

> Menu

Backend AI System

Relevant source files

This document provides a comprehensive overview of the backend AI system in the AgenteInteligente project. The backend system is built using n8n workflows that implement Retrieval-Augmented Generation (RAG) to provide intelligent responses to user queries in different domains, including customer service and tax-related information.

For frontend implementation details, see [Frontend System](#).

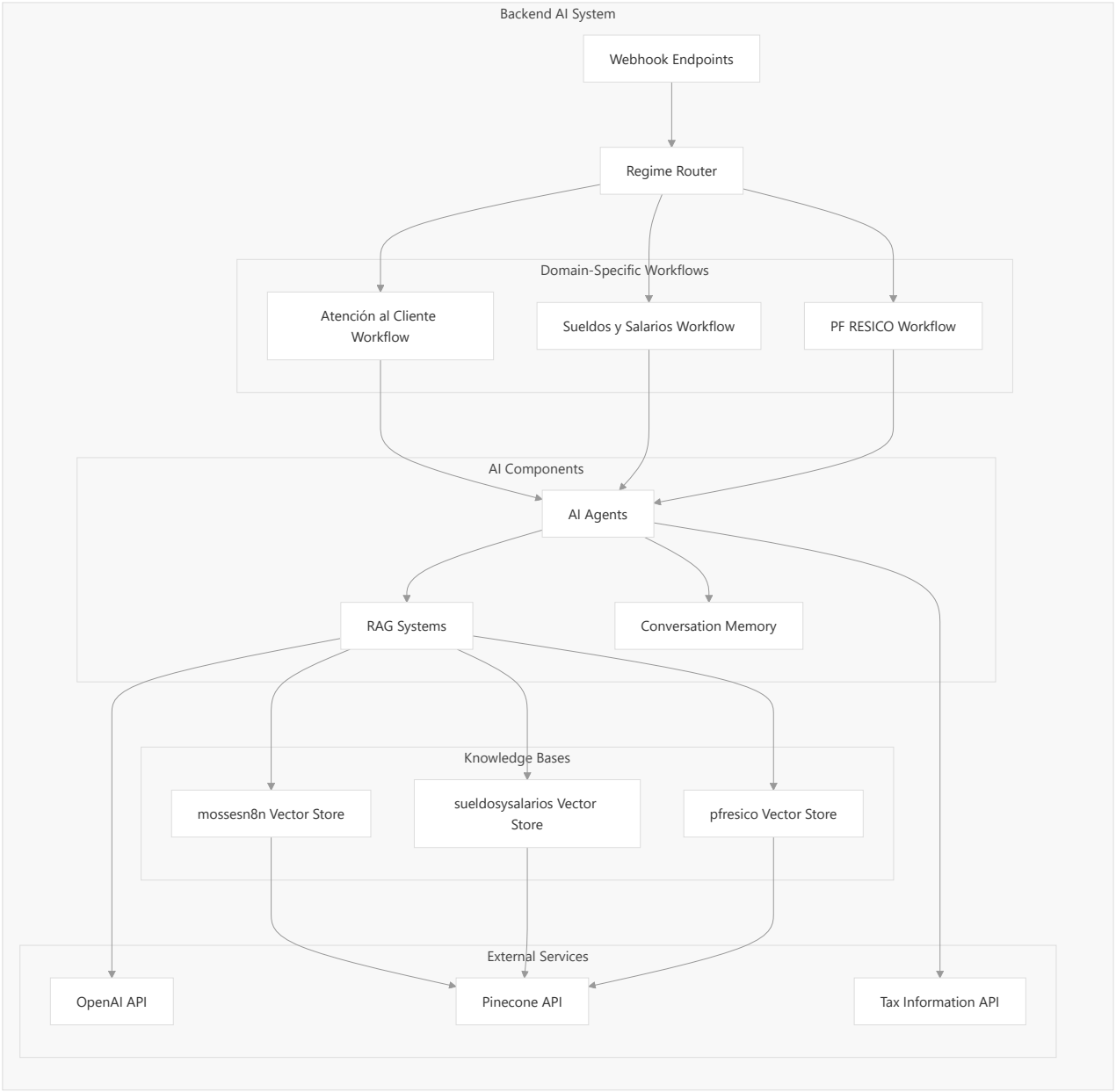
System Overview

The backend AI system processes incoming messages from users, routes them based on tax regime types, retrieves relevant knowledge from vector databases, and generates contextual responses using language models. The system is implemented as a collection of n8n workflows that handle different functional domains while following similar architectural patterns.

Ask Devin about axlRoman/AgenteInteligente

Deep Research





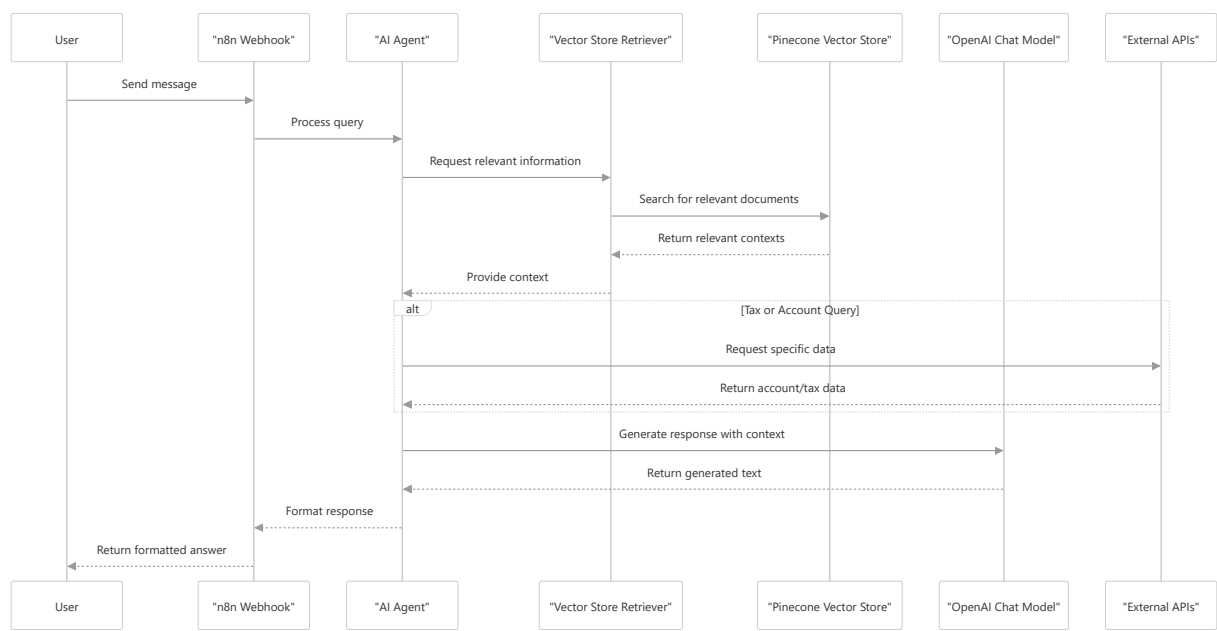
Sources: Atencion_al_Cliente.json | 1-550), PF_RESICO.json | 1-150),
Sueldos_y_Salarios.json | 1-150)

Retrieval-Augmented Generation Architecture

The system employs a RAG architecture that combines information retrieval with generative AI

Ask Devin about axlRoman/AgenteInteligente

Deep Research ☐



Sources: `Atencion_al_Cliente.json` | 10-130), `PF_RESICO.json` | 86-142)

Core Components

Webhook Endpoints

The system exposes several webhook endpoints that serve as entry points for different types of queries:

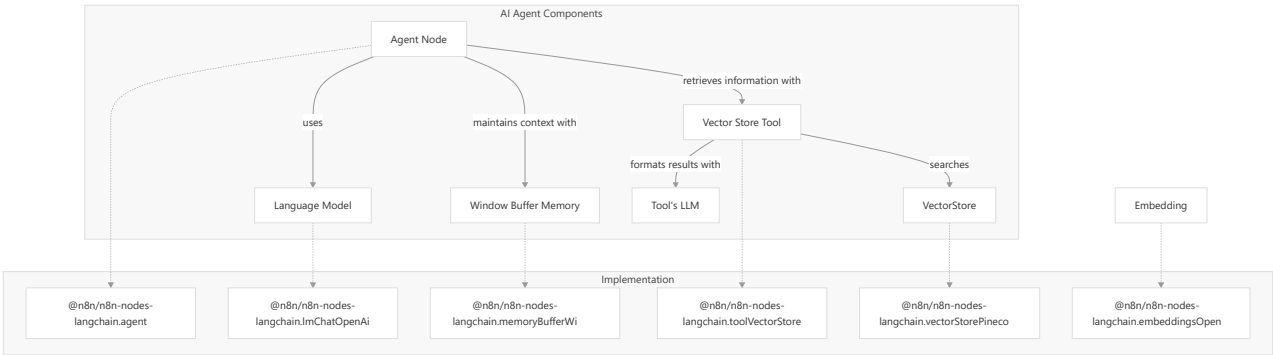
| Endpoint | Purpose | Implementation | |
|-------------------|-----------------------------------------------------|------------------------------|-----------|
| AtencionAlCliente | Handles general customer service inquiries | Atencion_al_Cliente.js on | 324-338 |
| SueldosYSalarios | Processes wage and salary related inquiries | Sueldos_y_Salarios.js n | 517-626 |
| pfResico | Manages queries related to the RESICO tax regime | PF_RESICO.json | 6-19 |
| regimenes | General entry point that routes to specific regimes | FlujoFinalMosses.js n | 1186-1268 |

Ask Devin about axlRoman/AgenteInteligente

Deep Research

The system uses LangChain's agent framework implemented in n8n to create conversational AI agents that can:

- Understand and respond to user queries
- Use tools (like vector stores) to retrieve information
- Maintain conversation context through memory systems
- Format responses appropriately



Sources: `Atencion_al_Cliente.json` | 4-17), `Atencion_al_Cliente.json` | 396-411)

Vector Stores

The system utilizes Pinecone vector stores to index and retrieve domain-specific knowledge:

| Vector Store Index | Purpose | Used By |
|-------------------------------|--------------------------------------|------------------------------|
| <code>mossesn8n</code> | General customer service information | Atención al Cliente workflow |
| <code>sueldosysalarios</code> | Salary and wage related information | Sueldos y Salarios workflow |
| <code>pfresico</code> | RESICO tax regime information | PF RESICO workflow |

Each vector store is accessed through Pinecone's API and utilizes OpenAI embeddings to convert text into vector representations.

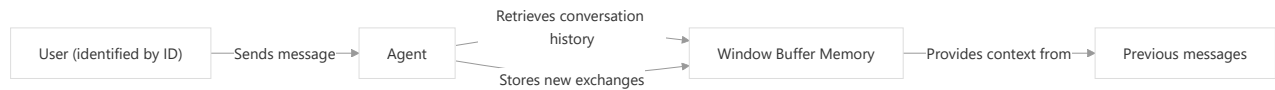
Sources: `Atencion_al_Cliente.json` | 196-217), `PF_RESICO.json` | 117-141),

Ask Devin about axlRoman/AgenteInteligente

Deep Research

The backend implements conversation memory using Window Buffer Memory, which:

- Maintains a history of back-and-forth interactions
- Uses session keys based on user identifiers
- Helps provide continuity and context awareness in conversations



Sources: `Atencion_al_Cliente.json` | 43-56), `PF_RESICO.json` | 615-626),
`Sueldos_y_Salarios.json` | 615-626)

Workflow Implementations

Domain-Specific Workflows

The system implements three primary domain-specific workflows:

1. Customer Service (Atención al Cliente)

- Handles general customer inquiries
- Uses the `mossesn8n` vector store index
- Provides support for customer service topics

2. Wages and Salaries (Sueldos y Salarios)

- Processes queries related to employment, wages, and salaries
- Uses the `sueldosysalarios` vector store index
- Can retrieve tax information related to employment income

3. RESICO Tax Regime (PF RESICO)

- Handles inquiries specific to the Simplified Tax Regime for Trust (RESICO)
- Uses the `pfresico` vector store index
- Specializes in tax regulations and calculations for this regime

Ask Devin about axlRoman/AgenteInteligente

Deep Research



Sources: Atencion_al_Cliente.json | 427-468), Sueldos_y_Salarios.json | 430-496)

Ask Devin about axlRoman/AgenteInteligente

Deep Research ☐

- Generating text responses via the Chat API (typically using gpt-4o-mini)
- Creating embeddings for document indexing and retrieval
- Processing and understanding user intents

These integrations are implemented through n8n nodes like:

- `@n8n/n8n-nodes-langchain.lmChatOpenAi`
- `@n8n/n8n-nodes-langchain.embeddingsOpenAi`

Sources: `Atencion_al_Cliente.json` | 21-41), `Atencion_al_Cliente.json` | 221-236)

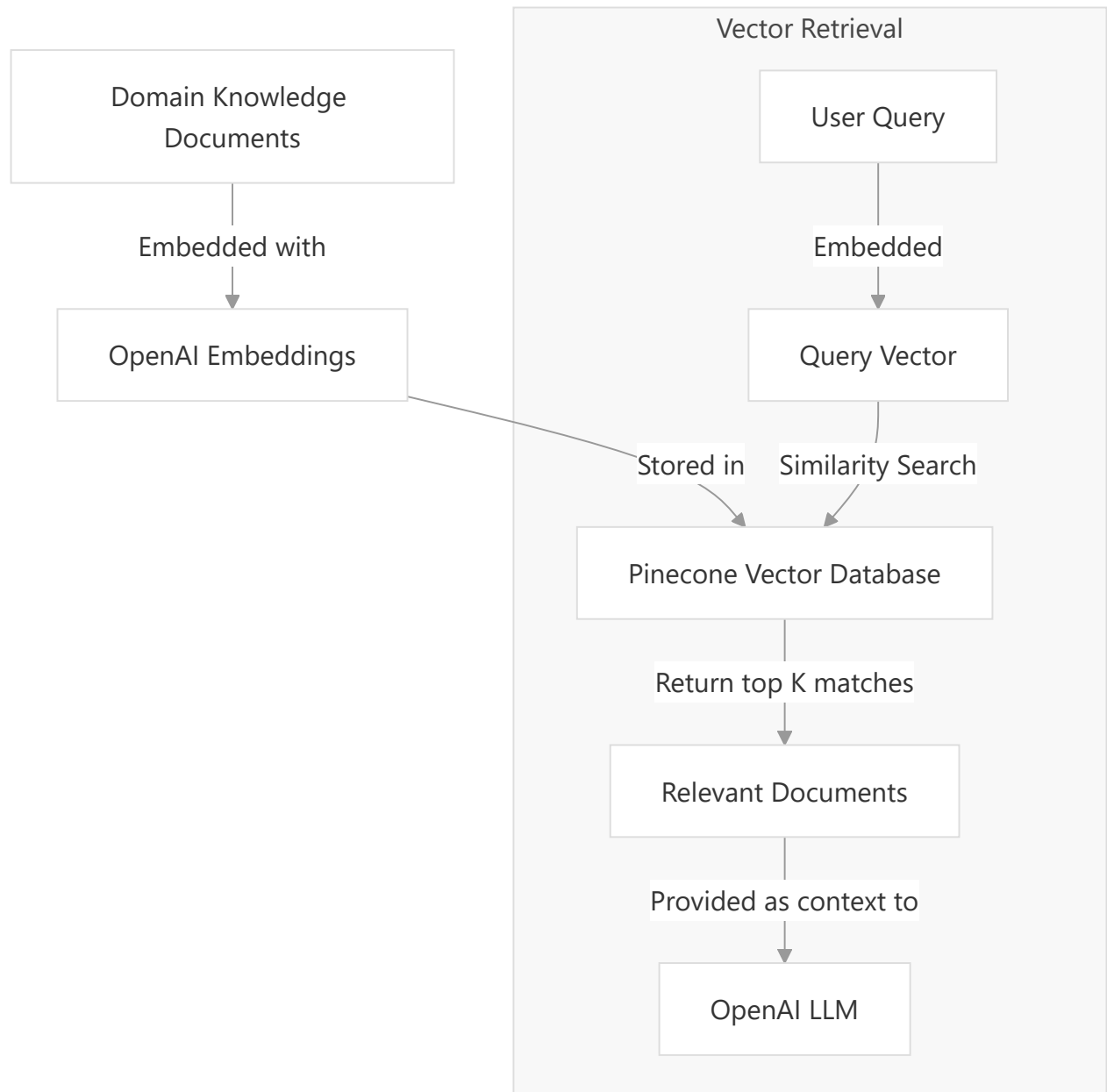
Pinecone Vector Database

The system uses Pinecone as its vector database with several indexes:

- Each domain has its own specialized index
- OpenAI embeddings are used to convert text to vectors
- Retrieval parameters like `topk` control how many documents are returned

Ask Devin about axlRoman/AgenteInteligente

Deep Research



Sources: `Atencion_al_Cliente.json` | 196-217), `Atencion_al_Cliente.json` | 266-275)

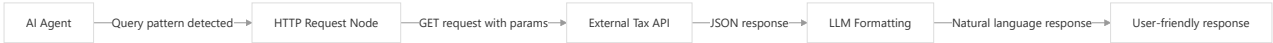
External APIs for Tax and Account Information

The system integrates with external tax information APIs to provide specific data about:

- User taxes (ISR, IVA)

Ask Devin about axlRoman/AgenteInteligente

Deep Research ☐



Sources: `Atencion_al_Cliente.json` | `547-567`), `Atencion_al_Cliente.json` | `733-746`)

Data Flow and Processing Logic

Input Processing

1. User sends a message through the frontend
2. Message arrives at n8n webhook endpoint
3. System routes the message based on the tax regime parameter
4. The appropriate workflow handles the message

DeepWiki axlRoman/AgenteInteligente

 Share



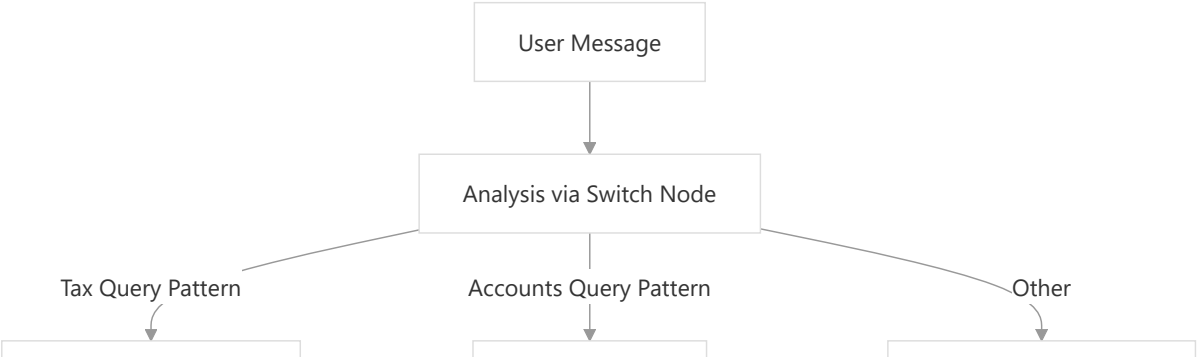
The workflows analyze queries using string pattern matching to determine:

- If it's a tax information request
- If it's an accounts receivable query
- If it's a general knowledge query

Each path has specialized processing logic.

Ask Devin about axlRoman/AgenteInteligente

Deep Research



Ask Devin about axlRoman/AgenteInteligente

Deep Research ☐