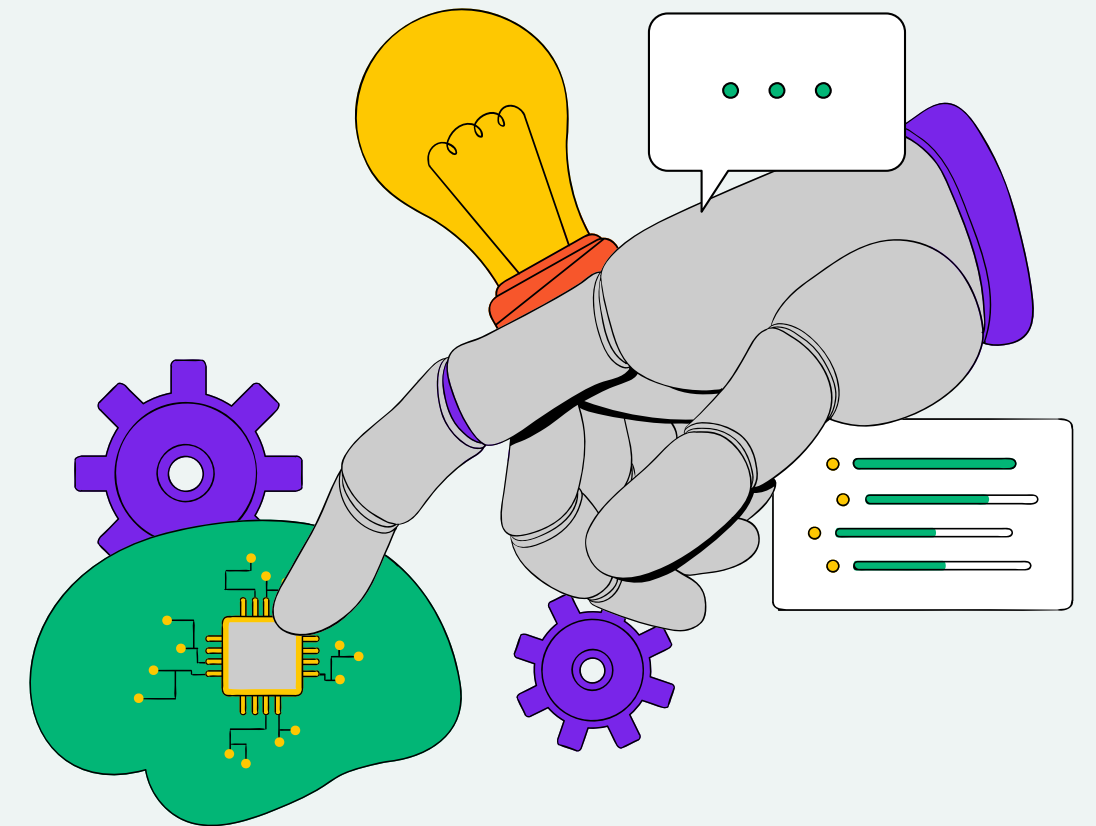
PRESENTATION OUTLINE

- Project Objective
- Problem Statement
- Architecture Diagram
- Use Cases:
- Architecture Diagram
- Performance Evaluation
- RAG Pipeline Architecture
- Handling Common Issues
- Latency, Cost, and Memory Management
- Future Proofing
- DEMO

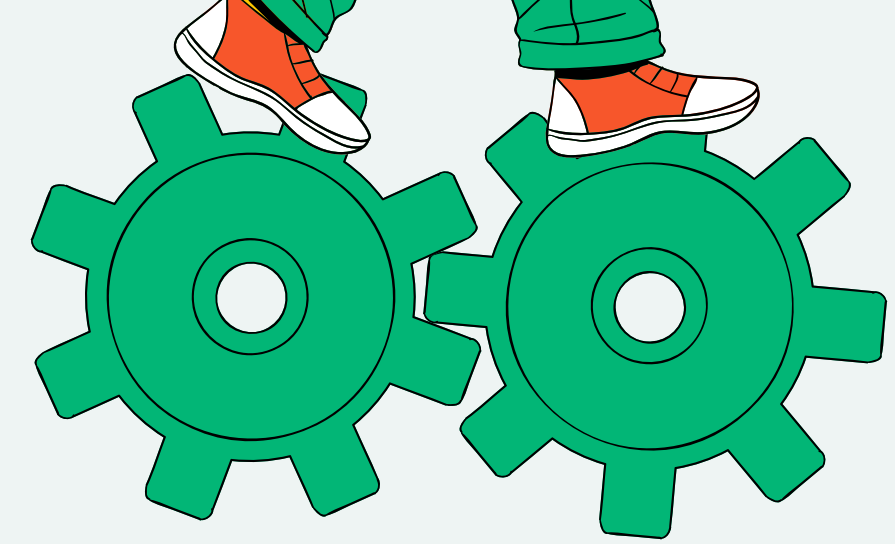


PROJECT OBJECTIVE

**Build a chatbot that can
translate YouTube videos
into text and speech that
enables natural language
queries.**



PROBLEM STATEMENT:



- **80% of users prefer text summaries over videos.**
- **Hearing-impaired and non-native speakers struggle with video content.**

ARCHITECTURE DIAGRAM

TEXT & AUDIO PROCESSING:



TRANSCRIPTION (WHISPER)

Converts audio to text.



VECTOR STORE (CHROMA)

Stores and indexes video content.



CHATBOT INTERFACE (CHATGPT + LANGCHAIN)

Provides conversational interface for users to ask questions about the video content

DESIGN DECISIONS:



WHISPER FOR AUDIO TRANSCRIPTUON

Chosen for accuracy and ease of use



CHROMA FOR VECTOR STORAGE

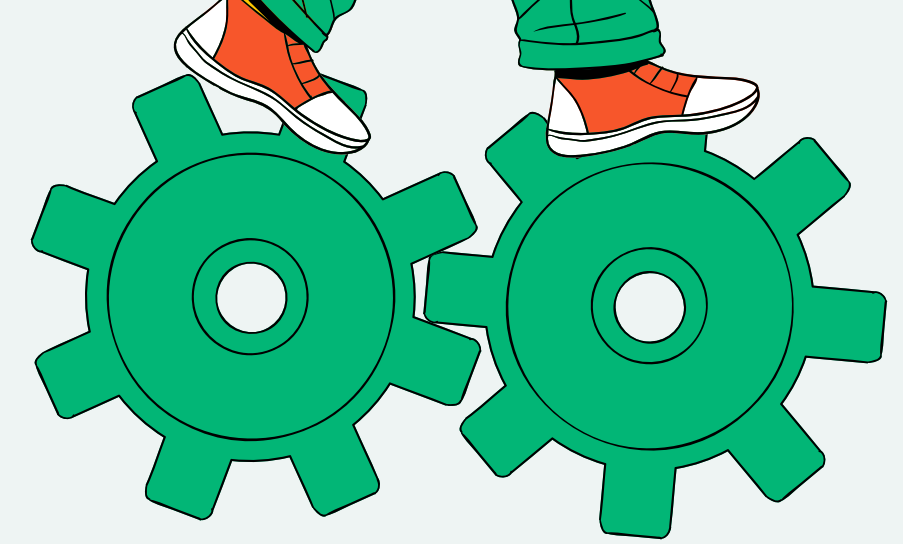
Selected for efficient similarity search



CHATGPT FOR Q&A

Based on its ability to generate coherent

USE CASES:



- **Accessibility for hearing-impaired**
- **Educational content indexing**
- **Multilingual support**

PERFORMANCE EVALUATION

Baseline:

Test bot's transcription and query accuracy.

Improvements:

Show how RAG boosts performance.



HANDLING COMMON ISSUES

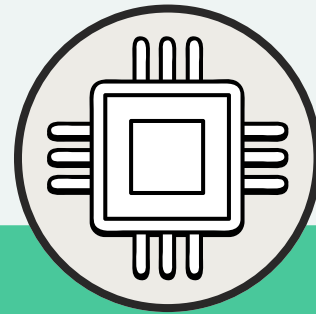
- Hallucinations

- Relevancy



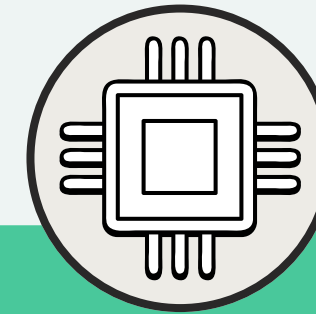
HALLUCINATIONS

CHALLENGE:



- **Model sometimes generates unrelated content.**

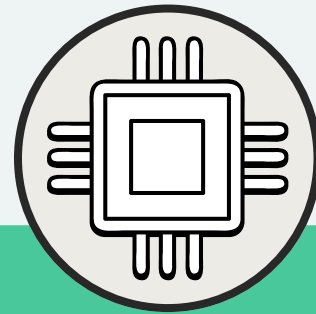
SOLUTION:



- **Used Chroma vector search for video-based answers.**

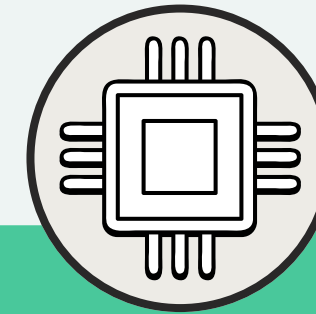
RELEVANCY

CHALLENGE:



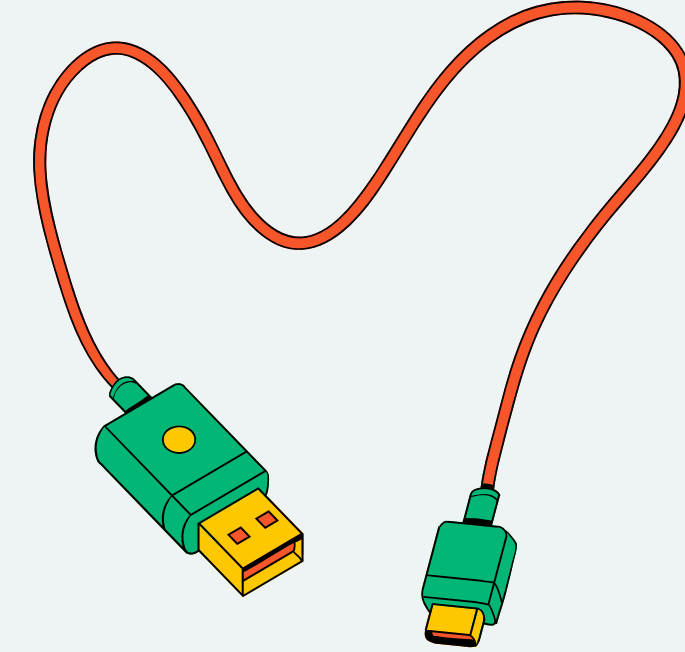
- Some user queries may not align with the video content.

SOLUTION:



- Refined search to use only relevant transcript segments for answers.

LATENCY, COST, AND MEMORY MANAGEMENT



LATENCY

- VIDEO PROCESSING
- TRANSCRIPTION (WHISPER)
- VECTOR SEARCH (CHROMA)
- CHATBOT (GPT)

COST CONSIDERATIONS

- WHISPER & GPT COSTS
- EFFICIENT LIBRARIES:

MEMORY REQUIREMENTS

- CHROMA VECTOR STORAGE
- MANAGING MEMORY

FUTURE PROOFING

Scalability:

- Improve vector store and parallel processing

Enhancements:

- Integrate images/other media.
- Enable real-time transcription.





DEMO



THANK YOU!

QUESTIONS?

