

Introduction and Objectives

- Develop an AI-powered chatbot (named NIRA) to assist employees in understanding HR policies, particularly leave policies.
- Utilize **Retrieval-Augmented Generation (RAG)** to provide accurate, context-based answers using a predefined document.
- Enhance accessibility of HR policy documents through conversational AI.

Research Gaps

- Lack of intelligent HR bots that strictly rely on company-specific documents without hallucination.
- Traditional chatbots are rule-based and do not handle complex queries contextually.
- Difficulty in retrieving specific policy details from long documents without manual search.

Methodology and Techniques

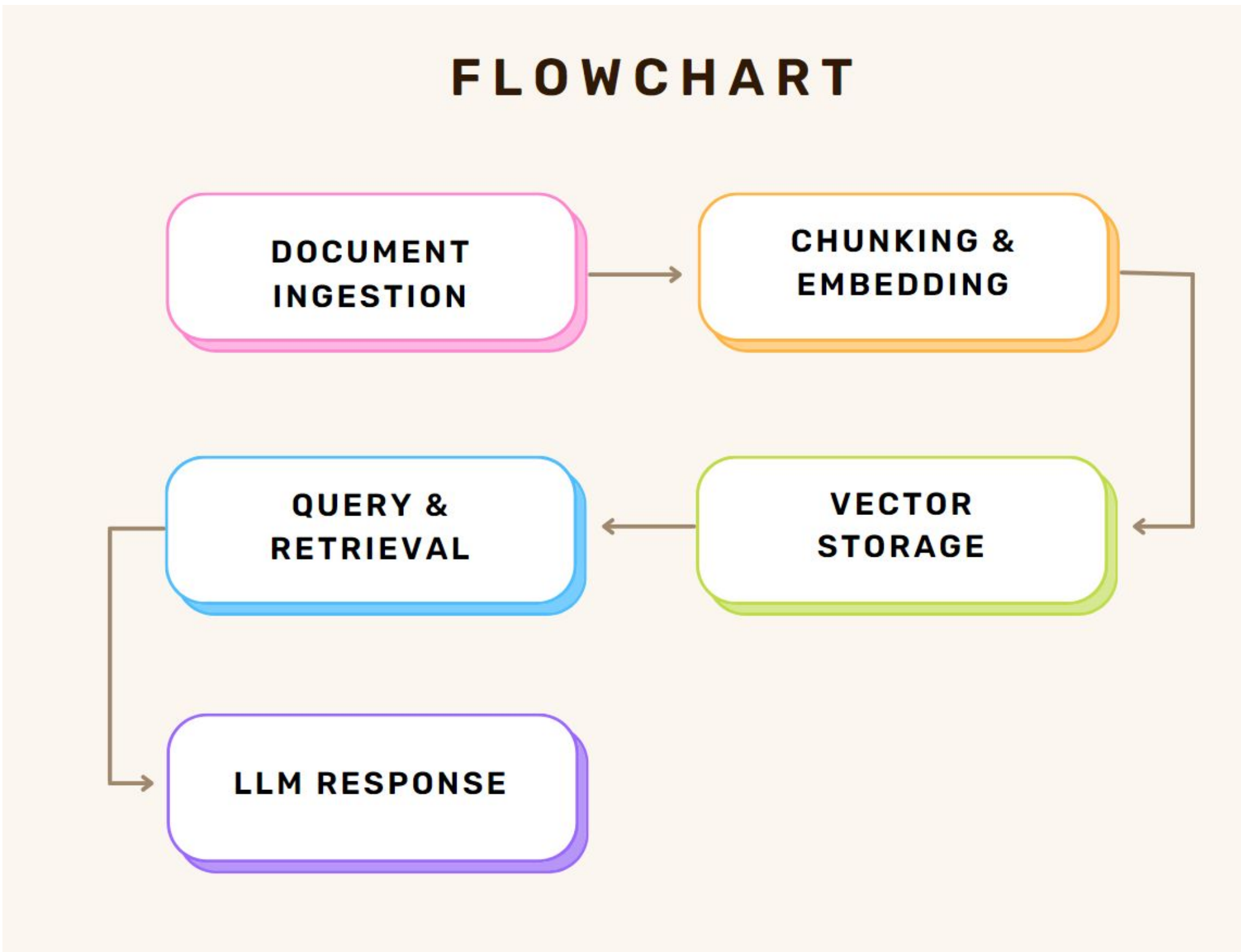


Fig 1. Flowchart of Methodology

Document Processing:

- Load HR policy documents using `TextLoader`.
- Split text into manageable chunks using `RecursiveCharacterTextSplitter`.

Embedding and Storage:

- Generate vector embeddings using `HuggingFaceEmbeddings` (`all-MiniLM-L6-v2`).
- Store and persist vectors using `ChromaDB`.

RAG Workflow:

- Retrieve top relevant chunks using similarity search.
- Construct a strict and detailed prompt using retrieved context.
- Generate structured answers using the **Ollama LLM API** (`qwen2.5:3b` model).

Interaction:

- User inputs a query about leave policy.
- Bot responds based solely on available document chunks.

Work Completed and Results

Successfully implemented:

- Successfully implemented a document ingestion and chunking pipeline to preprocess HR policy text for downstream processing.
- Created and persisted a semantic vector store using ChromaDB and HuggingFace embeddings for efficient retrieval.
- Designed prompt templates ensuring all responses remain accurate, structured, and policy-aligned.
- Developed a CLI-based AI assistant (NIRA) integrated with Qwen LLM via Ollama to answer leave-related HR queries.

Results:

- Accurately responds to employee queries using only the verified HR policy content.
- Delivers structured and professional answers aligned with company guidelines.
- Effectively rejects queries beyond the document scope, ensuring strict compliance and reliability.

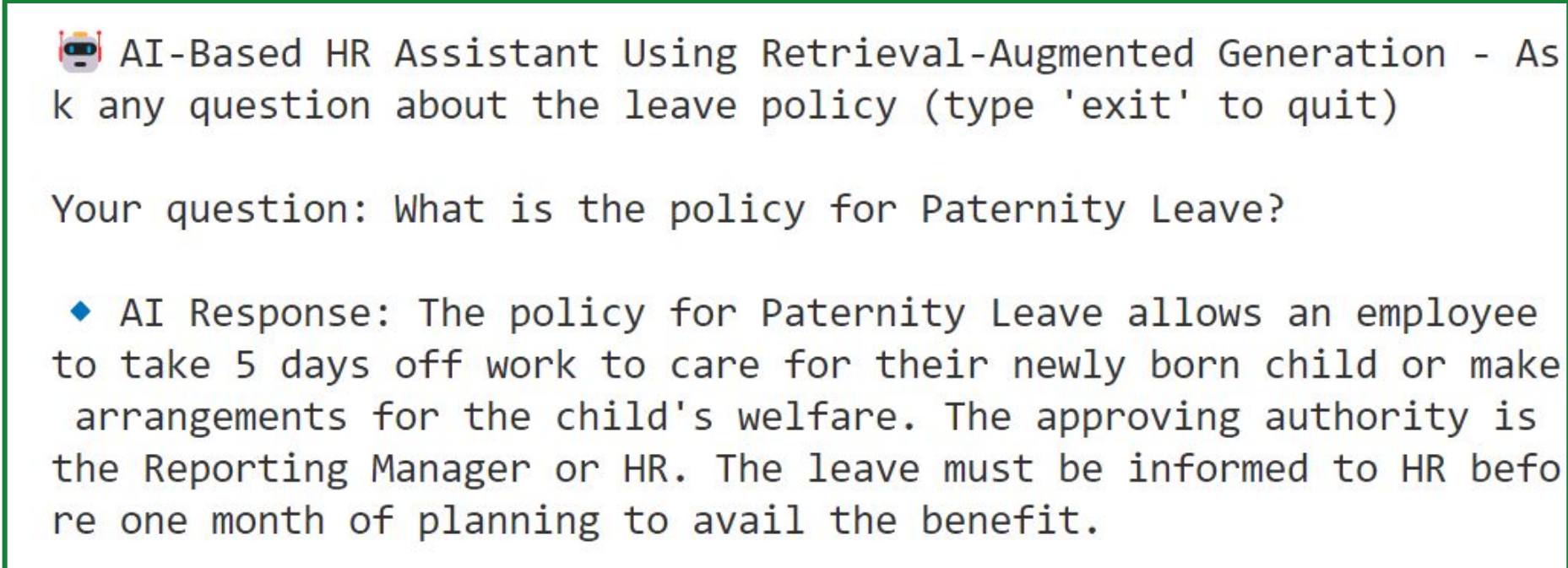


Fig 2. Chatbot Response

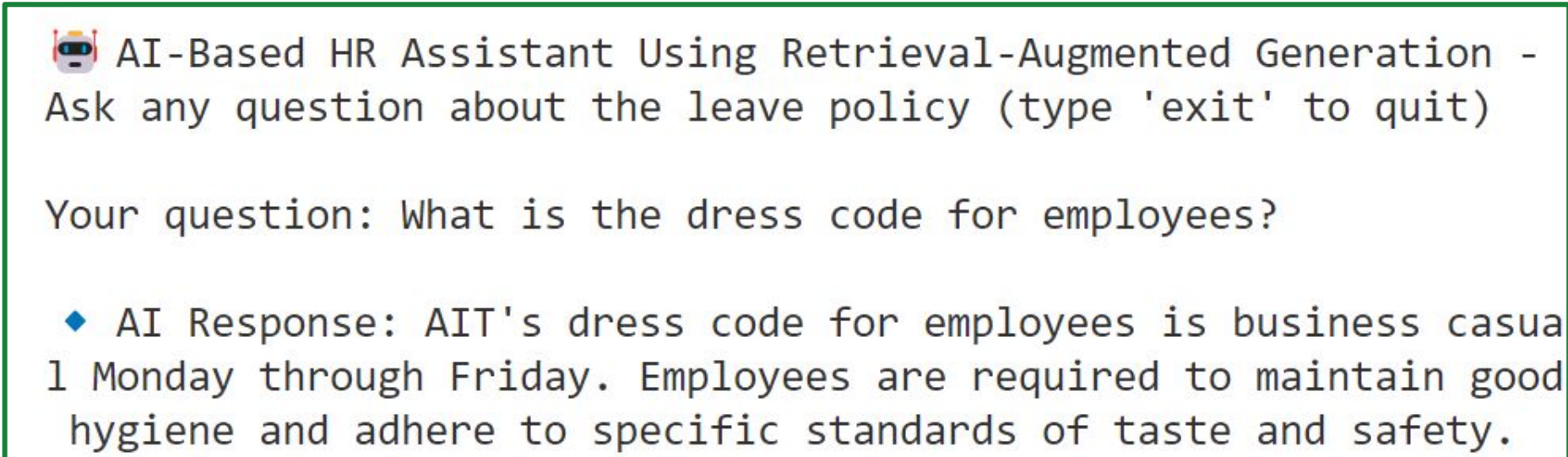


Fig 3. Chatbot Response for different query

Conclusion and Remaining Work

Conclusion:

- The system is effective for answering document-bound HR policy queries.
- Reduces manual effort and improves accessibility to HR information.

Remaining Work:

- Extend functionality to include more HR documents (e.g., onboarding, benefits).
- Add web-based integration for ease of use.

Bibliography/ References

- LangChain documentation: <https://docs.langchain.com/>
- HuggingFace Transformers: <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>
- Ollama API for LLMs: <https://ollama.com>
- ChromaDB: <https://www.trychroma.com/>
- RAG Concept: Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," NeurIPS 2020