🧠✨ Multi-Modal RAG Jira Plugin

Project Overview

This project implements a Multi-Modal Retrieval Augmented Generation (RAG) Jira plugin designed to provide intelligent, context-aware insights across Jira projects and enterprise documents.

The system ingests Jira project JSON data and multiple document formats, transforms them into embeddings, and stores them in Weaviate Cloud for efficient vector-based retrieval. The solution enables users to query Jira data and documents using natural language and receive accurate, summarized responses powered by LLMs.

📐 System Architecture & Data Flow

🔹 Data Ingestion & Indexing

1. **Jira project data is fetched using the Jira REST API**
2. **Raw JSON data is:**
    - **Cleaned**
    - **Transformed**
    - **Chunked**
3. **The processed data is embedded and stored in Weaviate Cloud as vector indexes**

In addition to Jira data, the system supports multi-modal document ingestion, including:

- 📄 **PDFs**
- 📝 **Word documents**
- 📊 **PowerPoint files**
- 🖼️ **Images**
- 🔊 **Audio files**

All supported files are ingested, embedded, and indexed into Weaviate Cloud, enabling unified semantic search across data types.

🔍 Query Processing & RAG Pipeline

1. **The React frontend sends user queries to backend APIs**
2. **The RAG pipeline performs:**
    - 🔍 **Semantic retrieval from Weaviate Cloud**
    - 🧠 **Context filtering and ranking**
3. **Only the user query + relevant retrieved context is sent to the LLM API**
4. **OpenAI models generate a concise, grounded response**
5. **The final summarized answer is returned to the frontend**

This approach ensures:

- **Reduced token usage**
- **Minimal hallucinations**
- **Highly relevant, context-aware responses**

⚙️ Key Technologies Used

- ⚛️ **React – User interface**
- 🛡️ **Django / FastAPI – Backend APIs & middleware**
- 🧠 **RAG Pipeline – Context-aware AI responses**
- 📦 **Weaviate Cloud – Vector database for embeddings**
- 🤖 **OpenAI – Language model for summarization and Q&A**
- 📜 **Jira REST API – Project and issue data source**

🌟 Key Features

- 🔗 **Jira-native AI querying**
- 📚 **Multi-modal document support**
- 🧠 **Semantic vector search**
- 📈 **Scalable cloud-based indexing**
- 🔐 **Secure and modular architecture**
- ⚡ **Optimized for enterprise workloads**

📜 Summary

This Multi-Modal RAG Jira Plugin provides a scalable and intelligent solution for querying Jira projects and enterprise documents using natural language. By combining Weaviate Cloud, RAG pipelines, and LLMs, the system delivers accurate, context-aware insights while maintaining performance, scalability, and maintainability.