



Sumário

1. NOME DO GRUPO.....	2
2. INTEGRANTES DO GRUPO.....	2
3. DESCRIÇÃO DO TEMA ESCOLHIDO.....	2
4. PÚBLICO-ALVO.....	2
5. JUSTIFICATIVA DO TEMA ESCOLHIDO.....	2
6. DETALHAMENTO DO QUE FOI DESENVOLVIDO.....	3
7. ELEMENTOS ADICIONAIS.....	5
7.1. Fluxograma.....	5
7.2. Pitch Deck	5
8. LINK PARA O REPOSITÓRIO DO GITHUB QUE CONTÉM O PROJETO.....	6

	<p><i>PROJETO FINAL - FAQ NF-E</i></p>	
--	--	--

1. NOME DO GRUPO

LDM SOLUÇÕES

2. INTEGRANTES DO GRUPO

Lenine De Maria (lenine.de.maria@ldmsolucoes.com.br)

Leonardo De Maria (leonardodmail@gmail.com)

3. DESCRIÇÃO DO TEMA ESCOLHIDO

Agente inteligente que utiliza IA para auxiliar o usuário através de um chat que responde a dúvidas sobre a emissão, validação e fiscalização da Nota Fiscal Eletrônica.

4. PÚBLICO-ALVO

- Contribuintes;
- Contadores;
- Profissionais de tecnologia da informação;
- Demais usuários que precisam esclarecer dúvidas relativas a Nota Fiscal Eletrônica.

5. JUSTIFICATIVA DO TEMA ESCOLHIDO

O acesso a informações sobre NF-e é complexo devido à dificuldade em interpretar as normas fiscais brasileiras, e a escassez de recursos confiáveis e acessíveis para a consulta torna ainda mais complicado o entendimento das NF-e.

6. DETALHAMENTO DO QUE FOI DESENVOLVIDO

1. Usuário

Ponto de entrada do sistema, onde o usuário interage para tirar dúvidas e consultar erros da Nota Fiscal Eletrônica (NFe) do estado de São Paulo (SP).

2. FastAPI (NFE API)

Interface web principal do sistema. Recebe as requisições do usuário, repassa ao backend e retorna as respostas formatadas. Atua como camada de comunicação entre o front-end e a lógica de processamento.

3. Query Processor (Processador de Consultas)

Controla o fluxo completo de uma consulta. Encaminha a pergunta do usuário para o pipeline de busca local (SQL + ChromaDB) e decide quando acionar a LLM refinadora. Também gerencia o histórico e a formatação das respostas.

4. Base SQL (SQLite - Base Estruturada)

Banco de dados relacional que contém todos os erros oficiais de XML da NFe publicados pelo SEFAZ-SP. É a primeira fonte de busca, realizando consultas diretas por código ou descrição de erro.

5. Base Vetorial (ChromaDB - Base Semântica)

Camada vetorial construída a partir da base SQL. Armazena embeddings semânticos dos erros e descrições, permitindo buscas aproximadas e contextuais quando a consulta não tem correspondência exata na base SQL.

6. LLM Refinadora (Modelo de Linguagem Local/Externo)

Responsável por melhorar a qualidade da consulta. Corrige erros de digitação, normaliza sinônimos e reformula a pergunta com base nos resultados parciais da busca SQL+ChromaDB. Em seguida, gera o texto final da resposta sobre o erro.

7. Pipeline de Ingestão de Dados (Data Ingestion Pipeline)

Subsistema automatizado responsável por coletar, processar e popular as bases de conhecimento.

Inclui as seguintes etapas:

- **7.1. Fontes de Dados:** URLs oficiais e sites do SEFAZ contendo PDFs de erros e manuais técnicos.
- **7.2. Document Crawler:** Baixa automaticamente os PDFs com erros conhecidos.

- **7.3. Processador de Documentos:** Extrai códigos e descrições de erros a partir das seções especificadas em um dicionário de seções.
- **7.4. Data Converter:** Estrutura os dados extraídos e os insere no banco SQL, além de gerar embeddings vetoriais para o ChromaDB.
- **7.5. Data Ingestion (Orquestração LangGraph):** Controla toda a automação do pipeline, desde o download até a carga final nas bases.

8. Response Formatter (Formatador de Respostas)

Recebe as saídas das bases de conhecimento e da LLM, unifica e apresenta ao usuário respostas curtas, claras e numeradas, sempre com indicação da fonte.

9. LangGraph Orchestrator (Orquestrador de Agentes)

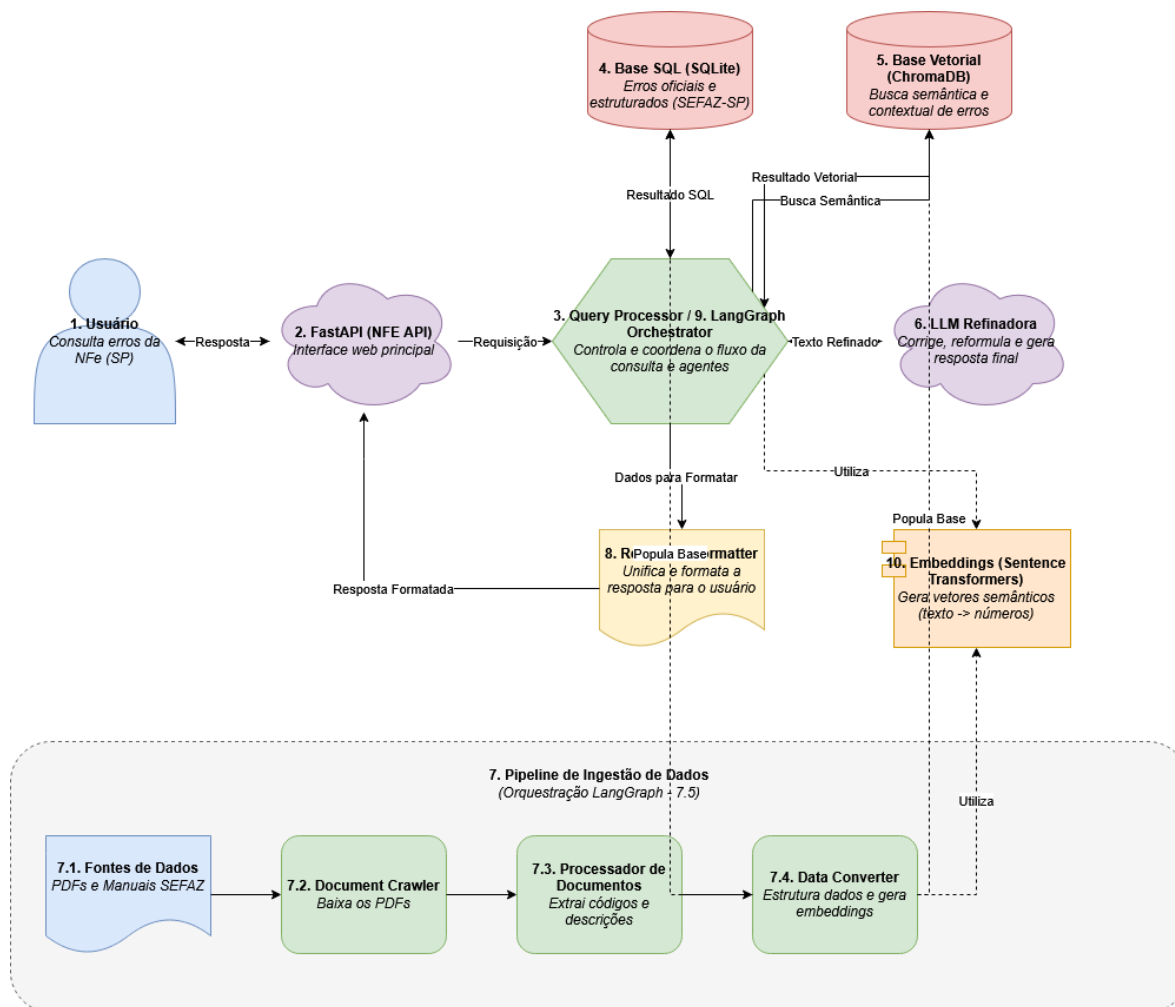
Coordena as etapas do sistema por meio de múltiplos agentes (Agent_nfe.py, Agent_pesquisa.py). Define o fluxo entre SQL, ChromaDB e LLM, controlando o estado e as transições do processo de consulta e ingestão.

10. Embeddings (Sentence Transformers)

Transforma os textos de erro e descrições em vetores numéricos, capturando o significado semântico e permitindo busca por similaridade no ChromaDB.



7. ELEMENTOS ADICIONAIS

7.1. Fluxograma



7.2. Pitch Deck

Encontra-se no repositório do projeto na pasta “Projeto Final - Artefatos”.

	<p><i>PROJETO FINAL - FAQ NF-E</i></p>	
--	--	--

8. LINK PARA O REPOSITÓRIO DO GITHUB QUE CONTÉM O PROJETO

https://github.com/ldmsolucoes/FAQ_NFE