D	ro۱	/e	ct	n ı	า8	r
	יטו	<i>,</i>	U U	v	10	

Diana Marcela Salguero Sanchez

Programación III

Semestre 2025-1
Universidad tecnológica de Pereira

Especificaciones de mi computador:

Procesador: AMD Ryzen 5 4500U with Radeon Graphics (2.38 GHz)

RAM: 8,00 GB

Windows 11 home

Especificaciones de la máquina virtual instalada

Nombre: n8n_proyecto

Sistema operativo: Ubuntu (64-bit)

Memoria base: 2048 MB

Memoria de video: 16 MB

Objetivos

Un RAG en n8n se crea para potenciar las capacidades de generación de texto de modelos de lenguaje con datos personalizados o específicos, como documentos, bases de datos, PDFs, archivos Excel, imágenes, artículos, etc.

RAG (Retrieval-Augmented Generation) es una técnica en la que el modelo de lenguaje no genera respuestas solo desde su entrenamiento, sino que primero busca ("retrieves") información en una fuente externa y luego genera ("augmented generation") una respuesta basada en esa información.

¿Por qué crear un RAG en n8n?

n8n (plataforma de automatización) tiene funciones como acceso a modelos de lenguaje, almacenamiento de datos vectorizados, y más. Crear un RAG en este entorno te permite:

- 1- Automatizar respuestas inteligentes basadas en tu propia información:
 - Documentación técnica.
 - Manuales internos.
 - Base de conocimientos de tu empresa.
- 2- Mantener el control y seguridad de tus datos:
 - Todo se almacena y ejecuta en tu entorno.
 - No dependes de APIs externas para acceder a tu información.
- 3- Crear asistentes personalizados o chatbots:
 - Que respondan con base en documentos cargados (PDFs, Word, etc.).
 - Integrados con otros servicios (Telegram, WhatsApp, web, etc.) vía n8n.

Proceso

Vamos a realizar la instalación sobre el servidor de Linux, para hacer este proceso vamos a instalar una máquina virtual sobre virtual box, durante la instalación de Ubuntu server lts24 se activa la opción de Docker compose, luego se crea una carpeta en cualquier ubicación y allí se crea el archivo llamado "dockercompose.yml", dentro de este archivo se debe ingresar lo siguiente para la instalación de n8n con ollama.

```
versión: '3.8'
services:
 n8n:
 image: n8nio/n8n
 restart: always
 ports:
  - "5678:5678"
 volumes:
  - n8n_data:/home/node/.n8n
 environment:
  - N8N_BASIC_AUTH_ACTIVE=true
  - N8N_BASIC_AUTH_USER=#user
  - N8N_BASIC_AUTH_PASSWORD=#Password
 user: node
 networks:
  - n8n_network
 ollama:
 image: ollama/ollama
 container_name: ollama
 restart: unless-stopped
 ports:
```

```
- "11434:11434"
volumes:
```

- ./ollama data:/root/.ollama

networks:

- n8n network

volumes:

n8n data:

networks:

n8n network:

driver: bridge

Despues de ingresar este texto en el archivo se debe verificar que este bien indentado y guardar los cambios, para instalarlo se debe hacer con el siguiente comando "Docker-compose up -d –build ollama".

Al realizar este proceso me di cuenta de que cuando instale Ubuntu no le asigne suficiente memoria y no tenia espacio para la instalación de ollama, por este motivo resolví haciéndolo directamente con Open Ai, ya que no había necesidad de instalarlo, simplemente se creaban las credenciales con las claves correspondientes en n8n y listo.

En este proyecto vamos a implementar un agente de IA con RAG utilizando n8n y supabase para la memoria de las conversaciones y la base de datos vectorial para poder digerir documentos para permitir que el agente entregue respuestas claras y precisas con respecto al archivo entregado.

Vamos a crear una cuenta en supabase y luego creamos un nuevo proyecto. Luego de crear el proyecto vamos a buscar la opción de Project settings y allí ingresamos a API settings y en la parte de connection string vamos a encontrar con código PSQL donde vamos a encontrar todas las claves para crear la credencial de postgress, luego de conectar el postgress este nos va a crear una tabla llamada (n8n_chat_histories) que es donde va a almacenar el historial de conversaciones de el agente. Para crear la base de datos vectorial para los archivos de supabase se debe crear las credenciales que las vamos a encontrar en la misma parte supabase en API DATA, allí vamos a encontrar un código de SQL que crea una extensión para crear la base de datos vectorial y luego crea la tabla (documents) en donde vamos a poder digerir y guardar

los documentos, este código lo copiamos y lo pegamos en la parque de supabase donde dice SQL editor y allí corremos este código y ya automáticamente nos crea la tabla.

Lo que sigue seria crear el nodo de Google drive donde vamos a insertar el documento. Para conectar el drive con n8n debemos crear una cuenta de Google cloud, después de crear el proyecto vamos a ingresar a la opción de APIS y servicios habilitados y allí vamos a buscar credenciales y primero necesitamos crear una pantalla de consentimiento donde debemos ingresar correo electrónico, tipo de publico que va a ser externos, nombre de la app que es el nombre del proyecto y allí debemos aceptar los términos y condiciones. El siguiente paso seria cread cada Api a cada servicio de Google, en este caso vamos a utilizar el de Google drive, entonces lo buscamos y lo habilitamos, para crear las credenciales debemos sacar la url directamente desde n8n es la parte de crear credenciales, luego de pegar esta url le damos en crear y allí nos va a mostrar una pestaña con el id del cliente y el client secret que debemos ahora copiarlos y pegarlos en n8n y allí ya solo seria darle en conectar y guardar. De esta manera ya tendríamos el Google drive conectado en n8n y así lo podemos hacer para cualquier servicio de Google como Google docs, Google callendar, etc.

Ya que tenemos todas las conexiones necesarias lo único que queda es conectar todos los nodos y poner a prueba el flujo de trabajo.

Pasos Resumidos:

- 1- Instalar virtual box
- 2- Instalar Ubuntu server
- 3- Durante la instalación de Ubuntu server lts24 se activa la opción de Docker compose
- 4- Crear archivo llamado "Docker-compose.yml" en una carpeta nueva con nombre cualquiera en mi caso n8n_proy
- 5- Ingresar el archivo de texto y guardar los cambios
- 6- Se debe crear una cuenta en supabase
- 7- Crear un nuevo proyecto
- 8- Crear credenciales de postgress
- 9- Crear tabla para los documentos
- 10- Crear cuenta de Google cloud
- 11- Crear proyecto nuevo
- 12- Conectar Google drive con n8n

Comandos utilizados en Ubuntu server

Ifconfig: muestra la dirección ip

clear: Limpiar la pantalla

sudo su: Ingresa como super usuario

ls: Ver carpetas

cd..: Atrás

cd "Nombre de la carpeta": Entra a una carpeta

mkdir: Crear una carpeta

nano "Nombre de la carpeta": Crea un archivo txt

cat "Nombre de la carpeta": Muestra lo que hay en el archivo txt

docker ps – Muestra el estado de Docker

docker-compose.yml up -d: Levanta el contenedor

Palabras nuevas:

Api: Es una interfaz que permite que dos aplicaciones o sistemas se comuniquen entre sí.

Workflow: es una serie de pasos o tareas organizadas y automatizadas que se ejecutan de forma secuencial o condicional para lograr un objetivo.

Supabase: Es una plataforma de backend como servicio (BaaS) que te permite crear rápidamente aplicaciones modernas con una base de datos PostgreSQL, autenticación, almacenamiento y funciones en tiempo real, todo listo para usar.

Postgress: es un sistema de gestión de bases de datos relacional (RDBMS) de código abierto, potente y ampliamente utilizado. Está diseñado para ser seguro, escalable y extensible, ideal tanto para pequeños proyectos como para sistemas empresariales complejos.

Trigger: Un trigger (disparador) es un evento que inicia automáticamente una acción o un flujo de trabajo en sistemas como bases de datos, herramientas de automatización (como n8n) o aplicaciones.

Environments: es un conjunto de variables o configuraciones que definen cómo y dónde se ejecuta una aplicación. Se usan para adaptar el comportamiento de tu app según el lugar donde corre (desarrollo, pruebas, producción, etc.).

N8n: es una plataforma de automatización de flujos de trabajo (Workflow) que te permite conectar aplicaciones y servicios entre sí sin necesidad de escribir mucho código.