

Relatório 2025 - Projeto Final I2A2

Nome do Grupo

Grupo YA

Integrantes do Grupo

Ana Júlia Oliveira Lins (analinsdev@gmail.com)

Yohanna de Oliveira Cavalcanti (yohannaoliveira12@gmail.com)

Introdução ao Tema

O tema escolhido foi “Extração de dados de notas fiscais físicas ou digitais em empresas de pequeno e médio porte”. O projeto consiste em um agente inteligente acompanhado de uma interface visual, desenvolvidos para facilitar o uso pelo usuário. No entanto, em cenários reais, o ideal seria que esse agente fosse integrado a uma API, permitindo sua utilização por sistemas internos da empresa.

Seguindo o contexto proposto na descrição do projeto, desenvolvemos uma solução capaz de ler e extrair automaticamente campos específicos de uma nota fiscal. Entre os dados capturados estão: informações do emitente e do destinatário, itens da nota (descrição, quantidade e valor), impostos (ICMS, IPI, PIS e COFINS), além de CFOP, CST e outros códigos fiscais.

Público alvo e justificativa do Tema Escolhido

O público-alvo do projeto será formado por pequenas empresas, com foco nos setores contábil e fiscal. Além dos dados básicos das NF-e, o sistema também realizará a extração de informações relevantes para essas áreas, como:

- Dados do emitente e do destinatário;
- Itens da nota (descrição, quantidade e valor);
- Impostos e demais códigos fiscais encontrados na nota fiscal.

Justificativa da Escolha do Tema

O projeto tem como objetivo automatizar a leitura de notas fiscais, tanto digitais quanto físicas, convertendo suas informações em formato JSON para fácil integração com outras ferramentas. Essa automação busca otimizar a

rotina dos usuários, reduzindo o tempo dedicado à extração manual de dados e minimizando erros humanos no processo.

No futuro, a aplicação poderá evoluir para uma API, ou ser conectada a uma, permitindo que outras ferramentas consumam diretamente os dados em JSON gerados. Dessa forma, será possível utilizá-la em aplicações web internas ou em soluções de análise de dados, viabilizando a geração de relatórios e gráficos com filtros por data de emissão, entre outros campos.

Análise das Etapas de Desenvolvimento

Para o desenvolvimento desse agente, usamos a linguagem python para a implementação da lógica do código, o Gemini AI como a nossa Inteligência Artificial e o Streamlit para ser o nosso frontend. Definimos modelos, helpers e funções de iniciação e comunicação com o Gemini para uma melhor extração de dados.

Como a solução será operadas, em passos:

1. (Usuário) Inserir a API KEY Gemini que será usada;
2. (Usuário) Escolher o download das notas fiscais (PDF, XML ou imagem (JPG/PNG);
3. (Agente) Irá ler as notas fiscais:
 - a. XML: o agente irá extrair o texto usando python;
 - b. Outros: o agente usará um prompt para extrair o texto da imagem ou PDF;
4. (Usuário) Clicar em "Processar Arquivos";
5. (Agente) Irá usar o prompt anterior para a extração dos dados relevantes a partir do texto gerado anteriormente e retornar um JSON para cada nota fiscal inserida;
6. (Plataforma) Irá mostrar os dados extraídos em formato de um array de JSONs;
7. (Usuário) Poderá visualizar o array e fazer o seu Download.

Links do projeto

Github (código):

https://github.com/AnaLinsDev/projeto_final_i1a2

Streamlit (deployed):

<https://projetofinali2a2.streamlit.app/>