Etapa 1: Planejamento e Requisitos

Definição de requisitos funcionais e não funcionais.

Escolha de tecnologias (OCR, LLM, frameworks de backend e frontend).

Levantamento de fontes de dados (tipos e formatos de documentos fiscais).

Etapa 2: Construção do Pipeline Automatizado

Agente 1 – Aquisição de Documentos:

Módulo de upload manual e download automático de fontes confiáveis (ex: SEFAZ).

Armazenamento em banco de dados estruturado.

Agente 2 – Extração e Aprendizado:

Aplicação de OCR (ex: Tesseract, AWS Textract, Google Vision).

Processamento com NLP para estruturação de dados.

Uso de LLMs para adaptação a novos layouts de notas fiscais.

Base de Conhecimento:

Criação e manutenção de banco de dados estruturado com os dados extraídos.

Indexação para rápido acesso pelas interfaces e LLM.

Agente 3 – Resposta e Interação:

Integração com LLM (ex: OpenAI GPT, Claude, Mistral) para resposta contextualizada a perguntas dos usuários.

Consultas baseadas nos dados da Base de Conhecimento.

Etapa 3: Desenvolvimento da Interface do Usuário

Interface Web responsiva (React, Vue ou Angular).

Funcionalidade de upload, visualização de documentos, e painel de perguntas e respostas.

Etapa 4: Testes

Testes unitários e de integração nos módulos de aquisição, extração e interação.

Validação da acurácia da extração de dados.

Testes de usabilidade da interface.

Etapa 5: Implantação e Monitoramento

Deploy em ambiente de produção (AWS, Azure ou Google Cloud).

Monitoramento de performance e logging de falhas.

Treinamento para usuários finais.

---

5. Tecnologias Sugeridas

OCR: Tesseract, Google Vision API, Amazon Textract.

NLP e LLM: spaCy, Transformers (HuggingFace), OpenAI GPT API.

Backend: Python (FastAPI, Flask) ou Node.js.

Frontend: React.js.

Banco de dados: PostgreSQL ou MongoDB.

Infraestrutura: Docker, Kubernetes, CI/CD (GitHub Actions, GitLab CI).

---

6. Cronograma Estimado (Macro)

Fase Duração Estimada

Planejamento e Requisitos 1 semana

Desenvolvimento dos Agentes 4 semanas

Interface e Integrações 2 semanas

Testes 1 semana

Implantação 1 semana

Total 9 semanas

---

7. Próximos Passos

1. Validação desta proposta com as partes interessadas.

2. Refinamento dos requisitos com base em casos de uso reais.

3. Alinhamento com equipe técnica para início do planejamento detalhado.