



# LLM Supported Report Generation from Expert's Recording - InsightBridge

**CSE 496**  
**Second Presentation**

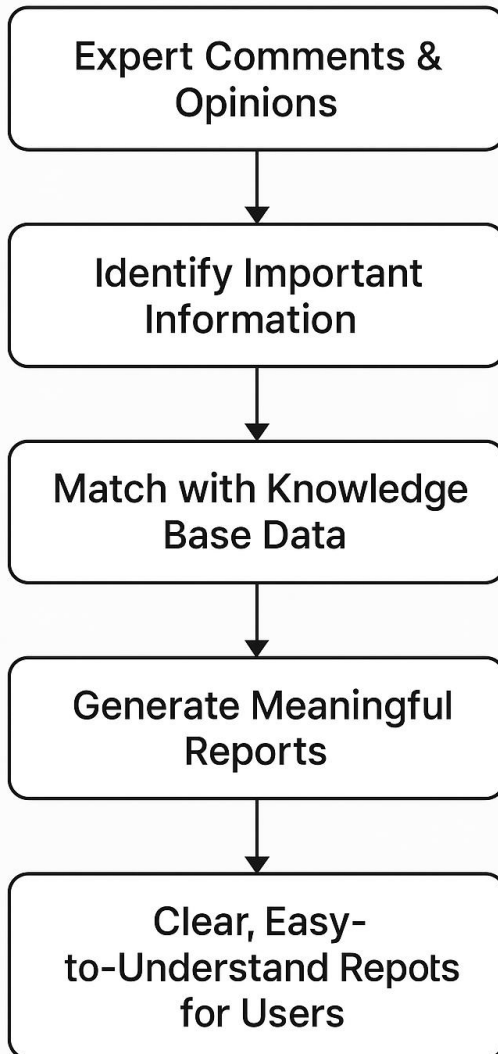
**Tugay Talha İçen**

**Project Advisor: Doc. Dr. HABİL KALKAN**  
**May 2025**



- Project Schema and Introduction
- Project Timeline
- Project Design Plan
- Project Requirements
- Success Criteria
- Resources





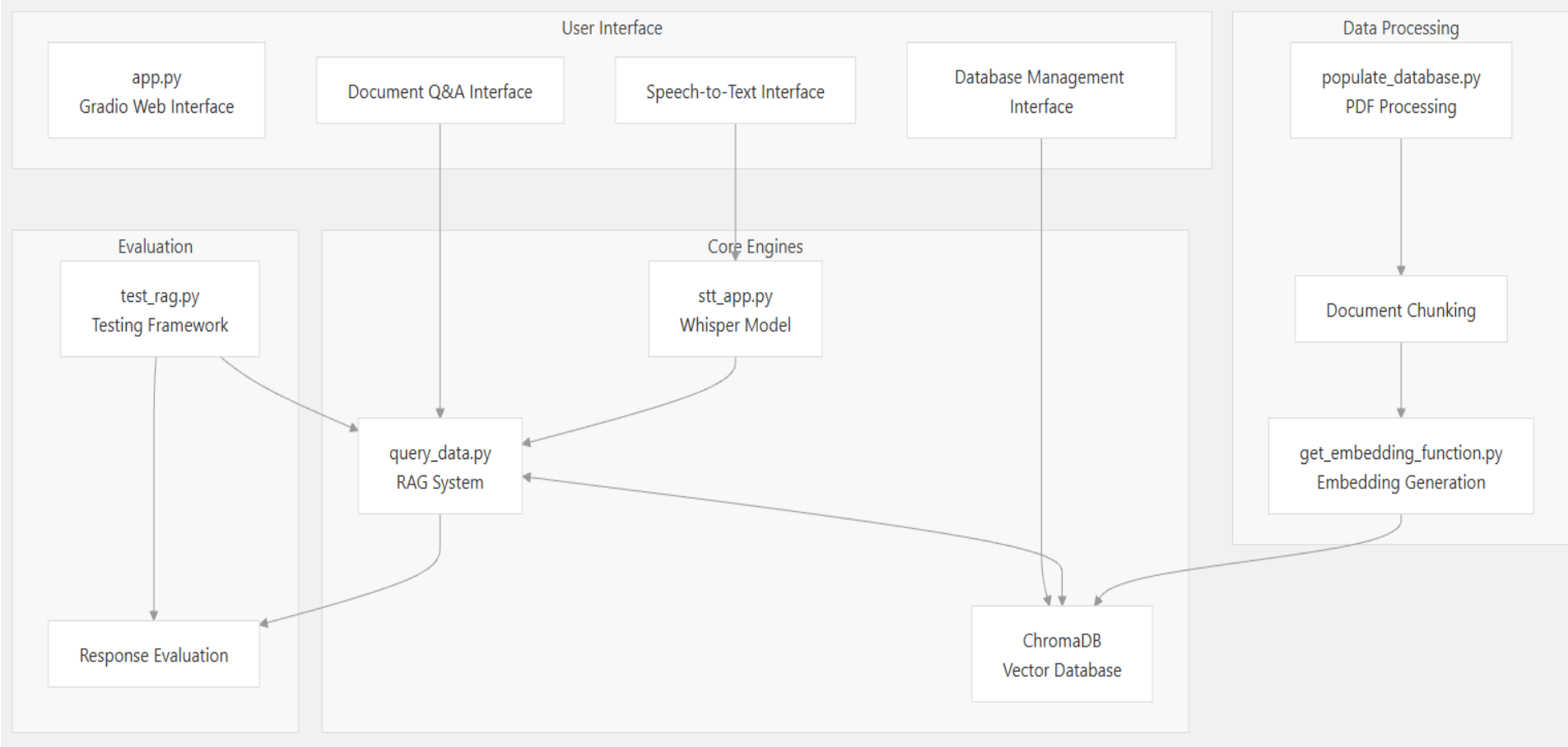
## **Description**

- Aim of this project is automatically create clear, structured reports by intelligently extracting valuable insights from expert comments and combining these with existing data in a structured knowledge base to generate a report.

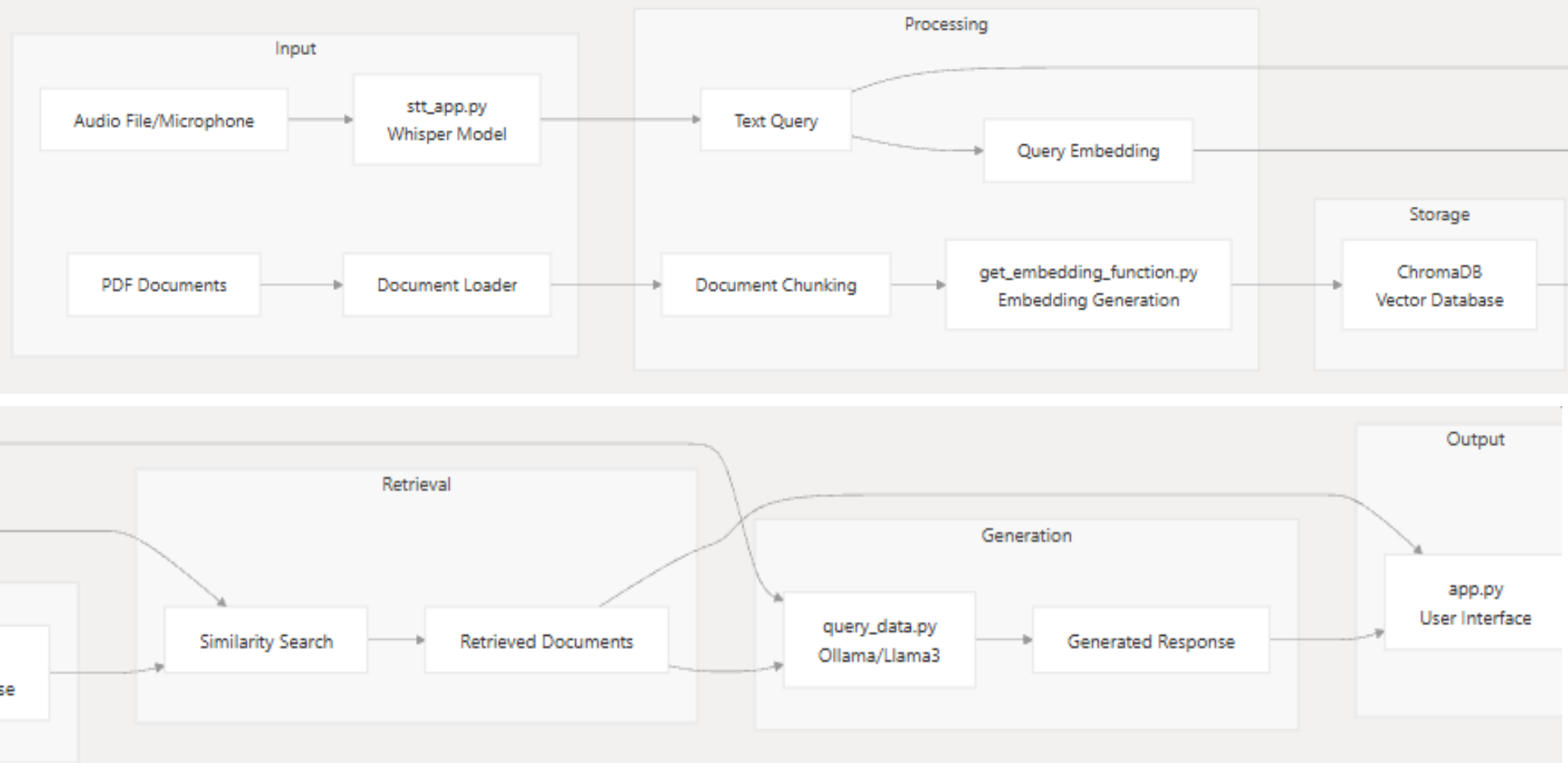
## **Purpose of the Project:**

- Transform expert comments into actionable insights.
- Automatically combine these insights with existing structured data.
- Generate simple, clear, and informative reports understandable by everyone.

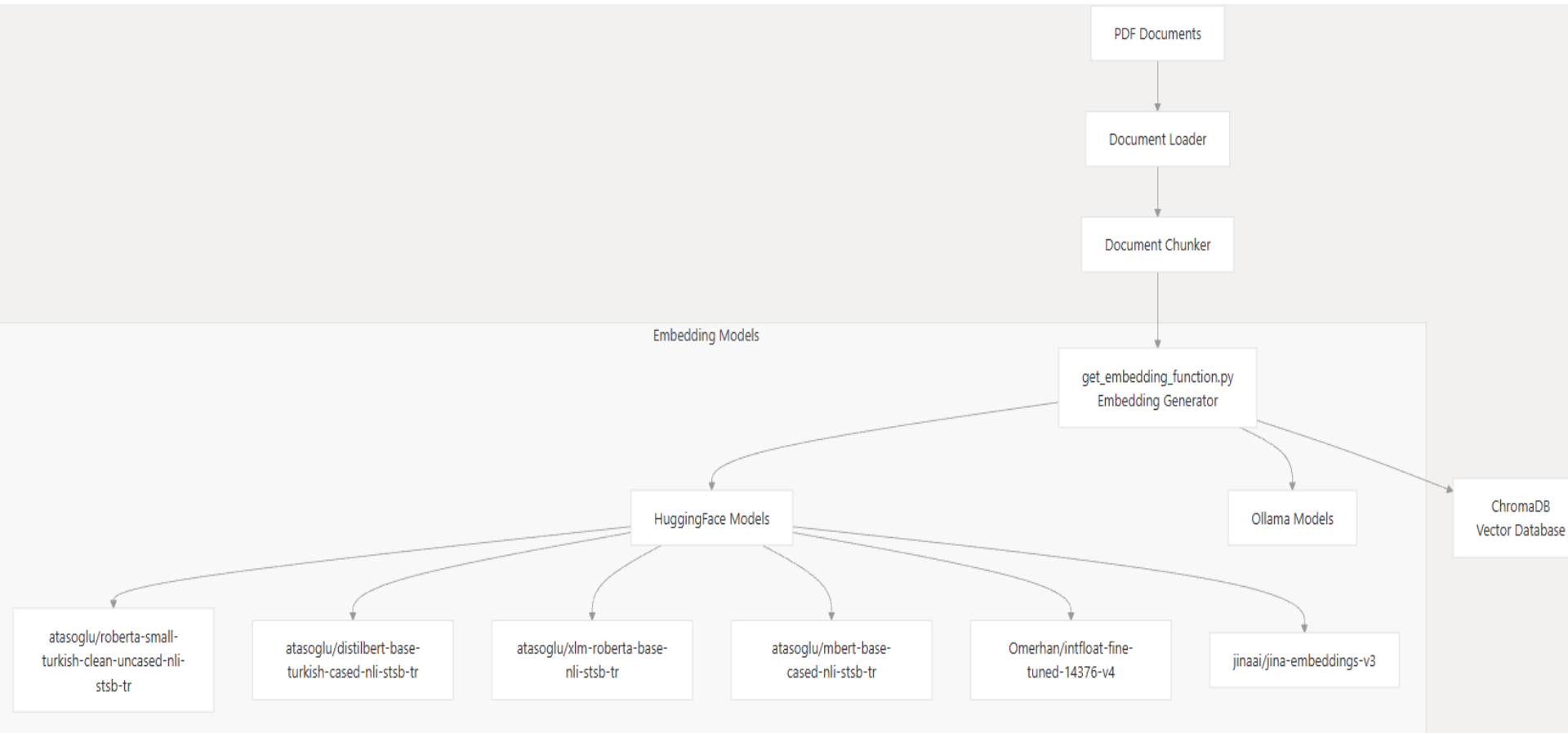
# System Architecture Overview



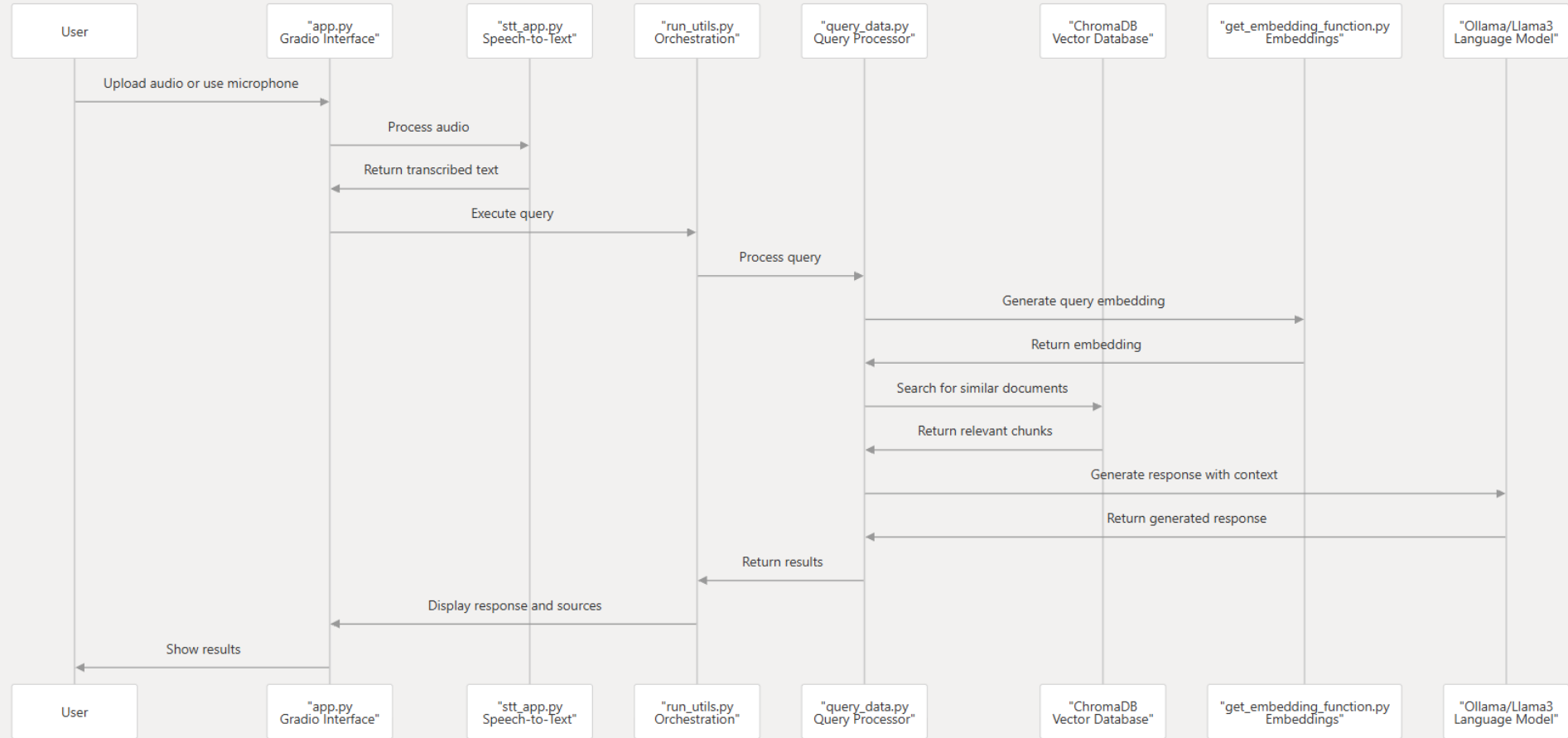
# Data Flow Architecture



# Document Management



# Technical Architecture



# Interface

Query Documents Database Management

Upload Audio File OR Record from Microphone



Drop Audio Here

- or -

Click to Upload



Language Code (Optional)

e.g., 'en', 'es', 'fr'. Leave blank to auto-detect.

Submit Query

Response

Retrieved Chunks

Source	Content	Relevance Score
--------	---------	-----------------

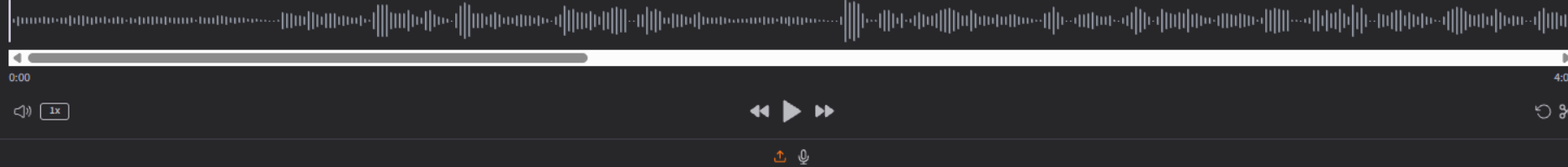




# Example Output - 1

Query Documents Database Management

Upload Audio File OR Record from Microphone



Language Code (Optional)

en

Submit Query

Response

**\*\*Findings:\*\***

**\*\*Coronal Views:\*\***

A radial tear is identified in the **\*\*anterolateral aspect of the lateral meniscus\*\***. The meniscus demonstrates **\*\*increased signal intensity\*\*** in the zone of the tear, which is best visualized on **\*\*fat-saturated images\*\***. The **\*\*iliotibial band\*\*** is noted, and there is evidence of **\*\*tearing\*\*** in the anterolateral portion of the lateral meniscus. There is no evidence of **\*\*meniscal extrusion\*\***, and the **\*\*root attachment\*\*** of the meniscus is intact.

**\*\*Sagittal Views:\*\***

The sagittal images show **\*\*disruption\*\*** of the lateral meniscus, with **\*\*increased signal intensity\*\*** and **\*\*loss of the normal "bow tie" appearance\*\***. The **\*\*anterior horn\*\*** of the lateral meniscus shows **\*\*increased signal\*\***, suggesting a **\*\*complete disruption\*\*** of the meniscus. The **\*\*patellar tendon\*\*** shows **\*\*signs of patellar tendonitis\*\*** at its insertion on the patella. The **\*\*anterior cruciate ligament (ACL)\*\*** and **\*\*posterior cruciate ligament (PCL)\*\*** are **\*\*intact\*\***. The **\*\*medial meniscus\*\*** appears **\*\*normal\*\***, with no increased signal intensity.

**\*\*Axial Views:\*\***

The axial images demonstrate a **\*\*complete disruption\*\*** of the **\*\*anterior aspect of the lateral meniscus\*\***, with **\*\*fluid within the meniscal substance\*\***, indicating a **\*\*radial tear\*\***. This disruption leads to **\*\*loss of hoop fibers\*\***, resulting in **\*\*instability of the meniscus\*\*** and impaired function as a **\*\*shock absorber\*\***. There is **\*\*joint swelling\*\*** in the **\*\*proximal region\*\***. A **\*\*small lesion\*\*** is noted in the **\*\*trochlear cartilage\*\***, with **\*\*subchondral bone swelling\*\***. The **\*\*articular cartilage\*\*** of the **\*\*patella\*\*** is **\*\*normal\*\***.

---

**\*\*Impression:\*\***

A **\*\*radial tear of the anterior aspect of the lateral meniscus\*\*** is identified, which is a **\*\*complete disruption\*\*** leading to **\*\*meniscal instability\*\***. This type of tear is **\*\*common in younger patients\*\*** and can be **\*\*catastrophic\*\*** due to the **\*\*loss of structural integrity\*\*** of the meniscus. No significant abnormalities are noted in the **\*\*medial meniscus\*\***, **\*\*ACL\*\***, or **\*\*PCL\*\***. There is also evidence of **\*\*patellar tendonitis\*\*** and a **\*\*small trochlear cartilage lesion\*\***.

Retrieved Chunks

Source	Content	Relevance Score
data\10_Radiology-Fundamentals-Introduction-to-ImagingTechnology-Fourth-Edition-2012.pdf:41:1	tissue detail to be displayed. Each individual picture of a CT study is referred to as a section or an axial "slice." This is because the picture must be interpreted as if the patient has been completely sectioned in an axial plane, like a loaf of bread, with the viewer looking at the section from the feet toward the head (Fig. 5.1 ). Pixels and Voxels If you look closely at the CT scan, you will realize that the picture is actually made up of thousands of little squares called "pixels" (picture elements). Each pixel represents tissue that is about 1 mm or less on each of the two sides, and is assigned 5 Computed Tomography Objectives: 1. State how the densities on a CT scan are assigned	0.1695

# Example Output - Transcribe

start to see some increase in signal intensity in the zone of the meniscus here where the tear is located. So this is the iliotibial band. And we can start to see evidence of some tearing here of the anterolateral aspect of the lateral meniscus. And that's kind of a subtle finding, and that's what we usually see with these is we don't see it as well in the coronal views. We want to follow the meniscus back and see if there's any extrusion of the meniscus. And then follow it all the way posteriorly to verify that the root attachment right here is still intact. So we can see in general the only thing that really is a signal that there's a radial tear is this change on the anterolateral aspect. The next view that I'll go through is a sagittal view. And this is starting at the far lateral aspect. So as we follow this along, we start to see some appearance here of the lateral meniscus. And there's evidence of some disruption here. So the normal appearance would be a dark signal, almost looking like a bow tie. And as we come along, you'll see that there's some disruption here and some increased signal intensity. And then we'll start to see more normal appearance. So the anterior horn here has some increased signal within the substance, but it's difficult to tell if that's a tear just looking at that one cut. And as we go more towards the midline, we see that there's evidence of patellar tendonitis right here at the insertion of the patellar tendon on the patella. Here's the ACL. The ACL looks normal. Here's the PCL, normal appearance. And then we follow all the way over medially, and we can see the medial meniscus looks normal. And here's that bow tie appearance here where there really isn't any increased signal intensity. Within the medial meniscus. So the main cut we'd look for to look for radial tears is to look at the axial cuts. So I'm starting down distal here. Here's the fibular head. This is the patellar tendon. And as I start to course proximate, I want to stop right at the joint line and see this. So we'll course up image by image. We're starting to see a little bit of the meniscus here ghosting in. And this is a cut right through the center of the joint. So this is the anterior **inner meniscus** ligament. This is the anterior horn. And we can see that there's fluid here, which indicates that there is a complete disruption, which is a radial tear of the anterior aspect of the lateral meniscus. And because it's completely disrupted, it disrupts all the hoop fibers. And it makes the whole meniscus unstable so it doesn't act as a normal shock absorber. As we see more proximally, we can see some swelling within the joint along here. This is a trochlea cartilage with a little bit of a trochlea lesion right here with some swelling in

%99+ Accuracy

~700 Word    1 Mistake



# Example Output - 3

## Radiology Report

**Patient:** Unspecified  
**Study:** MRI of the Right Knee  
**Examined By:** [Doctor's Name]  
**Date:** [Insert Date]  
**Reported By:** [Radiologist's Name]

### Clinical Indication:

Assessment of a radial tear of the meniscus on MRI.

### Findings:

#### Coronal Views:

A radial tear is identified in the **anterolateral aspect of the lateral meniscus**. The meniscus demonstrates **increased signal intensity** in the zone of the tear, which is best visualized on **fat-saturated images**. The **iliotibial band** is noted, and there is evidence of **tearing** in the anterolateral portion of the lateral meniscus. There is no evidence of **meniscal extrusion**, and the **root attachment** of the meniscus is intact.

#### Sagittal Views:

The sagittal images show **disruption** of the lateral meniscus, with **increased signal intensity** and **loss of the normal "bow tie" appearance**. The **anterior horn** of the lateral meniscus shows **increased signal**, suggesting a **complete disruption** of the meniscus. The **patellar tendon** shows **signs of patellar tendonitis** at its insertion on the patella. The **anterior cruciate ligament (ACL)** and **posterior cruciate ligament (PCL)** are **intact**. The **medial meniscus** appears **normal**, with no increased signal intensity.

#### Axial Views:

The axial images demonstrate a **complete disruption** of the **anterior aspect of the lateral meniscus**, with **fluid within the meniscal substance**, indicating a **radial tear**. This disruption leads to **loss of hoop fibers**, resulting in **instability of the meniscus** and impaired function as a **shock absorber**. There is **joint swelling** in the **proximal region**. A **small lesion** is noted in the **trochlear cartilage**, with **subchondral bone swelling**. The **articular cartilage** of the **patella** is **normal**.

### Impression:

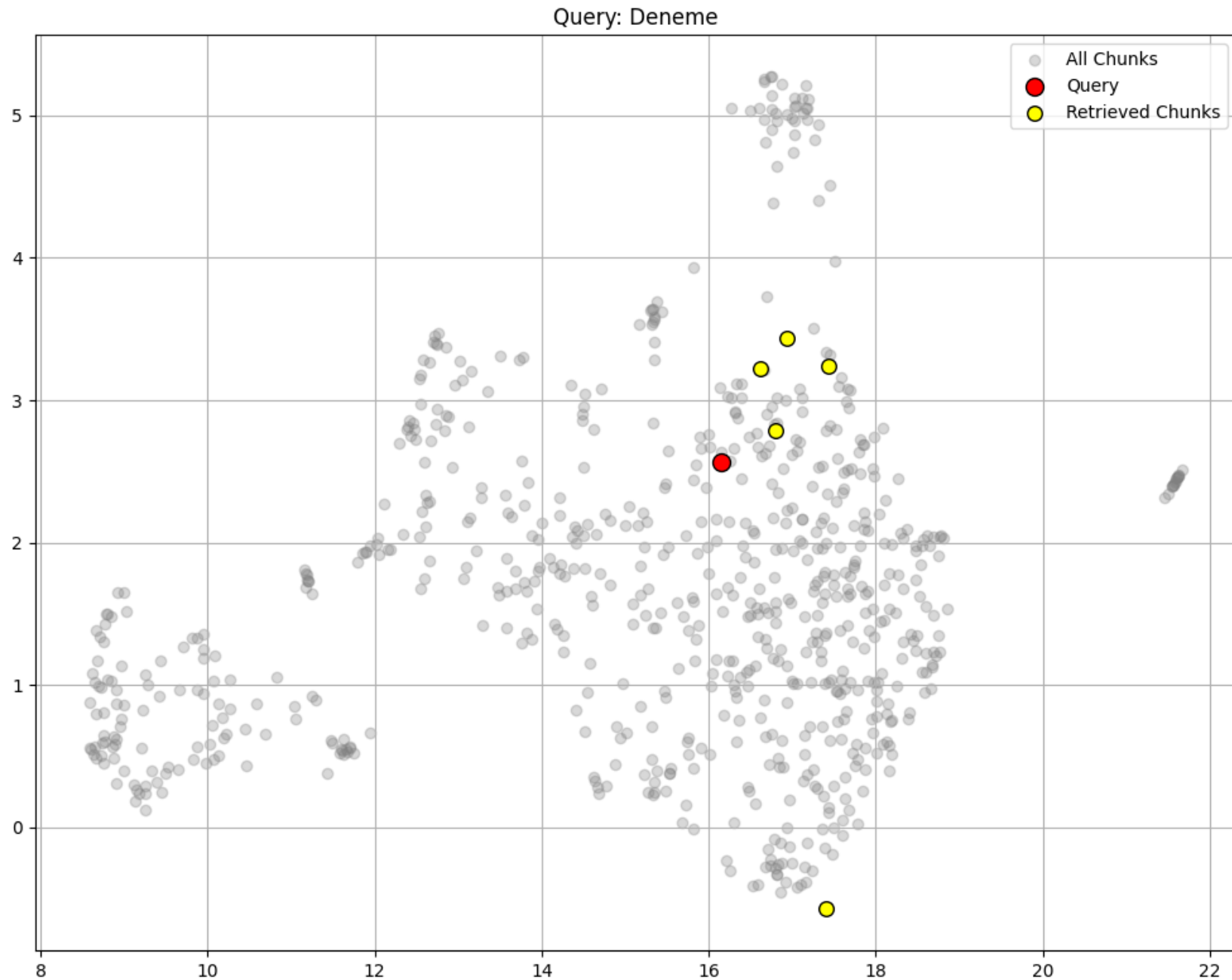
A **radial tear of the anterior aspect of the lateral meniscus** is identified, which is a **complete disruption** leading to **meniscal instability**. This type of tear is **common in younger patients** and can be **catastrophic** due to the **loss of structural integrity** of the meniscus. No significant abnormalities are noted in the **medial meniscus**, **ACL**, or **PCL**. There is also evidence of **patellar tendonitis** and a **small trochlear cartilage lesion**.

### Recommendation:

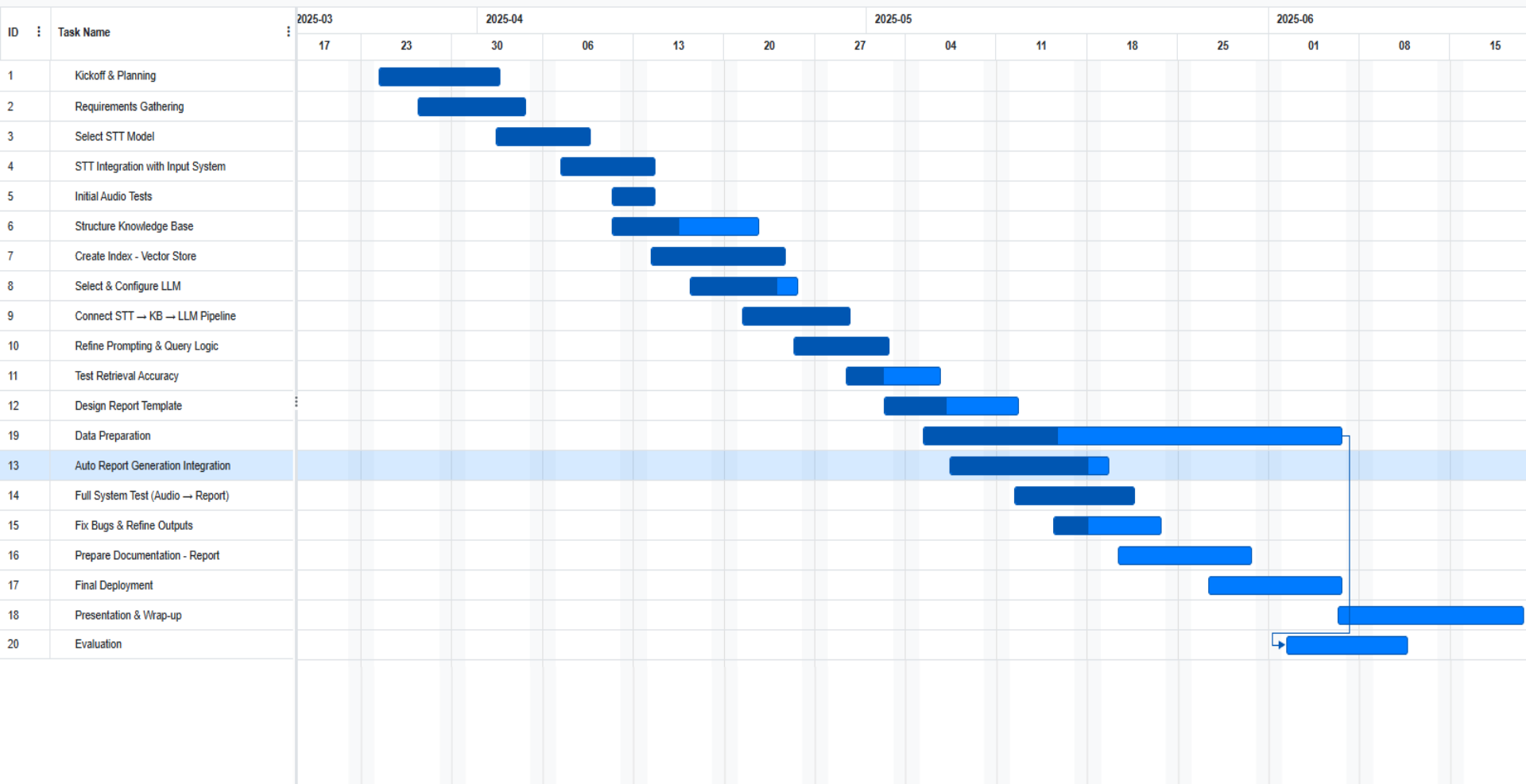
Further clinical correlation and possible **surgical evaluation** are recommended given the **complete nature of the radial tear**. Consideration of **arthroscopic evaluation** may be warranted depending on the patient's symptoms and functional status.



# Example Output – Embedding Visualization



# Project Timeline



# Data Things

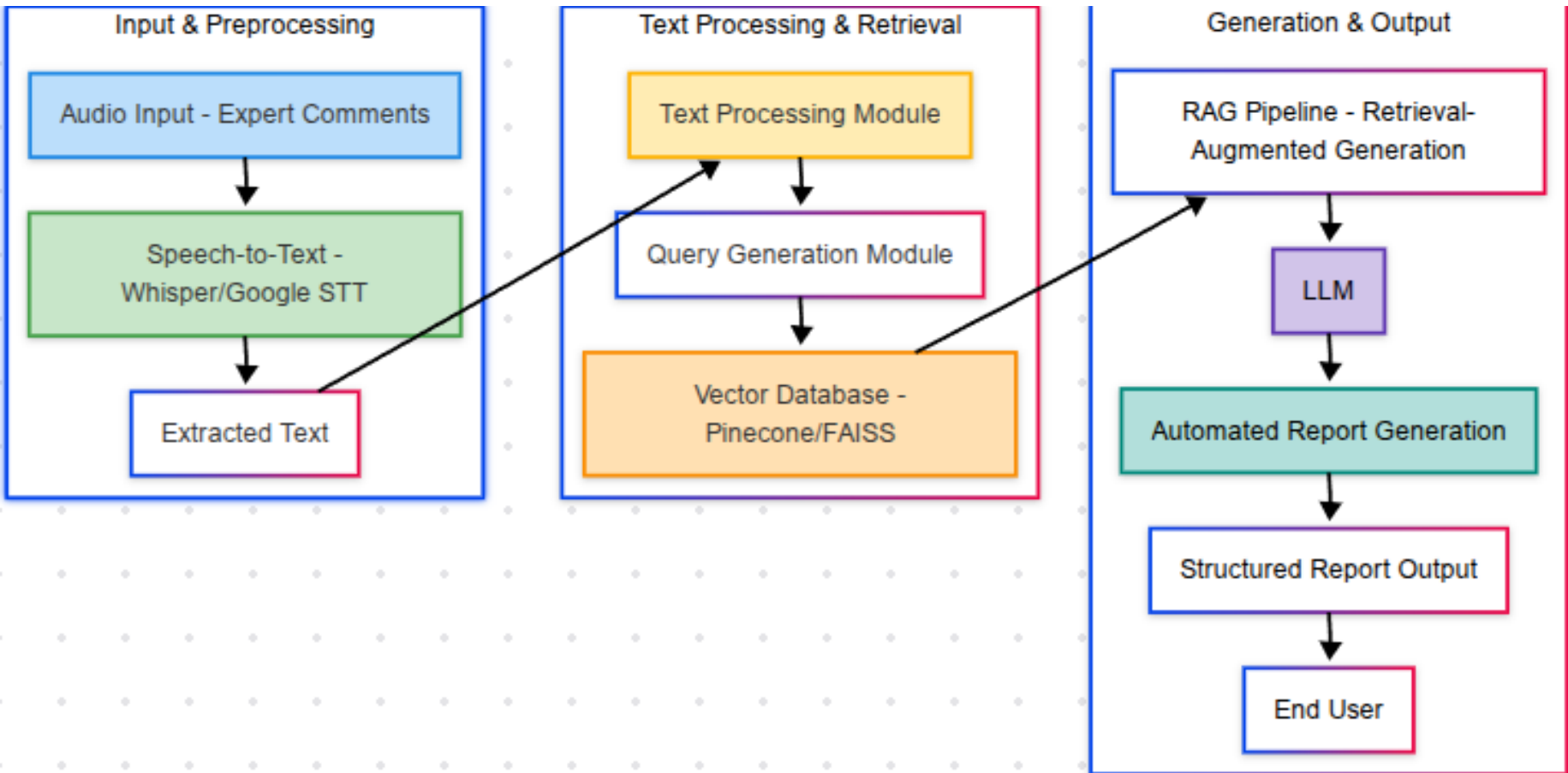
```
</Journal>
<ArticleTitle>Indiana University Chest X-ray Collection</ArticleTitle>
▼<Abstract>
  <AbstractText Label="COMPARISON">CT XXXX image XXXX, CT abdomen which included the lower thorax XXXX.</AbstractText>
  <AbstractText Label="INDICATION">Chest pain today. History of stent placement 7+ years ago.</AbstractText>
  <AbstractText Label="FINDINGS">The XXXX examination consists of frontal and lateral radiographs of the chest. The cardiac silhouette is not
  Calcified granuloma is again seen in the right upper lobe. There is no consolidation, pleural effusion or pneumothorax.</AbstractText>
  <AbstractText Label="IMPRESSION">Increased size of density in the left cardiophrenic XXXX. Primary differential considerations include incr
  is recommended. These findings and recommendations were discussed XXXX. XXXX by Dr. XXXX XXXX telephone at XXXX p.m. XXXX/XXXX. Dr. XXXX<XX
</Abstract>
<Affiliation>Indiana University</Affiliation>
▼<AuthorList CompleteYN="Y">
  ▼<Author ValidYN="Y">
    <LastName>Kohli</LastName>
```

Legal – Courtroom, Education – How-To Tutorials, Business – Finance Calls,  
Business – Meetings, Aviation – ATC





# Project Design Plan



## 🔧 Implementation Overview:

### Step 1: Speech-to-Text (STT)

Use pre-trained STT models to transform expert audio comments into accurate textual format.

### •Step 2: Text Processing

Clean and analyze the extracted text, identifying key topics, summarizing content, and preparing data for querying.

### •Step 3: Query Generation

Generate meaningful queries automatically from the processed text to retrieve relevant information from our knowledge base.

### •Step 4: Knowledge Retrieval (RAG)

Employ **Retrieval-Augmented Generation (RAG)** by querying a vector database to find relevant information based on semantic similarity.

### •Step 5: Large Language Model (LLM)

Integrate a pre-trained LLM to synthesize retrieved knowledge with the original expert insights into a coherent narrative.






### •Step 6: Automated Report Generation

Automatically produce structured, easily understandable reports for the end-user.







## Technological Components:




-  **Speech-to-Text Model** (*Pre-trained*)
  - Whisper (API or Local) / Google Cloud STT API
-  **Large Language Model (LLM)** (*Pre-trained*)
  - GPT-4o / OpenAI API or similar generative local/API model
-  **Vector Database** (*Semantic search & retrieval*)
  - Pinecone / FAISS / Weaviate/ChromaDB
-  **Text Processing Tools**
  - NLP libraries (spaCy, Hugging Face Transformers, NLTK)
-  **Report Generation & Templating**
  - Jinja2, python-docx, or similar templating tools

## Data Requirements:





-  **Audio Data** (*Expert Comments*)
  - Audio files clearly recorded, properly labeled, in formats like .wav or .mp3
-  **Knowledge Base** (*Existing Structured Data*)
  - Structured data or documents to be converted into a semantic searchable knowledge base



## Technical Environment:




-  **Programming Language**
  - Python (preferred, recommended 3.9+)
-  **Cloud or Local Infrastructure**
  - Cloud: AWS / Azure / GCP
  - Local: Sufficient CPU/GPU resources if deploying locally
-  **Development Environment**
  - VS Code / PyCharm / Jupyter Notebook

## What Should I be capable of?:


-  Basic Knowledge of NLP & AI
-  Python programming experience
-  API integration & data engineering skills
-  Familiarity with cloud deployment (optional)



## Accuracy and Performance:

-  **Speech-to-Text Accuracy:**
  - Minimum **85% transcription accuracy** on clearly recorded audio inputs.
-  **Information Retrieval (RAG) Accuracy:**
  - Relevant knowledge retrieved correctly at least **70% of the time** when queried.
-  **Report Generation Quality:**
  - Generated reports should clearly address key topics and expert insights accurately, with at least **70% accuracy based on manual validation from doctors if possible.**



## Performance and Efficiency:

-  **Response Time:**
  - The entire pipeline (audio-to-report) should produce outputs within **3 - 15 minutes per audio file** (audio length up to ~5 minutes, 2-3 minutes for cloud infrastructure with small models and up to 15 minutes for mid-large local models depend on hardware).






## Usability and Presentation:

-  **Report Clarity:**
  - Generated reports are understandable and rated as “clear and helpful” by **at least 70% of test users**.
-  **Ease of Use:**
  - %70 of Non-technical end-users can generate reports successfully after **minimal training (less than 15 minutes)**.



## Integration and Stability:

-  **Easy Integration:**
  - Clearly documented setup that can be reproduced within **1-2 hour by developers familiar with Python**.



- [openai.com/blog/whisper](https://openai.com/blog/whisper)
- [cloud.google.com/speech-to-text/docs](https://cloud.google.com/speech-to-text/docs)
- [python.langchain.com/docs/use\\_cases/question\\_answering](https://python.langchain.com/docs/use_cases/question_answering)
- [github.com/openai/openai-cookbook/blob/main/examples/Semantic\\_text\\_search\\_using\\_embeddings.ipynb](https://github.com/openai/openai-cookbook/blob/main/examples/Semantic_text_search_using_embeddings.ipynb)
- [pinecone.io/docs](https://pinecone.io/docs)
- [platform.openai.com/docs/api-reference](https://platform.openai.com/docs/api-reference)
- [huggingface.co/docs/transformers/index](https://huggingface.co/docs/transformers/index)
- [spacy.io/usage/spacy-101](https://spacy.io/usage/spacy-101)
- [nltk.org/book](https://nltk.org/book)
- [jinja.palletsprojects.com/en/3.1.x](https://jinja.palletsprojects.com/en/3.1.x)
- [python-docx.readthedocs.io/en/latest](https://python-docx.readthedocs.io/en/latest)
- [mermaid.js.org](https://mermaid.js.org)
- [lucidchart.com/pages/examples/basic-flowchart-template](https://lucidchart.com/pages/examples/basic-flowchart-template)



- <https://www.kaggle.com/datasets/paultimothymooney/medical-speech-transcription-and-intent>
- <https://www.kaggle.com/datasets/tboyle10/medicaltranscriptions>
- <https://physionet.org/content/egd-cxr/1.0.0/>
- <https://physionet.org/content/b2ai-voice/>
- <https://physionet.org/content/mimic-cxr>
- <https://www.oyez.org/>
- [https://huggingface.co/datasets/Jzuluaga/atcosim\\_corpus](https://huggingface.co/datasets/Jzuluaga/atcosim_corpus)
- <https://ieee-dataport.org/documents/fearless-steps-apollo-11-landing-audio>
- <https://datasets.kensho.com/datasets/spgispeech>
- <http://groups.inf.ed.ac.uk/ami/corpus/>
- <https://www.openslr.org/51>
- <https://huggingface.co/datasets/nbroad/mediasum>
- <https://www.kaggle.com/datasets/azmayensabil/doctor-patient-conversation-large>
- <https://github.com/srvk/how2-dataset>
- <https://commonvoice.mozilla.org/en/datasets>
- <https://www.openslr.org/12>
- <https://openi.nlm.nih.gov/>
- <https://huggingface.co/datasets/speechcolab/gigaspeech>
- <https://catalog.ldc.upenn.edu/LDC97S62>
- <https://catalog.ldc.upenn.edu/LDC2004S13>
- <https://www.kaggle.com/datasets/shuyangli94/interview-npr-media-dialog-transcripts>
- <https://keithito.com/LJ-Speech-Dataset/>
- <https://github.com/UCSD-AI4H/Medical-Dialogue-System>



- <https://www.ted.com/translate>
- <https://catalog.ldc.upenn.edu/LDC2012S06>
- <https://www.openslr.org/108/>
- <https://huggingface.co/datasets/ysdede/khanacademy-turkish-math>
- <https://huggingface.co/datasets/ysdede/khanacademy-turkish>
- [https://huggingface.co/datasets/mozilla-foundation/common\\_voice\\_17\\_0/viewer/tr/train](https://huggingface.co/datasets/mozilla-foundation/common_voice_17_0/viewer/tr/train)
- <https://catalogue.elra.info/en-us/repository/browse/ELRA-S0206/>
- <https://www.clarin.eu/parlamint>
- <https://github.com/onurgu/turkish-parliament-texts>
- <https://www.kaggle.com/datasets/batuhankalem/turkish-law-dataset-for-llm-finetuning>
- <https://huggingface.co/datasets/KocLab-Bilkent/turkish-constitutional-court>
- <http://www.tnc.org.tr/>
- <https://huggingface.co/incidelen/MedTurkQuAD>
- <https://huggingface.co/KocLab-Bilkent/turkish-constitutional-court>
- <http://tscorpus.com>
- <https://dergipark.org.tr/tr/pub/veterinary>
- <https://saglik.gov.tr/tr/112-egitim-kitabi>
- <https://huggingface.co/datasets/yusufbaykaloglu/turkish-finance-dataset>
- [http://kemik.yildiz.edu.tr/datasets/42bin\\_haber.rar](http://kemik.yildiz.edu.tr/datasets/42bin_haber.rar)
- <https://archive.ics.uci.edu/ml/datasets/TTC-3600>
- <https://www.kap.org.tr/tr/Duyuru/AyrintiliVeri>
- <https://www.tcmb.gov.tr/wps/wcm/connect/tr/tcmb+tr/main/meetings/minutes>
- <https://data.mendeley.com/datasets/twnertc/1>
- <https://github.com/bilkent-nlp/turkish-writings-dataset>



# THANK YOU FOR LISTENING

