WhatsApp AI Agent with n8n + OpenAI + Pinecone + Google Docs

Overview

This project demonstrates a fully automated **Retrieval-Augmented Generation (RAG) chatbot** built on **n8n** that responds to WhatsApp messages with contextually relevant information drawn from company documents.
It integrates **WhatsApp Business API**, **OpenAI**, **Pinecone**, and **Google Docs** to create an intelligent assistant that can search, reason, and respond conversationally in real time.

Workflow Architecture

1. Document Indexing (Docs → Pinecone)

- **Trigger:** Manual or scheduled execution in n8n.
- **Fetch Document:** Retrieves reference documents from Google Docs.
- **Text Processing:** Splits the content using a Recursive Character Text Splitter for better embedding performance.
- **Embedding Creation:** Converts text chunks into vector embeddings using **OpenAI Embeddings**.
- **Storage:** Saves the embeddings in **Pinecone Vector Store** under a defined namespace (`waagent`).
  (File: `Docs to Pinecone.json`)

  Docs to Pinecone

2. Chat Interaction (WhatsApp → AI Agent)

- **Trigger:** Incoming WhatsApp messages via **WhatsApp Business API**.
- **Memory:** Uses session-based memory to maintain conversation context across messages.
- **Retrieval:** Pulls the most relevant document chunks from Pinecone based on user query embeddings.
- **Response Generation:** Uses **OpenAI GPT-4 model** to compose a grounded, natural language response combining context and intent.
- **Execution Tools:**
  - **Wikipedia** – for factual lookups.
  - **Calculator** – for computations.
  - **Gmail Tool** – for drafting and sending structured emails.
- **Output:** The final AI-generated response is sent back to the user via WhatsApp.
  (File: `WA Service Docs.json`)

  WA Service Docs

Tech Stack

- **n8n** – Orchestrates the workflows and connects all nodes.
- **OpenAI API** – Generates embeddings and chat responses.
- **Pinecone** – Manages semantic search and vector retrieval.
- **Google Docs** – Serves as the data source for reference material.
- **WhatsApp Business API** – Provides real-time message interface with users.

Key Features

- Context-aware chatbot responding directly via WhatsApp.
- Dynamic RAG pipeline for document-based question answering.
- Multi-tool integration (Wikipedia, Calculator, Gmail).
- Session memory to maintain conversational continuity.
- Modular workflows for document ingestion and live query handling.

Use Cases

- **Customer support bots** powered by internal documentation.
- **Enterprise knowledge assistants** for HR, IT, or sales teams.
- **Automated business communication** via WhatsApp.
- **Internal query automation** for operational data retrieval.

Outcome

Once deployed, the workflow allows any authorized user to ask questions on WhatsApp and receive **real-time, accurate, and context-grounded answers**.
The system automatically fetches data from internal documents, processes it via embeddings, and delivers a personalized and intelligent conversational experience.