

FACULDADE SANTO AGOSTINHO  
Curso: Engenharia de Software  
Disciplina: Projeto de Banco de Dados  
Professor: Anderson Soares  
Integrantes: Nicolas Alexandrino e Ítalo Gabriel

## **RELATÓRIO TÉCNICO – SISTEMA DE CHATBOT COM BANCO DE DADOS VETORIAL**

### **1. Problema Escolhido**

O problema escolhido consiste em desenvolver um chatbot capaz de responder perguntas de forma inteligente a partir de documentos enviados pelo usuário. O objetivo é permitir que um arquivo enviado via Google Drive seja processado, convertido em embeddings e armazenado em um banco vetorial no Supabase. A IA, então, realiza buscas semânticas para responder perguntas com base no conteúdo armazenado.

### **2. Tecnologias Utilizadas**

- n8n – Automação e construção do workflow.
- Google Drive – Upload dos arquivos que serão indexados.
- Supabase + pgvector – Armazenamento vetorial dos embeddings.
- OpenAI – Geração dos embeddings e respostas via LLM.

### **3. O que São Embeddings**

Embeddings são representações numéricas de textos em forma de vetores de alta dimensionalidade. Diferente de palavras literais, embeddings capturam significado, contexto e relações semânticas. Isso permite que dois textos com significados semelhantes tenham vetores próximos no espaço vetorial.

### **4. Banco Vetorial Escolhido e Justificativa**

Foi utilizado o Supabase com a extensão pgvector. Motivos:

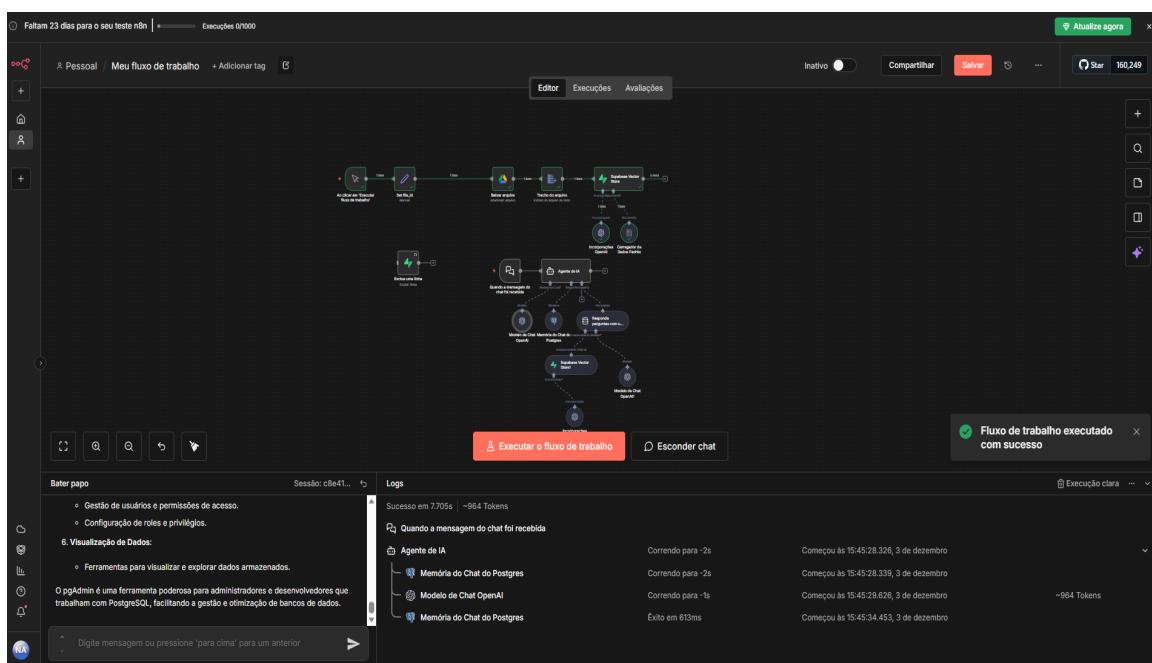
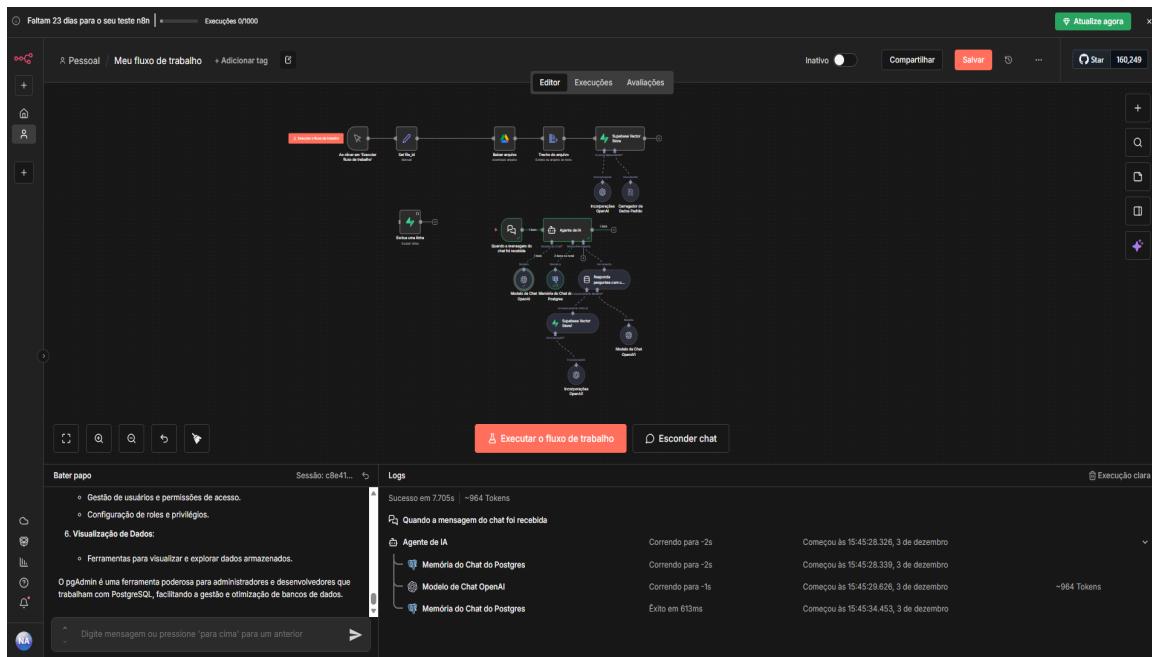
- É gratuito e fácil de configurar.
- Integra bem com aplicações modernas e com o n8n.
- Permite consultas por similaridade (cosine, Euclidean, inner product).
- Suporte nativo à extensão pgvector.

### **5. Arquitetura da Solução**

1. Usuário envia um arquivo para o Google Drive.
2. O n8n detecta o upload.
3. O arquivo é convertido para texto.
4. O texto é fragmentado e enviado para geração de embeddings.
5. Os embeddings são armazenados no Supabase.
6. O usuário envia uma pergunta ao chatbot.
7. A pergunta é convertida em embedding.
8. O Supabase retorna os vetores mais semelhantes.
9. A IA responde usando RAG (Retrieval Augmented Generation).

## 6. Prints do Workflow Funcionando

Abaixo estão capturas do fluxo funcionando dentro do n8n.



The screenshot shows a pgAdmin interface with the following details:

- Logs:** Successo em 7705s | ~984 Tokens
- Quando a mensagem do chat foi recebida:**
  - Agentes de IA: Correndo para -2s | Começou às 15:45:28.326, 3 de dezembro
  - Memória do Chat do Postgres: Correndo para -2s | Começou às 15:45:28.339, 3 de dezembro
  - Modelo de Chat OpenAI: Correndo para -1s | Começou às 15:45:29.626, 3 de dezembro | ~984 Tokens
  - Memória do Chat do Postgres: Êxito em 610ms | Começou às 15:45:34.453, 3 de dezembro
- Sidebar (Exemplos de bancos de dados vetoriais incluem pvector, Pinecone e Qdrant.)**
  - qualas a funcionalidades do pgadmin?**
  - O pgAdmin é uma ferramenta de administração e desenvolvimento para PostgreSQL. Suas principais funcionalidades incluem:
    - Gerenciamento de Banco de Dados:
      - Criação, alteração e exclusão de bancos de dados, tabelas, índices e outros objetos.
    - Consultas SQL:
      - Editor SQL com destaque de sintaxe e autocompletar.
      - Execução de scripts SQL e visualização de resultados.
    - Monitoramento:
      - Visualização de estatísticas de desempenho e uso.
      - Monitoramento de atividades do servidor.
    - Backup e Restauração:
      - Criação de backups e restauração de dados.
    - Segurança:
      - Gestão de usuários e permissões de acesso.
      - Configuração de roles e privilégios.
    - Visualização de Dados:
      - Ferramentas para visualizar e explorar dados armazenados.

## 7. Explicação da Busca Semântica

A busca semântica funciona comparando o embedding da pergunta com os embeddings armazenados no banco. Em vez de buscar palavras exatas, o sistema busca significados semelhantes. Isso torna o chatbot capaz de encontrar respostas relacionadas mesmo quando a pergunta é formulada de maneira diferente do texto original.

## 8. Resultados e Conclusões

O sistema desenvolvido demonstrou com sucesso a capacidade de indexar documentos, gerar embeddings e responder perguntas de forma contextualizada. O uso do n8n simplificou a criação do pipeline, enquanto Supabase e OpenAI garantiram robustez no armazenamento vetorial e qualidade nas respostas da IA. O chatbot é escalável e pode ser expandido para múltiplas aplicações, como educação, saúde, direito e atendimento ao cliente.