

What is a RAG Chatbot?

RAG = Retrieval-Augmented Generation

Instead of relying only on the model's memory, a RAG chatbot:

1. **Retrieves relevant documents** from a knowledge base (using vector embeddings)
2. **Feeds them into the language model**
3. The model **generates an answer** based on those documents

This is perfect for:

- Customer support bots
- Internal knowledge assistants
- AI tutors
- FAQ bots for your website

Components & Tools

Component	Tool / Service
Cloud Storage	Google Drive
Trigger System	Google Drive Trigger
Vector Store	Pinecone
Embeddings	OpenAI (text-embedding-ada-002)
Data Loader	Default Data Loader
Text Splitter	Recursive Character Text Splitter
Chat Model	OpenAI Chat Model (e.g., GPT-4 / GPT-3.5)
Memory	Simple Memory
Language	JSON-based backend configuration

Component	Tool / Service
Frontend	None (CLI or system integration only)

Key Features

- Automated ingestion from Google Drive
- RAG pipeline for up-to-date document-backed answers
- OpenAI-powered understanding + generation
- Memory-based continuity across chat turns
- No frontend needed; callable via API or CLI

Use Case – Webai RAG Bot

The Webai Agency RAG Chatbot AI Agent is a **backend-first, Retrieval-Augmented Generation (RAG) chatbot system**. It is designed to dynamically process and retrieve knowledge from uploaded documents (via Google Drive) using OpenAI embeddings and respond to user queries with grounded responses via a simple AI agent architecture.

For your **Webai Agency**, a RAG chatbot can:

- Answer FAQs automatically from a structured PDF or Notion doc
- Provide support on website/SEO issues
- Onboard new clients with custom responses
- Allow sales teams to pull up key project/service info instantly

Webai Agency RAG Chatbot – Overview

What it Does:

- Answers questions about:
 - Services (e.g., “Do you offer SEO?”)
 - Pricing (e.g., “How much is a landing page?”)
 - Timelines (e.g., “How long does website delivery take?”)
 - Support (e.g., “What happens after launch?”)
 - Future features (e.g., “When will AI bots be available?”)

- Pulls answers directly from a vectorized knowledge base
- Responds in a brand-aligned, professional tone