



Microsoft

byte^{XL}

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

presents

HackXcelerate 2024

Empower Innovation with Microsoft and ByteXL

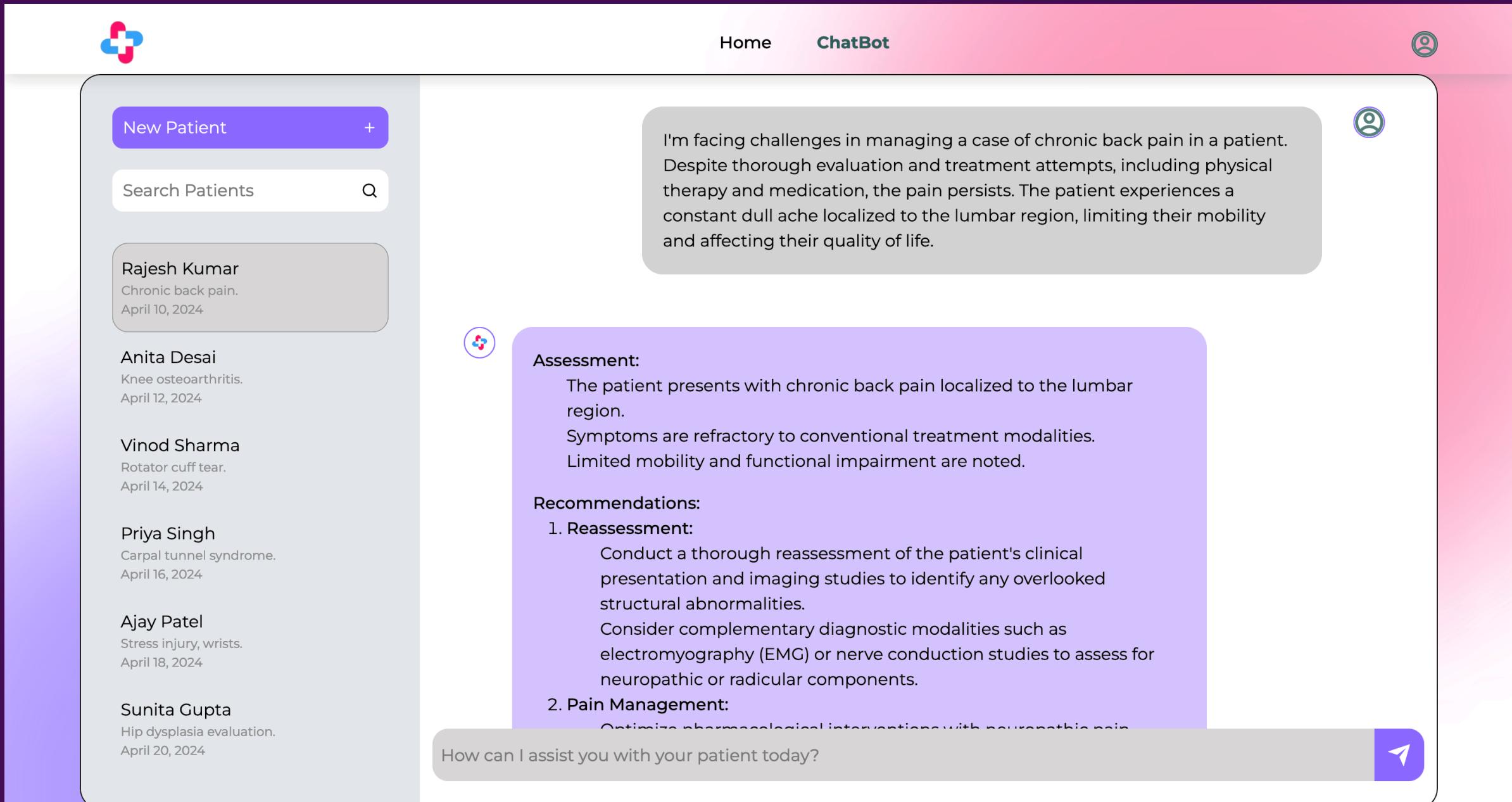
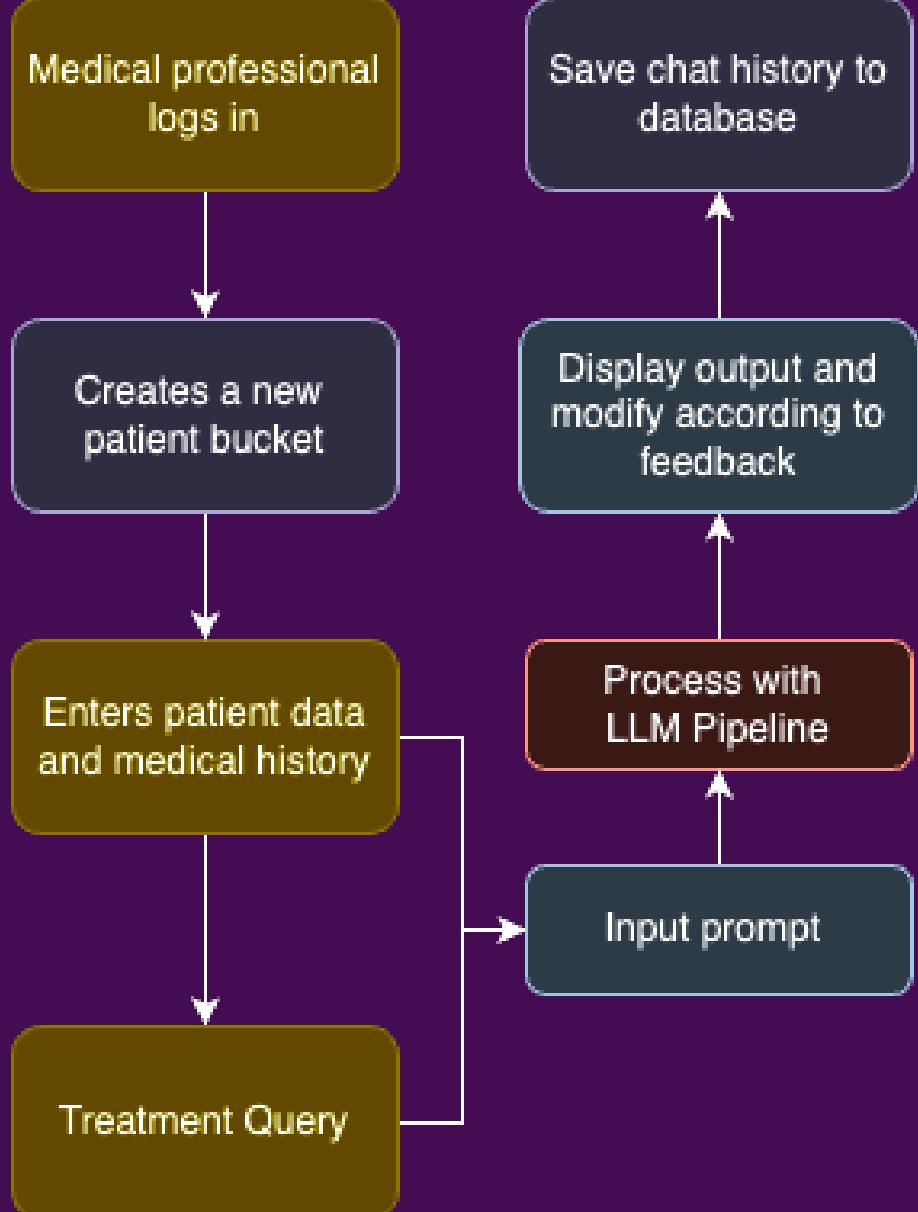


Med-RAGa

AI – Enhanced Diagnostic & Treatment Planning System

**Ctrl+Shift+Win
Omkar Kabde
CBIT**

PROTOTYPE OVERVIEW



Sample interaction with chatbot

User flow representation

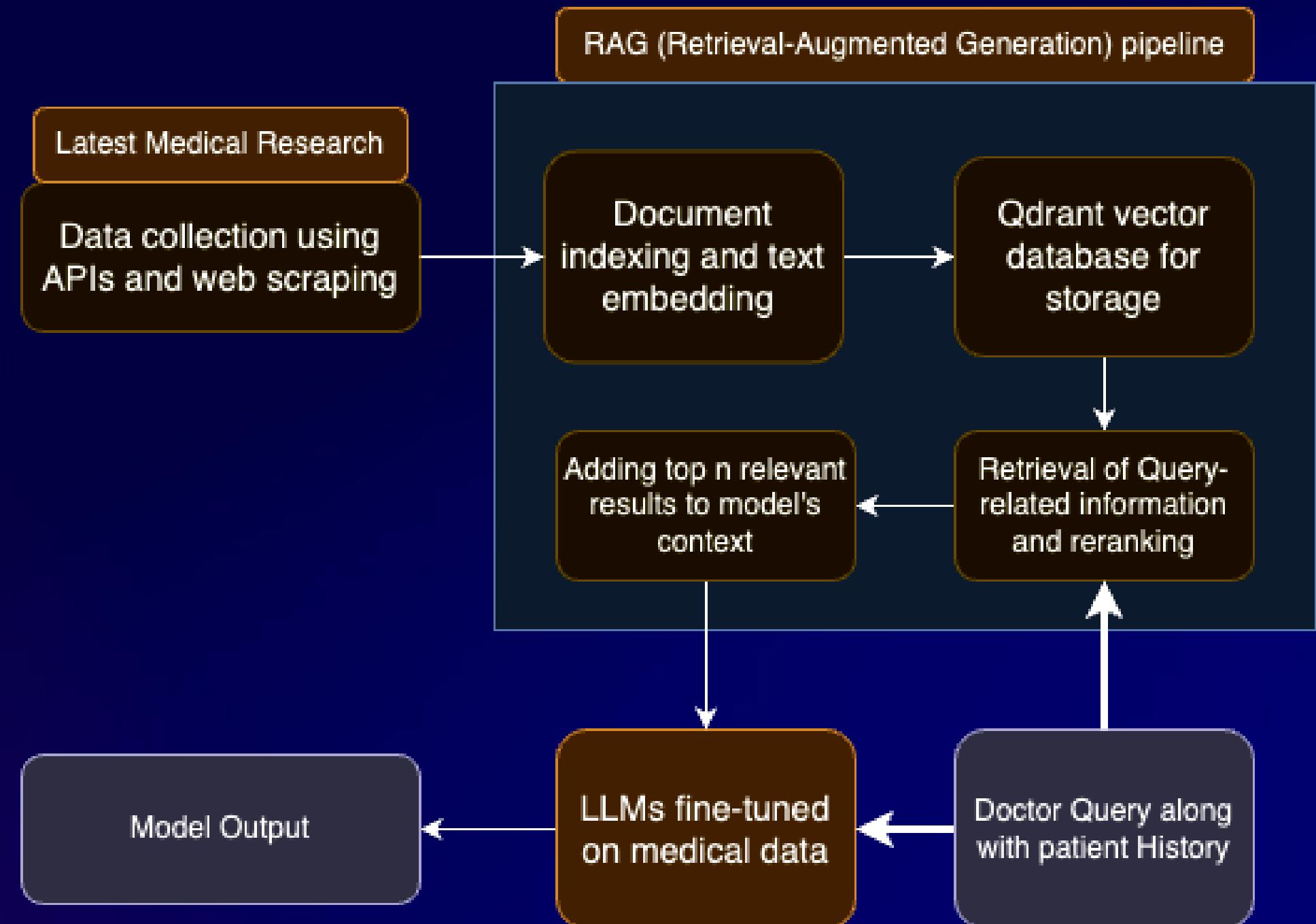
INNOVATION/UNIQUENESS

1. RAG (Retrieval-Augmented Generation): The use of RAG in the project combines the strengths of both retrieval-based and generation-based models to enhance the quality and relevance of the information provided to doctors. In the RAG pipeline, the retrieval component first identifies the most relevant and recent medical research articles and information from the indexed database. The retrieved information is augmented with the doctor's query and the patient's medical history. This augmented context provides a comprehensive understanding of the specific medical scenario at hand.
2. Getting latest medical information: Our solution uses APIs and web scraping to collect the newest medical research from well-known and trusted medical journals and websites. This helps ensure that the information is accurate and up-to-date, which is very important for correct diagnosis and treatment planning.
3. Personalizing for each patient: The model takes into account the unique medical history and details of each patient. By including a summary of the patient's history in the prompt, it can provide personalized diagnoses and treatment plans that are suitable for that specific patient. This allows the model to better understand the context of the doctor's questions and the patient's medical history.
4. Trusted information sources: Our solution only collects information from well-known and trusted medical journals and websites. This ensures that the information used for diagnosis and treatment planning is accurate and reliable, which makes the recommendations more trustworthy.

ABSTRACT

Latest medical research is obtained using APIs and web scraping. Select reputed medical journals and websites are allowed to ensure the authenticity of the data. PDF files and articles are downloaded and then indexed. These are then converted to text embeddings using Cohere and stored as vectors in a vector database (Qdrant).

When a doctor query's the prototype, it first retrieves the most relevant information and adds it to the LLM's context window. The LLM takes the new context window and the original prompt, along with the summarized patient history, and returns the output.



TECH STACK

- API – arXiv, PubMed
- Web Scraping - Selenium, Playwright, BeautifulSoup
- Text Embeddings – Cohere
- Vector Database – Qdrant
- Retriever Model – BioMistral

Backend

- Langchain
- FastAPI
- Python

Front-End

- React
- Javascript

Database

- Firestore

USECASES

1. **Automated Treatment Suggestion:** Medical professionals can use the application to automatically generate treatment suggestions. This helps in reducing the time spent on researching and planning treatments, allowing healthcare providers to focus more on patient care
2. **Training and Continuous Learning:** Doctors and other medical professionals can use the application as a learning tool to understand possible treatment options and approaches, particularly for complex cases, by interacting with the LLM and observing its recommendations
3. **Research Contribution:** By analyzing the aggregate data on patient outcomes and treatment efficiency, the application can contribute to medical research.

TEAM CONTACT DETAILS

Team Lead

Name: Omkar Kabde

Phone Number: 8500383520

Team Member-1

Name: Chittappu Abhiram Reddy

Phone Number: 9392549463

Team Member-2

Name: Mohammed Imaduddin

Phone Number: 9052812005

All team members are from

Chaitanya Bharathi Institute of Technology

