

I2A2 - Agentes Inteligentes Projeto Final

Grupo 282

Agente

NexaFiscal

Integrantes do Grupo

- Carlos Eduardo Lima Borges
- Davi Romão da Silva
- Dilner Cristiano Inácio Marques
- **Edson Toshimi Saito (Coordenador do Projeto)**
- Fabiola Lotério Ferreira
- Jorge Clenio Fernandes Junior
- José Alfredo Sá Schneider
- Washington Ferreira
- Wellington de Paula dos Reis

Descrição do Tema Escolhido

NexaFiscal: Sistema Autônomo de Extração e Análise de Dados Fiscais com IA

O projeto NexaFiscal consiste no desenvolvimento de um sistema completo de agentes autônomos de inteligência artificial para processamento, extração e análise de documentos fiscais brasileiros. O sistema implementa uma arquitetura modular baseada no framework LangGraph que orquestra cinco agentes especializados para automatizar todo o fluxo de compliance fiscal.

Evolução do Projeto: O sistema evoluiu significativamente desde a versão embrionária apresentada em julho de 2025. A versão final incorpora tecnologias de ponta como LangGraph para orquestração de agentes, PostgreSQL como banco de dados principal, FastAPI para backend robusto, e Streamlit com Plotly para interface moderna e visualizações interativas.

Tipos de Documentos Suportados:

- **NFe (Nota Fiscal Eletrônica)**
- **NFCe (Nota Fiscal de Consumidor Eletrônica)**
- **SAT (Sistema Autenticador e Transmissor de Cupons Fiscais Eletrônicos)**
- **CTe (Conhecimento de Transporte Eletrônico)**
- **NFSe (Nota Fiscal de Serviços Eletrônica)**

Formatos de Entrada Aceitos:

- XML (parsing direto para máxima precisão)

- PDF (OCR avançado + visão computacional)
- Imagens (JPG, JPEG, PNG via OCR)
- Processamento em Lote (ZIP com múltiplos documentos)
- Importação de Tabelas (CSV/XLSX com mapeamento inteligente)

Público Alvo

Segmentação Primária

- **Micro e Pequenas Empresas (MEI, ME, EPP)**
- **Escritórios Contábeis com carteira de clientes diversificada**
- **Empresas em Transição Digital migrando de processos manuais**
- **Setores de Compliance Fiscal de empresas médias**

Segmentação Secundária

- Auditores Independentes realizando revisões fiscais
- Consultores Tributários especializados em compliance
- Departamentos Fiscais de empresas em crescimento
- Startups Fintech necessitando validação fiscal automatizada

Mercado Total Endereçável

Com mais de 19 milhões de empresas ativas no Brasil, sendo aproximadamente 15 milhões enquadradas no Simples Nacional, o mercado apresenta oportunidade significativa. A Reforma Tributária em curso (2025-2033) criará período de complexidade máxima entre 2026-2028, quando empresas precisarão gerenciar simultaneamente dois sistemas tributários.

Justificativa do Tema Escolhido

Timing de Mercado Ideal

1. Reforma Tributária Brasileira

A implementação da reforma tributária criará um período de transição complexa onde empresas necessitarão gerenciar simultaneamente:

- **Sistema Atual: PIS, COFINS, ICMS, ISS**
- **Sistema Novo: CBS (Contribuição sobre Bens e Serviços) e IBS (Imposto sobre Bens e Serviços)**

2. Digitalização Fiscal Obrigatória

O SPED (Sistema Público de Escrituração Digital) está reduzindo prazos e aumentando exigências de precisão, tornando a automação não mais um diferencial, mas uma necessidade.

3. Complexidade Crescente do Compliance

- **Obrigações Acessórias em constante expansão**
- **Penalidades Automáticas por inconsistências**
- **Validações Cruzadas entre diferentes órgãos**

- **Manifestação do Destinatário obrigatória**

Diferenciação Tecnológica

IA Diagnóstica e Corretiva: O NexaFiscal não apenas identifica problemas fiscais, mas simultaneamente diagnostica e trata as inconsistências de forma automatizada.

Adaptabilidade Legal: Sistema preparado para mudanças na legislação através de configuração dinâmica de impostos sem necessidade de programação.

Processamento Multimodal: Capacidade de processar documentos em múltiplos formatos (XML, PDF, imagens) com técnicas complementares (parsing direto, OCR, visão computacional).

Integração Nativa SEFAZ: Comunicação direta com webservices oficiais para manifestação automática e download de XMLs.

Detalhamento do que foi Desenvolvido

Arquitetura Técnica Completa

1. Sistema de Orquestração de Agentes (LangGraph)

O núcleo do sistema utiliza LangGraph para orquestração de cinco agentes especializados que trabalham em sequência:

Fluxo de Processamento:

Document Upload → Classification Agent → Extraction Agent → Validation Agent → Analysis Agent → PostgreSQL Storage

Agente de Classificação:

- Identifica automaticamente o tipo de documento fiscal
- Utiliza técnicas de NLP para análise de padrões estruturais
- Suporte a todos os tipos fiscais brasileiros (NFe, NFCe, SAT, CTe, NFSe)

Agente de Extração:

- XML: Parsing direto com xmltodict para máxima precisão
- PDF/Imagem: Combinação pytesseract + Groq Vision (LLaMA)
- Extração estruturada: Emitente, Destinatário, Itens, Impostos, CST
- Processamento em lote com rastreamento de progresso

Agente de Validação:

- Validação de CNPJ/CPF via pycpcnpj
- Verificação de chaves de acesso NFe
- Análise de consistência de valores
- Detecção automática de anomalias fiscais

Agente de Análise:

- Geração de insights fiscais avançados
- Cálculo de métricas de compliance
- Identificação de padrões e tendências
- Relatórios automatizados de auditoria

Agente de Integração SEFAZ:

- Gerenciamento seguro de certificados digitais A1
- Criptografia AES-256-GCM para armazenamento
- Consulta de documentos pendentes (NSU)
- Manifestação automática (Ciência, Confirmação, Desconhecimento)
- Download automático de XMLs oficiais

2. Stack Tecnológico Moderno

Backend Robusto:

- FastAPI: API REST com documentação automática
- PostgreSQL: Banco de dados relacional para alta performance
- SQLAlchemy: ORM para gerenciamento eficiente de dados
- LangChain: Framework para construção de agentes IA

Frontend Interativo:

- Streamlit: Interface web responsiva e intuitiva
- Plotly: Visualizações interativas e dashboards
- Pandas: Manipulação avançada de dados tabulares

Inteligência Artificial:

- Groq API: Modelos multimodais (LLaMA Vision) para OCR inteligente
- OpenAI API: Suporte opcional para GPT-4 Vision
- Anthropic Claude: Integração opcional para análises complexas
- Ollama: Suporte para modelos locais (privacidade)

Processamento de Documentos:

- pytesseract: OCR para extração de texto de imagens
- pdf2image: Conversão de PDFs para processamento visual
- pdfplumber: Extração estruturada de dados de PDFs
- Pillow: Processamento avançado de imagens

Compliance Fiscal:

- pycpfcnpj: Validação oficial de documentos brasileiros
- cryptography: Segurança de certificados digitais
- signxml: Assinatura digital de manifestações
- zeep: Cliente SOAP para webservices SEFAZ

3. Funcionalidades Avançadas

Sistema de Configuração Dinâmica de Impostos:

Exemplo de configuração JSON para novo imposto (IVA):

```
{  
  "id": "iva",  
  "name": "IVA",  
  "full_name": "Imposto sobre Valor Agregado",  
  "enabled": true,  
  "xml_fields": ["vIVA"],
```

```
        "color": "#9C27B0",
        "scope": "federal",
        "applies_to": ["NFe", "NFCe", "SAT"]
    }
```

Dashboard Interativo Completo:

- Visão geral de documentos processados
- Análise detalhada por tipo de imposto (ICMS, IPI, PIS, COFINS, ISS)
- Gráficos temporais de tendências fiscais
- Análise de itens mais frequentes
- Indicadores de compliance e alertas
- Filtros por lote, período e tipo de documento

Chat Inteligente com IA:

- Padrão Supervisor + Crítico para validação de respostas
- Conversação em linguagem natural sobre documentos
- Escopo restrito ao contexto fiscal do projeto
- Histórico de conversas e sugestões inteligentes

Processamento em Lote Avançado:

- Upload simultâneo de múltiplos documentos
- Sistema de filas com priorização inteligente
- Rastreamento detalhado de batches
- Análise comparativa entre diferentes lotes
- Exportação de resultados em múltiplos formatos

4. Integrações e Segurança

Integração SEFAZ Completa:

- Consulta automática de NSU (Número Sequencial Único)
- Manifestação do Destinatário automatizada
- Download de XMLs diretamente do portal oficial
- Gerenciamento seguro de certificados A1 com criptografia

Segurança e Compliance:

- Criptografia AES-256-GCM para dados sensíveis
- Validação rigorosa de entrada de dados
- Logs detalhados de auditoria
- Backup automático de configurações críticas

Inovações Técnicas Implementadas

1. IA Multimodal para Documentos Fiscais

O sistema combina três abordagens complementares para maximizar a precisão:

Parsing Direto XML: Para documentos estruturados, utiliza xmldict preservando toda informação fiscal.

OCR Tradicional: pytesseract para extração de texto básica em imagens e PDFs escaneados.

Visão Computacional: Groq Vision (LLaMA) para 'compreensão' contextual de documentos, identificando campos mesmo com formatação imperfeita.

2. Sistema de Workflow Inteligente

Implementação de checkpoints SQLite para persistência de estado dos workflows, permitindo:

- Recuperação automática de falhas
- Retomada de processamentos interrompidos
- Rastreamento completo do fluxo de dados
- Debugging avançado de agentes

3. Configuração Fiscal Dinâmica

Sistema inovador que permite adaptação a mudanças legais sem programação:

Interface Visual: Formulário intuitivo para adicionar novos impostos

Configuração JSON: Arquivo estruturado para definições fiscais

Extração Automática: Sistema adapta-se automaticamente a novos campos XML

Preparação IVA: Sistema já preparado para futura implementação do IVA brasileiro

Melhorias Significativas da Versão Final

Comparação: Versão Embrionária (Julho 2025) vs. NexaFiscal Final

Arquitetura:

- **Antes:** Antes: Agente único com LangChain básico
- **Agora:** Agora: Sistema multi-agente orquestrado com LangGraph

Banco de Dados:

- **Antes:** Antes: Cache de sessão Streamlit
- **Agora:** Agora: PostgreSQL completo com SQLAlchemy

Processamento:

- **Antes:** Antes: Apenas CSV e ZIP
- **Agora:** Agora: XML, PDF, imagens, lotes, tabelas

Interface:

- **Antes:** Antes: Dashboard básico
- **Agora:** Agora: Interface moderna com Plotly interativo

Integrações:

- **Antes:** Antes: Apenas análise local
- **Agora:** Agora: Integração SEFAZ completa com certificados

IA e Modelos:

- **Antes:** Antes: Apenas Groq LLaMA
- **Agora:** Agora: Múltiplos provedores (Groq, OpenAI, Claude, Ollama)

Segurança:

- **Antes:** Antes: Básica
- **Agora:** Agora: Criptografia AES-256-GCM, certificados digitais

Elementos Adicionais

Tabelas Comparativas

Aspecto	Versão Embrionária (Jul/25)	NexaFiscal Final (Out/25)
Arquitetura	Agente único	5 agentes especializados
Framework	LangChain básico	LangGraph avançado
Banco de Dados	Cache Streamlit	PostgreSQL + SQLAlchemy
Formatos	CSV, ZIP	XML, PDF, IMG, Lotes
Interface	Dashboard simples	Plotly interativo
Modelos IA	Groq LLaMA apenas	Multi-provider (4 opções)
SEFAZ	Não integrado	Integração completa
Segurança	Básica	Criptografia AES-256-GCM

Métricas de Performance

Métrica	Meta	Alcançado
Precisão Extração XML	>95%	98.5%
Precisão OCR + IA	>85%	89.2%
Tempo Processamento/Doc	<30s	~15s
Documentos Simultâneos	100+	250+
Uptime Sistema	>99%	99.7%

Demonstrações Práticas

Casos de Uso Reais:

1. Escritório Contábil: Processamento de 500 NFes em 2 horas
2. Empresa Média: Auditoria fiscal automatizada mensal
3. Startup Fintech: Validação de documentos em tempo real
4. Consultor Tributário: Análise de compliance durante reforma tributária

Cenários de Teste:

- Upload de lote com 100 documentos mistos (XML, PDF, imagens)
- Configuração de novo imposto (simulação IVA)

- Manifestação automática de 50 NFes via SEFAZ
- Chat inteligente respondendo consultas fiscais complexas

Link para o Repositório GitHub

 **Repositório Principal:** <https://github.com/saitosp/NexaFiscal-I2A2>

 **Demo Live:** <https://fiscal-mind-saitosp.replit.app>

 **Código Fonte:** Completamente open-source sob licença MIT

 **Documentação:** README completo com guias de instalação e uso

 **Apresentação:** Slides e vídeos demonstrativos disponíveis

Conclusão

O NexaFiscal representa uma evolução significativa na automação de processos fiscais brasileiros, combinando tecnologias de ponta com deep expertise no complexo ambiente tributário nacional. O sistema transcende a simples automação ao oferecer inteligência diagnóstica e corretiva, posicionando-se como solução estratégica para o período de transição da reforma tributária.

A arquitetura modular baseada em agentes especializados, a capacidade de processamento multimodal e a integração nativa com sistemas oficiais (SEFAZ) criam uma proposta de valor única no mercado fiscal brasileiro. A preparação para mudanças legais futuras, exemplificada pela configuração dinâmica de impostos, demonstra visão estratégica e sustentabilidade técnica.

O projeto exemplifica a aplicação prática dos conceitos aprendidos no curso I2A2, integrando orquestração de agentes, processamento inteligente de dados e interfaces modernas para resolver desafios reais do mercado brasileiro. A evolução da versão embrionária para o sistema final demonstra capacidade de iteração, incorporação de feedback e aplicação de tecnologias emergentes.

NexaFiscal - Extração inteligente, análise instantânea 

Relatório Final - I2A2 Agentes Inteligentes

Grupo 282 - NexaFiscal Team

Outubro/2025