



Earnings Q&A

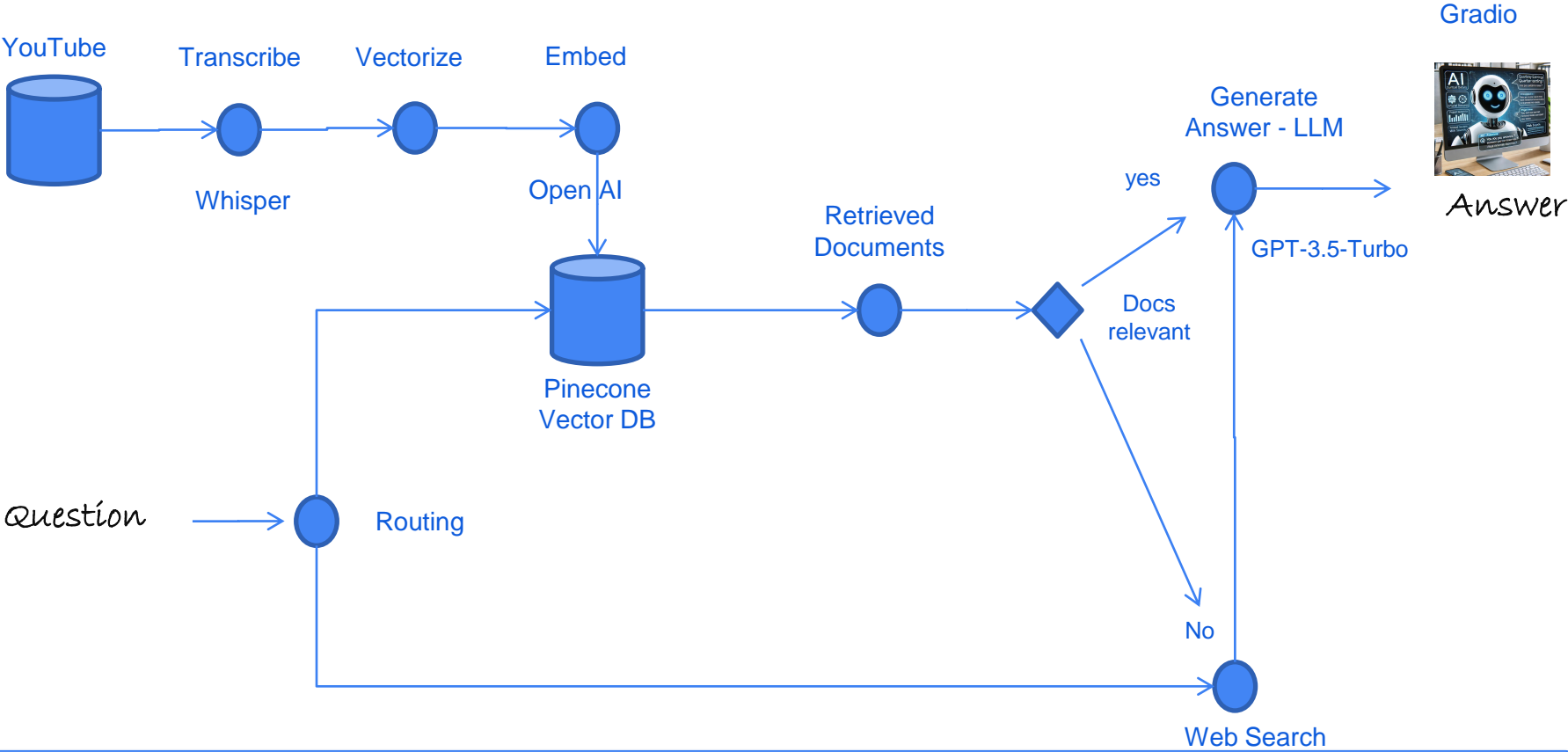
*Karthikeyan
Karuppusamy (KK)*

Overview & Scope

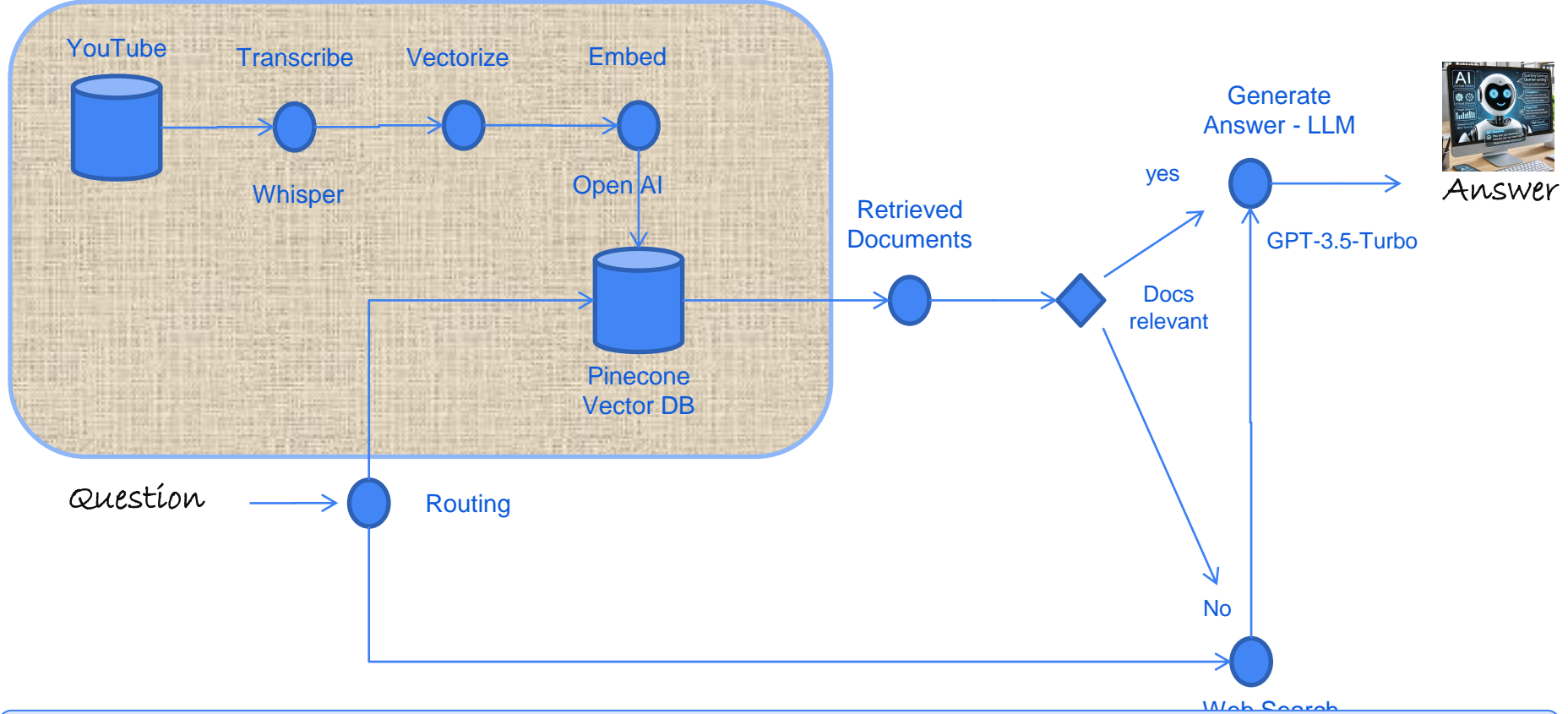
Build a conversational chatbot for Earnings Q&A

- with standout features (multi-source retrieval, dynamic answer generation, etc.)
- supplement Earning QA with Generic QA
- Implement Langchain ReAct agent & Tools
- Deploy on Gradio

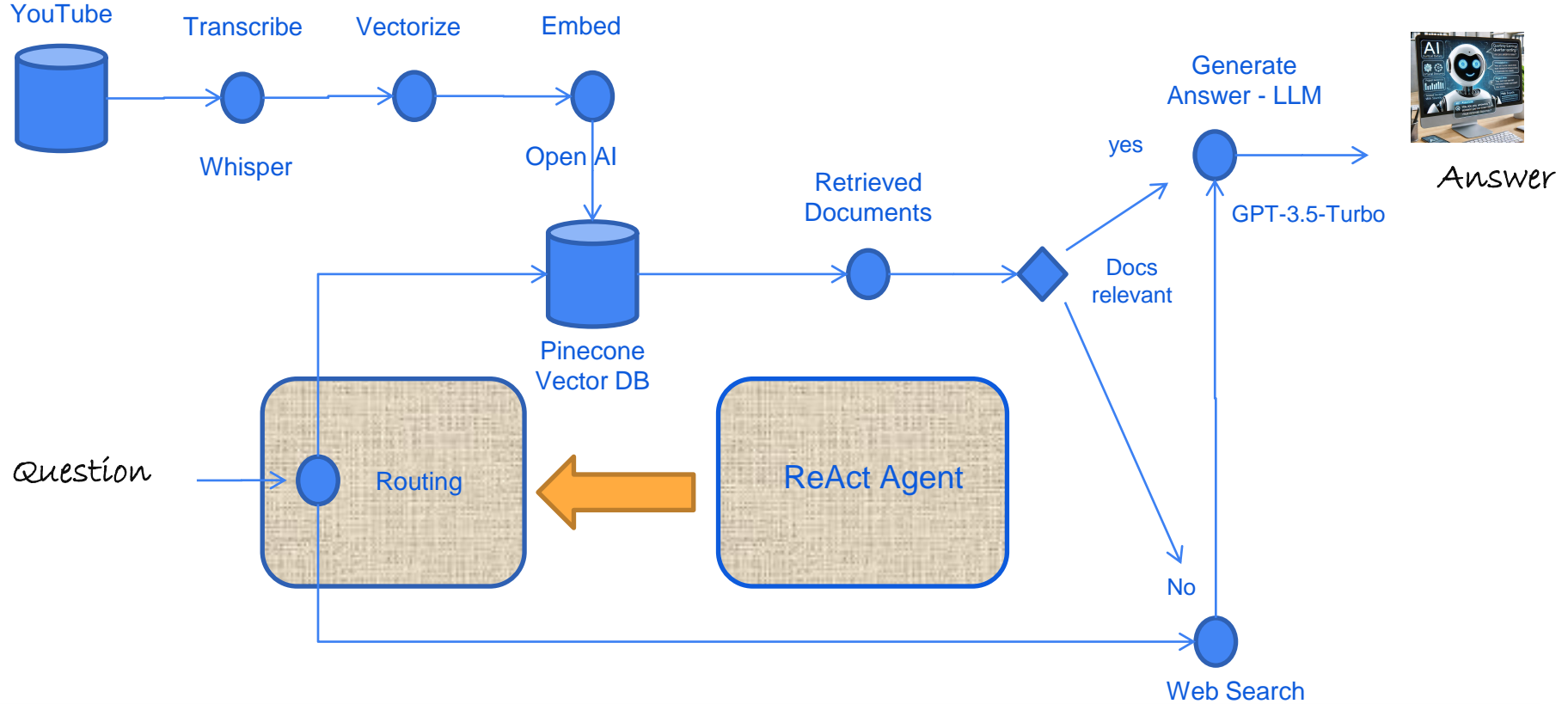
Solution Architecture



Dataset



ReAct Agent



YouTube



Transcribe

Wh

```
# Initialize the router instructions
router_instructions = (
    "You are an expert at routing a user question to a vectorstore or web search.\n\n"
    "The vectorstore contains documents related to Accenture company Quarterly results, "
    "Revenue, Earnings for the last four quarters, sales, digital projects, cloud projects "
    "and employee details.\n\n"
    "Use the vectorstore for questions on these topics. For all else, and especially for "
    "current events, use web-search.\n\n"
    "Return JSON with a single key, 'datasource', that is either 'websearch' or 'vectorstore' "
    "depending on the question."
)
```

```
# Function to route query based on instructions
def route_query(query):
    """route query to capture vectordb or websearch"""
    routing_response = chat_model.invoke(
        [SystemMessage(content=router_instructions), HumanMessage(content=query)]
    )
    route_decision = json.loads(routing_response.content)
    return route_decision["datasource"]
```

Generate
LLM



Answer

GPT-3.5-Turbo

Question

Routing

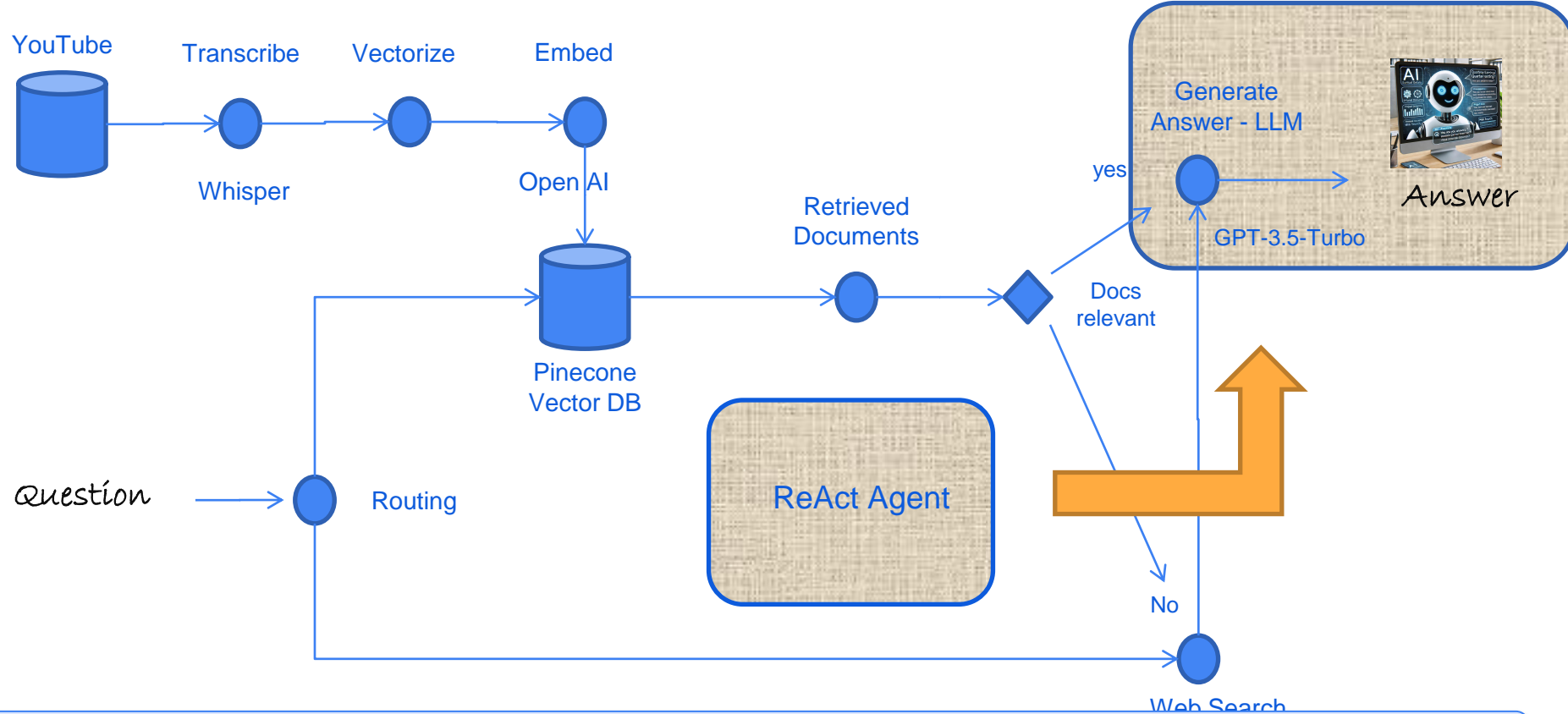
Vector
DB

ReAct Agent

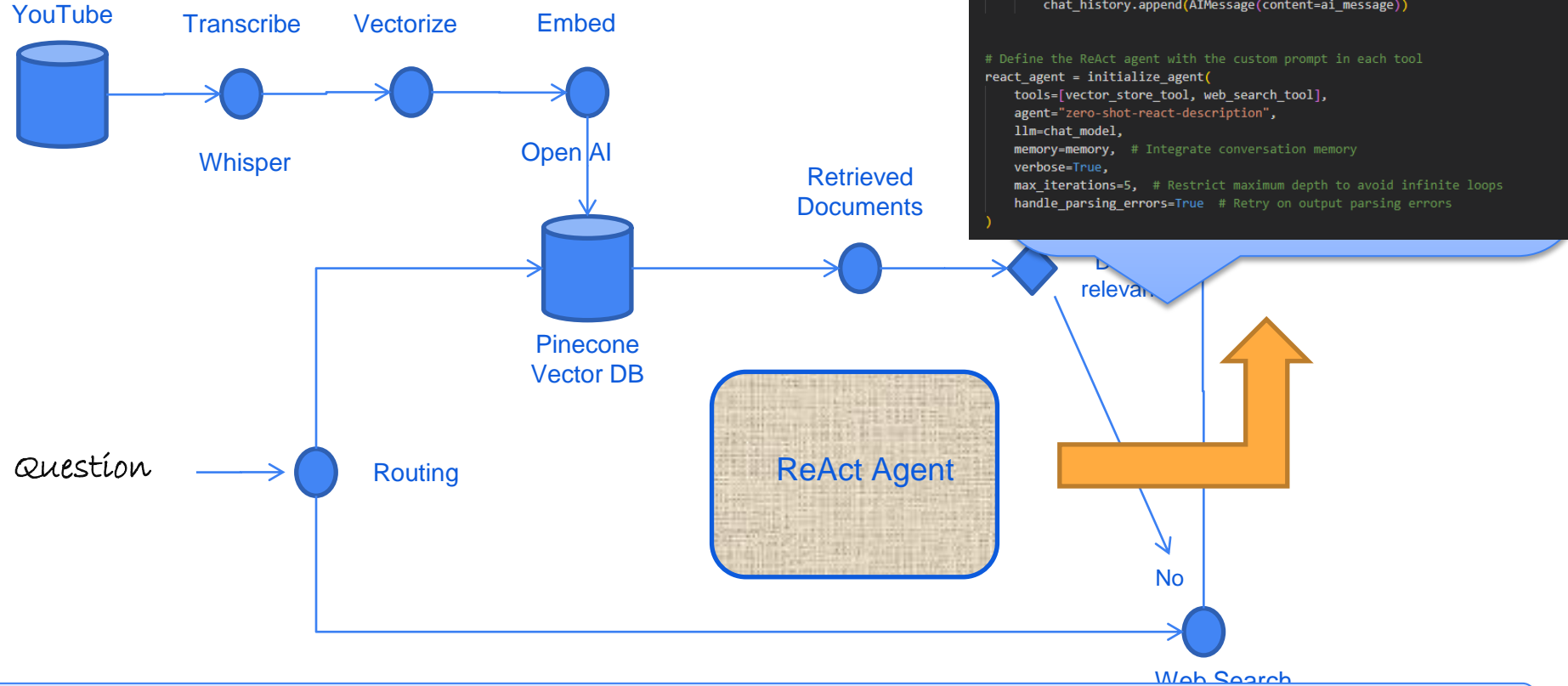
No

Web Search

Generation & Chatbot



Generation & Chatbot



Evaluation

- Input Dataset Optimization
- Router node implementation
- Langsmith for Tracing & debugging
- Giskard testing & Evaluation – results below.

Consolidated Input

```
# Summary with the results.  
report.correctness_by_question_type()
```

✓ 0.0s

correctness	
question_type	
complex	0.000000
conversational	0.500000
distracting element	0.000000
double	0.000000
simple	0.333333
situational	0.500000



Split Input

```
# Summary with the results.  
report.correctness_by_question_type()
```

✓ 0.0s

correctness	
question_type	
complex	0.5
conversational	0.0
distracting element	0.5
double	0.0
simple	1.0
situational	0.0

Giskard evaluation: contd.,

Question 1: Could you provide me with the specific monetary amount that Accenture earned during the last quarter of the year 2024?

Agent Answer 1: Accenture earned \$65 billion in the last quarter of the year 2024.

Question 2: Considering the unavailability of the EPS for the fourth quarter of 2024, could you provide the current total employee count in Accenture?

Agent Answer 2: Approximately 733,000 employees work at Accenture.

Question 3: As an investor, I'm interested in the size of GenAI project deals for Accenture in the year 2024. Can you provide this information?

Agent Answer 3: Over \$300 million

Question 4: As a financial analyst, I'm trying to find out the EPS of Accenture for the fourth quarter of 2024. Can you provide this information?

Agent Answer 4: The EPS of Accenture for the fourth quarter of 2024 was \$2.79.

Question 5: What were the earnings and EPS of Accenture in the fourth quarter of 2024?

Agent Answer 5: In the fourth quarter of 2024, Accenture reported adjusted earnings per share of \$11.67.

Question 6: What was their size?

Agent Answer 6: Accenture has approximately 733,000 employees.

Failure 1: The agent provided a specific amount for Accenture's earnings in the last quarter of 2024, but the reference answer indicates that this information was not provided in the doc

Failure 2: The agent stated that approximately 733,000 employees work at Accenture, but the correct number is 774,000.

Failure 3: The agent stated that the size of GenAI project deals for Accenture in the year 2024 was over \$300 million, but should have said it was \$2 billion.

Failure 4: The agent provided an EPS value for Accenture in the fourth quarter of 2024, but the reference answer indicates that this information is not available in the provided document

Failure 5: The agent provided specific earnings per share for Accenture in the fourth quarter of 2024, but the reference answer indicates that this information was not provided in the doc

Failure 6: The agent provided the number of employees at Accenture, but the user was asking about the size of the GenAI project deals for Accenture in 2024, which was \$2 billion.

Project Approach

Sprint	Goal	Tasks
Sprint 1	Initial Setup and Processing of a Single YouTube Video	<ul style="list-style-type: none">- Loaded one short (5min/20min) YouTube video to understand processing flow.- Extracted audio, transcribed with Whisper.- Generated embeddings and stored them in Pinecone as a trial run
Sprint 2	Scaling to Multiple YouTube Videos	<ul style="list-style-type: none">- Processed multiple YouTube videos relevant to Accenture's earnings.- Implemented an automated loop for audio extraction and transcription.- Implemented ConversationalQA agent
Sprint 3	Basic Chat functionality	<ul style="list-style-type: none">- Implemented ReAct agent.- Implemented Routing functionality.- Implemented Chat Function
Sprint 1	Chatbot and Gradio integration	<ul style="list-style-type: none">- Switched to ConversationBufferWindowMemory to retain recent context more efficiently.- Integrated Gradio for a web-based user interface.
Sprint 1	Evaluation & Tracking with Langsmith and Giskard	<ul style="list-style-type: none">- Added LangSmith to track interactions for debugging and analysis.- Conducted initial evaluations with Giskard.- Fine-tuned the input dataset based on Giskard evaluation.

Deployment:

Gradio link : <https://d50010d4b9a96855ac.gradio.live>

Next Steps:

- Enable Audio functionality for Chatbot
- Deploy it in HuggingFace or AWS. Make it more UI friendly.
- Use of Opensource LLM models for cost efficiency and data security.

Thank You!!!