A dark blue vertical bar on the left side of the page. A blue arrow points to the right from this bar, containing the date.

30/10/2025

Relatório do Projeto

Several thin, curved lines in dark blue and light grey that originate from the bottom left and sweep upwards and to the right.

Grupo: **MetaMind AI Agents**

Sumário

Grupo e participantes.....	2
Link de acesso ao Github.....	2
Descrição do Tema Escolhido	2
Público Alvo.....	2
Justificativa do Tema Escolhido	3
Detalhamento do Agente	3
Elementos Adicionais: tabelas, gráficos, diagramas	10
Anexos.....	11

Grupo e participantes

Nome do Grupo: **MetaMind AI Agents**

Participantes do Grupo:

Henry Monteiro – henry.monteiro@meta.com.br
Henrique Sardelli – henrique.sardelli@meta.com.br
Dario Volnei – dario.yamacake@meta.com.br
Douglas Kondo – douglas.kondo@meta.com.br
Felipe Santos – felipe.rsantos@meta.com.br
Lyndon Tavares - lyndon.tavares@meta.com.br
Elizabeth Marinho – elizabeth.marinho@meta.com.br
Bruno Kenzo – wakoteam@yahoo.com.br

Link de acesso ao Github

<https://github.com/dmkondo/MetaMindAIAgents/tree/main/Projeto%20Final%20-%20Artefatos>

Descrição do Tema Escolhido

O tema escolhido para o trabalho final foi **EXTRAÇÃO DE DADOS**. Esse tema indica algumas funcionalidades como:

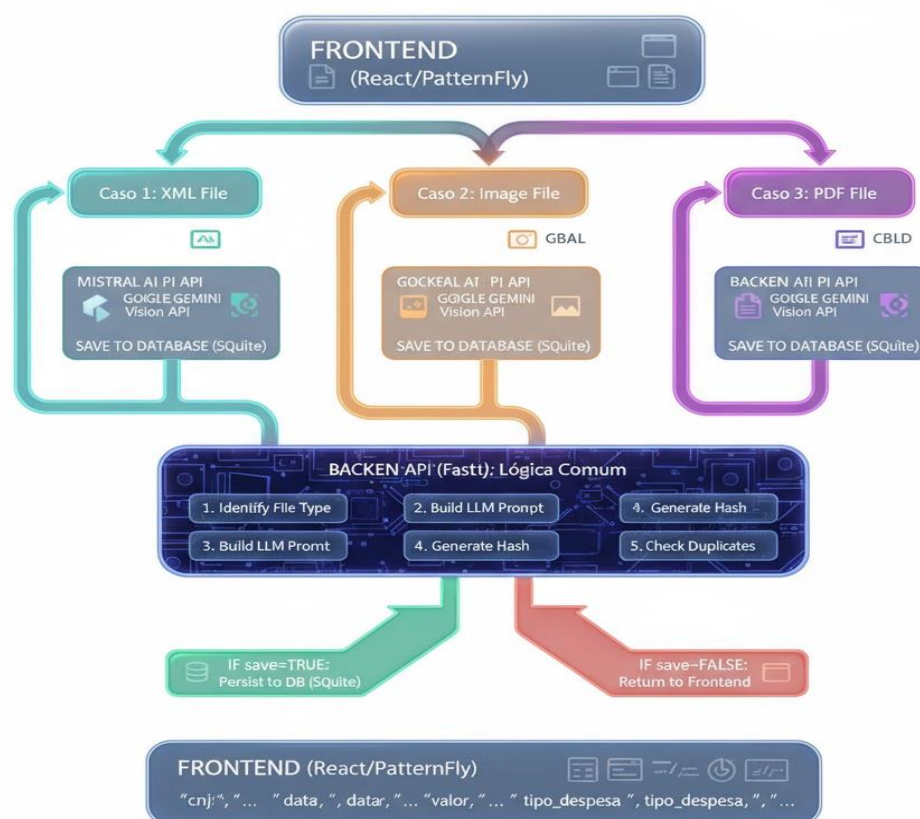
1. Recuperar documentos fiscais em fontes conhecidas;
2. Utilizar OCR (Reconhecimento Óptico de Caracteres) em conjunto com NLP (Processamento de Linguagem Natural) para extrair dados relevantes dos documentos:
 - a. Informações do emitente e destinatário;
 - b. Itens da nota (descrição, quantidade, valor);
 - c. Impostos (ICMS, IPI, PIS, COFINS);
 - d. CFOP, CST e outros códigos fiscais.
3. Desafio:
 - a. Como tornar o agente capaz de se adaptar a diferentes layouts e formatos de documentos;
 - b. Como se adaptar às mudanças legais (ex. IVA)

O grupo desenvolveu, para este trabalho de conclusão conforme o tema escolhido e descrito acima, um agente de IA que realiza a extração de dados a partir da leitura de notas fiscais preenchendo automaticamente um formulário de reembolso com identificação automática do valor, data e tipo de despesa.

Público Alvo

Colaboradores de empresas que necessitam realizar a prestação de contas para o registro de seus relatórios de reembolso.

INFOGRÁFICO: Fluxo Consolidado - Método extract_invoice_data



Justificativa do Tema Escolhido

Otimização do tempo e redução de erros em operações manuais tornando-as automatizadas através da leitura de documentos fiscais e inserção de dados com o uso de IA.

Detalhamento do Agente

Desenvolver uma API (Python) integrada com IA tendo um endpoint que receberá a imagem de uma nota fiscal e acionará um agente de IA para a extração e identificação dos dados persistindo-os em uma base de dados, juntamente com um segundo endpoint para consulta das mesmas.

Este documento técnico apresenta a arquitetura e os principais fluxos da API METAMIND, uma solução robusta para extração inteligente de dados de notas fiscais (imagens, PDFs e XMLs) utilizando modelos de linguagem (LLMs) avançados, como Google Gemini e Mistral AI. A API é a espinha dorsal de um sistema que visa automatizar e otimizar o processamento de documentos fiscais, fornecendo dados estruturados para análise e auditoria.

O frontend, desenvolvido em **React** e estilizado com **PatternFly**, oferece uma interface de usuário intuitiva para interagir com a API, visualizar dados extraídos e gerenciar configurações.

Arquitetura da Solução

A solução é composta por um conjunto de serviços interconectados, projetados para escalabilidade, desempenho e inteligência.

Componentes Principais:

- **Frontend (React + PatternFly):**
 - **Tecnologia:** React JS.
 - **Biblioteca UI:** PatternFly (design system baseado em React Components e CSS).
 - **Função:** Interface de usuário para upload de documentos, visualização de dados extraídos, edição e gerenciamento de configurações.
- **Backend (FastAPI):**
 - **Tecnologia:** Python com framework FastAPI.
 - **Função:** Gateway de API, orquestração de processos de extração, interação com LLMs, persistência de dados.
 - **Middleware:** CORS configurado para permitir comunicação com o frontend.
- **Modelos de Linguagem (LLMs):**
 - **Google Gemini (Flash/Pro Vision):** Utilizado para processamento multimodal (imagem e PDF) e extração de dados complexos de documentos.
 - **Mistral AI (Medium):** Utilizado para chat geral e extração de dados baseada em texto (via OCR ou XML).
- **Base de Dados (SQLite):**
 - **Tecnologia:** SQLite (para demonstração/desenvolvimento; pode ser migrado para um banco de dados mais robusto em produção).
 - **Função:** Armazenamento dos dados de notas fiscais extraídos e configurações da aplicação.
- **Processamento de Imagem/OCR:**
 - **Pillow (PIL):** Manipulação de imagens.
 - **Pytesseract:** Execução de OCR em imagens quando necessário (especialmente para `extract_invoice_data_with_mistral`).

Função `Extract_invoice_data`

A função assíncrona `extract_invoice_data` é a principal responsável por processar e extrair dados de notas fiscais digitais, utilizando modelos de Linguagem Grande Multimodais (LLMs) do Google Gemini para realizar OCR (Reconhecimento Óptico de Caracteres) e classificação.

1. Visão Geral e Parâmetros

Parâmetro	Tipo	Descrição
-----------	------	-----------

file	UploadFile	O arquivo da nota fiscal (Imagem, XML ou PDF) enviado pelo usuário.
save	bool	Flag booleana que determina a finalidade: True para salvar no banco de dados (status=PENDENTE), ou False para apenas checar/extrair dados iniciais (status=CHECKING) e verificar duplicidade.
session	Session	A sessão do banco de dados (SQLAlchemy) para consultas e persistência.

Exportar para as Planilhas

Objetivo: Extrair CNPJ, data de emissão, valor total e classificar o tipo_despesa da nota fiscal, verificar duplicidade e, opcionalmente, salvar no banco de dados.

2. Extração de Dados e Classificação (LLM - Gemini)

A função utiliza a biblioteca genai e o modelo **GEMINI_PRO_VISION_MODEL** para o processamento, dividindo a lógica em três casos principais baseados no content_type do arquivo:

✓ XML (text/xml, application/xml)

- Lê o conteúdo do XML como texto (xml_text).
- O LLM é chamado para analisar o **texto XML puro** e extrair os campos necessários, seguindo o prompt de JSON estrito.
- Usa o modelo de texto multimodal, que é eficiente para estruturar dados a partir de texto.

✓ Imagem (image/*)

- Lê o conteúdo binário da imagem (image_data).
- O LLM é chamado com o **conteúdo binário da imagem** (image_parts).
- O modelo utiliza suas capacidades de Visão para realizar **OCR** na imagem e, em seguida, extrair e classificar os dados.

✓ PDF (application/pdf)

- Lê o conteúdo binário do PDF (pdf_data).
- O LLM é chamado com o **conteúdo binário do PDF** (pdf_parts).
- O modelo Multimodal realiza a leitura e extração de dados diretamente do PDF, atuando como um **OCR avançado e extrator de estrutura**.

3. Processamento do Resultado e Hash

Para todos os formatos (XML, Imagem, PDF):

1. **Geração de Prompt:** O prompt é dinamicamente carregado da tabela de Configurations do banco de dados (se existir) ou usa um *fallback* padrão. O prompt exige que o LLM responda **SOMENTE em JSON estrito** com os campos: `cnpj`, `data` (DD/MM/AAAA), `valor`, `tipo_despesa` e `explicacao`.
2. **Decodificação de JSON:** A resposta crua do LLM (`raw_response`) é tratada para remover blocos de código (````json`) e, em seguida, decodificada usando `json.loads()`.
3. **Conversão de Valor:** O campo `valor` é convertido de forma segura para um tipo `float`.
4. **Geração de Hash:** Um `hash_value` (identificador único do arquivo) é gerado a partir do conteúdo binário do arquivo usando a função `gerar_hash_imagem`.

4. Persistência e Checagem de Duplicidade

Após a extração bem-sucedida, a função aplica a lógica de persistência:

1. **Checagem de Duplicidade:** O banco de dados é consultado usando o `imagem_hash` gerado para verificar se a nota fiscal já existe (existente).
2. **Lógica Condicional com save:**
 - **Se `save=True` (Tentativa de salvar):** Se o arquivo for um duplicado (existente), uma `HTTPException(400)` é levantada (erro, arquivo já cadastrado). Caso contrário, a nova fatura é salva com `status="PENDENTE"`.
 - **Se `save=False` (Modo de checagem/preview):**
 - Se o arquivo for um duplicado, o objeto existente é retornado imediatamente.
 - Se for um arquivo novo, um novo objeto `Invoice` é criado, mas **não é persistido**, recebendo o `status="CHECKING"` e é retornado para a pré-visualização no frontend.

Este design garante que o frontend receba dados pré-extraídos e validados, independentemente de ser um arquivo novo ou um duplicado, otimizando o fluxo de trabalho do usuário.

3. Endpoints da API METAMIND

Esta tabela detalha os endpoints disponíveis na API, agrupados por suas tags funcionais.

Tag	Endpoint	Método	Descrição
		HTTP	

Interação com LLM	/chat/mistral	POST	Recebe uma requisição de chat e a encaminha para a API do Mistral AI, retornando a resposta do modelo. Ideal para interações conversacionais ou geração de texto.
Interação com LLM	/invoices/extract/mistral	POST	Recebe uma imagem de nota fiscal, realiza OCR (via PyTesseract) e extrai CNPJ, data e valor total usando o modelo Mistral AI. Retorna os dados extraídos em formato JSON. (Não persiste no DB).
Interação com LLM	/invoices/extract/save	POST	Recebe um arquivo de nota fiscal (imagem, PDF ou XML), extrai CNPJ, data, valor total e tipo de despesa (ALIMENTACAO, VEICULO, ESCRITORIO) utilizando o Google Gemini Vision. Salva os dados extraídos no banco de dados.
Interação com LLM	/invoices/extract/check	POST	Recebe um arquivo de nota fiscal (imagem, PDF ou XML), extrai CNPJ, data, valor total e tipo de despesa (ALIMENTACAO, VEICULO, ESCRITORIO) utilizando o Google Gemini Vision. Não salva os dados no banco de dados, retornando-os para verificação.
Crud	/invoices	GET	Retorna uma lista de todas as notas fiscais (documentos) que foram extraídas e salvas no banco de dados.
Crud	/invoices/{id}	GET	Retorna os detalhes de um documento de nota fiscal específico, identificado pelo seu ID.
Crud	/invoices/add	POST	Adiciona um novo documento de nota fiscal manualmente ao banco de dados, utilizando os dados fornecidos no corpo da requisição.

Crud	/invoices/{id}	PUT	Atualiza parcialmente os dados de um documento de nota fiscal existente, identificado pelo seu ID. Permite modificar campos como CNPJ, tipo de despesa, data de emissão, valor total e status.
Crud	/invoices/{id}	DELETE	Exclui permanentemente um documento de nota fiscal do banco de dados, identificado pelo seu ID.
Configuração	/configuration	PUT	Atualiza o prompt de extração de dados padrão usado pelos modelos de LLM. Isso permite personalizar a forma como os dados são extraídos sem alterar o código da API.
Configuração	/configuration	GET	Retorna o prompt de extração de dados atualmente configurado na aplicação.

4. Front-end

Exemplo de front-end em React usando design system Patterfly.

4.1 Adicionar Nota Fiscal

Ao fazer upload da nota fiscal, é chamado endpoint de extração e checagem. O endpoint verifica o tipo de arquivo: pdf, xml ou imagem. Ao extrair os dados (data emissão, valor total, cnpj), também faz a classificação da categoria de despesa.

Exatracão Inteligente

Adicionar Nota Fiscal

X

Faça upload de notas fiscais para extração de dados.

ID

Tipo de Despesa *

CNPJ *

Data de Emissão *

DD/MM/AAAA

Valor Pago *

Imagem Hash

Status

CHECKING

Salvar

Cancelar

Upload Imagem

talo-de-conta-de-luz-na-foto-talo-de-conta-de-luz-da-copel-P42A50.jpg

Selecione...

Limpar Formulário

Processando extração de dados...

(Clique para ampliar)

Girar esquerda

Girar direita

Adicionar Nota Fiscal

Faça upload de notas fiscais para extração de dados.

ID

Tipo de Despesa *

Energia

CNPJ *

04368898000106

Data de Emissão *

16/02/2018

Valor Pago *

175,57

Imagem Hash

dffc9479025701ac6514796319b6997f1

Status

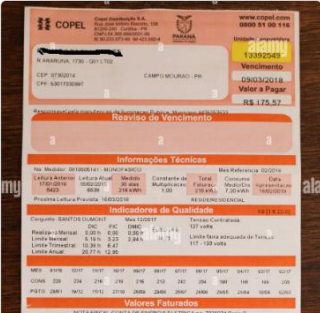
CHECKING

Upload Imagem

talo-de-conta-de-luz-na-foto-talo-de-conta-de-luz-da-copel-P42A50.jpg

Selecione...

Limpar Formulário



Elementos Adicionais: tabelas, gráficos, diagramas

Banco de dados: *SQLite*.

Utilizar um framework – *FastAPI*, para estabelecer um padrão para o *Swagger*, assim conseguimos ter uma *API*. A integração com o *agent*, será feita via *langchain*, quanto ao modelo estamos avaliando a melhor solução entre os modelos *Mitre* ou *Gemini* ou *ChatGP*. Para o OCR estamos estudando também usar a biblioteca *pytesseract*, que é uma biblioteca *free* que viabiliza diminuir o custo para o desenvolvimento do nosso *agent*.

Projeto de Conclusão do Curso

O projeto final deve ser entregue por cada um dos grupos e deverá apresentar a síntese do aprendizado obtido pelos integrantes.

Ele deve estar em consonância com temas apresentados na aula de 15/05/2025, ou seja, alinhado a pelo menos um destes temas.

Vocês devem se lembrar que já fizeram por duas vezes a apresentação da evolução dos seus projetos: em 12/06/2025, vocês enviaram a primeira versão do projeto e em 24/07/2025, enviaram a segunda versão, com as atualizações em função do que aprenderam até aquele momento.

Para tornar mais fácil a verificação dos temas, após o detalhamento da entrega, reproduzimos os slides apresentados em 15/05.

A Entrega

Agora, vocês devem entregar o que foi conseguido pelo grupo durante o curso.

O que deve ser produzido:

- Relatório do Projeto
- Apresentação em formato PPTX
- Vídeo de apresentação para banca examinadora
- Estes artefatos (relatório, apresentação e vídeo) deverão estar armazenados em uma pasta no repositório do projeto no Github. A pasta deve ter o nome "Projeto Final - Artefatos".
- O repositório do Github deve permitir acesso público e em seu readme.md deve se destacar que o projeto se encontra sob a licença MIT.

A Entrega

O Relatório do Projeto

- Cada grupo deverá gerar um relatório em formato PDF
- O arquivo deverá ter o nome:
I2A2_Agentes_Inteligentes_Projeto_Final_<Nome_do_Grupo>.pdf
- No relatório deverá ser apresentado:
 - Nome do Grupo
 - Integrantes do Grupo
 - Descrição do Tema Escolhido (o que foi feito)
 - Público alvo (quem irá utilizar a solução)
 - Justificativa do Tema Escolhido (por que isto é importante, o que agrega de valor ao público alvo)
 - Detalhamento do que foi desenvolvido (quais as funções, como a solução é operada, etc.)
 - Elementos adicionais: tabelas, gráficos, diagramas
 - Link para o repositório do github que contém o projeto

A Entrega

A Apresentação e o Vídeo

- O grupo também deverá gerar uma apresentação de slides em formato de *pitch deck* para apresentação para a banca examinadora.
- O arquivo da apresentação deve estar em formato PPTX e deverá ter o nome: I2A2_Agentes_Inteligentes_Projeto_Final_<Nome_do_Grupo>.pptx.
- O grupo também deve gravar um vídeo de apresentação de seu pitch com duração máxima de 5 minutos.
- O vídeo deverá ter o seguinte nome: I2A2_Agentes_Inteligentes_Projeto_Final_<Nome_do_Grupo>.mp4.
- Detalhes sobre como construir pitches em <https://sebrae.com.br/sites/PortalSebrae/ufs/pe/artigos/pitch-deck-o-que-e-e-como-estruturar-o-seu,4787c413c7fd4710VgnVCM1000004c00210aRCRD>

A Entrega

O Envio

- O envio deve ser realizada por e-mail.
 - O responsável pelo grupo deverá enviar uma mensagem para o endereço challenges@i2a2.academy, com cópia para si mesmo.
 - O título do e-mail deverá ser "I2A2 - Agentes Autônomos - Projeto Final - <Nome do Grupo>".
 - No corpo do e-mail deve ser enviado apenas o endereço do repositório do projeto no Github, onde deverão estar armazenados todos os artefatos, inclusive o relatório em formato PDF, apresentação em formato PPTX e vídeo em formato mp4.
 - Opcionalmente, o responsável pelo grupo poderá copiar os demais membros do grupo e desta forma gerar um "protocolo de entrega" adicional.
- O limite para entrega é 30/10/2025 às 23h59.
- Esta atividade é obrigatória para a finalização do curso. A ausência da entrega descredencia todos os elementos do GRUPO para receber o certificado de conclusão do curso.
- Lembrando novamente que o repositório do Github deve permitir acesso público e no readme.md deve estar destacado que o seu conteúdo se encontra sob a licença MIT.

Temas para os Projetos

Slides apresentados em 15/05/2025



15/05/2025 – Em que vamos trabalhar

• Objetivo:

- Automatizar o processamento e análise de documentos fiscais
- Podem ser documentos físicos ou eletrônicos (ex.: XML de NFe/NFCe/CTe/MDF-e)
- Foco em otimização/aprimoramento:
 - Redução de erros manuais na escrituração;
 - Otimização de tempo no fechamento contábil e fiscal;
 - Detecção de inconsistências fiscais (valores, CFOP, CST, NCM etc.);
 - Integração com ERPs de mercado e sistemas contábeis (Domínio, Alterdata, Protheus, etc.).

15/05/2025 – Temas que podem ser tratados

• Atividades alvo (temas):

- Extração de dados
- Validação e Auditoria
- Classificação, Categorização e Customização por ramo de atividade
- Automação de Processos Fiscais/Contábeis
- Ferramentas Gerenciais

15/05/2025 – Temas que podem ser tratados

• Extração de Dados:

- Recuperar documentos fiscais em fontes conhecidas
- Utilizar OCR (Reconhecimento Óptico de Caracteres) em conjunto com NLP (Processamento de Linguagem Natural) para extrair dados relevantes dos documentos:
 - Informações do emitente e destinatário
 - Itens da nota (descrição, quantidade, valor)
 - Impostos (ICMS, IPI, PIS, COFINS)
 - CFOP, CST e outros códigos fiscais
- Desafios:
 - Como tornar o agente capaz de se adaptar a diferentes layouts e formatos de documentos.
 - Como se adaptar às mudanças legais (ex. IVA)

15/05/2025 – Temas que podem ser tratados

• Validação e Auditoria:

- Agentes que verificam a consistência dos dados, comparando-os com regras fiscais e cadastros de clientes/fornecedores.
- Identificar e sugerir correção para erros comuns, como:
 - Cálculo incorreto de impostos
 - Códigos fiscais inconsistentes
 - Divergências entre pedido de compra e nota fiscal
- Produzir relatórios de auditoria, destacando possíveis problemas e áreas de risco e enviando-os aos responsáveis (que podem ser outros agentes)
- Desafios:
 - Identificar maiores agressores e sugerir melhorias
 - Adaptar às mudanças legais ou do ambiente de negócios

15/05/2025 – Temas que podem ser tratados

• Classificação e Categorização e Customização por ramo de atividade:

- Classificar automaticamente os documentos fiscais por tipo (compra, venda, serviço) e por centros de custos.
- Organizar e o arquivar corretamente os documentos.
- Realizar ações customização por ramo de atividade. Ex.:
 - Agronegócio: Monitoramento de CFOPs específicos do setor (venda de produtos agrícolas, insumos), cálculo de impostos com particularidades do agronegócio.
 - Setor Automotivo: Validação de notas fiscais de peças e serviços automotivos, conferência de códigos de peças e compatibilidade com as atividades da empresa.
 - Indústria: Apuração de impostos específicos da indústria (IPI, Substituição Tributária, etc.), Geração de insumos para cálculo de custos de produção
- Desafios:
 - Adaptar às mudanças legais ou do ambiente de negócios
 - Como tratar ramos de atividade específicos – órgãos públicos, terceiro setor, etc.

15/05/2025 – Temas que podem ser tratados

• Automação de Processos Fiscais/Contábeis:

- Lançamentos Contábeis:
 - Os agentes geram automaticamente os lançamentos contábeis a partir dos dados obtidos nos documentos fiscais.
- Apuração de Impostos:
 - Os agentes calculam os impostos a pagar e a recuperar, gerando guias de recolhimento.
 - Automação a entrega de obrigações acessórias (SPED Fiscal, EFD Contribuições).
- Conciliação Bancária:
 - Cruzamento dos dados das notas fiscais com os extratos bancários, facilitando a conciliação.
 - Identificação pagamentos e recebimentos pendentes.
- Desafios:
 - Manter os agentes atualizados em relação a critérios contábeis e obrigações acessórias?
 - Como se beneficiar do Open Banking?
 - Como garantir a segurança dos processos?

15/05/2025 – Temas que podem ser tratados

- Ferramentas Gerenciais:

- Relatórios Personalizados:

- Geração de relatórios personalizados, com informações relevantes para o seu setor.
 - Utilizar informações internas
 - Agregar informações externas relevantes para a empresa

- Análises preditivas e simulações de cenários.

- Assistente Consultor Especializado:

- Suporte para dúvidas e decisões estratégicas.
 - Informações sobre contabilidade e tributação.

- Desafios:

- Como garantir a qualidade das informações apresentadas?
 - Como maximizar a experiência do usuário