

Plano Parcial do Projeto Final de Inteligência de Negócios

DISCIPLINA: Inteligência de Negócios

CURSO: Ciências da Computação e Engenharia de Software

TURNO: Matutino

DOCENTE RESPONSÁVEL: Lucas dos Santos Althoff

Integrantes e Responsabilidades:

- Mateus De Araújo Almeida
- Aluno2
- Aluno3

Introdução

O presente relatório parcial do projeto de Inteligência de Negócios foi desenvolvido para criar uma solução baseada em coleta e análise de dados de processos judiciais públicos. Esse projeto visa solucionar problemas de tomada de decisão para otimização de processos no setor jurídico.

Objetivos

1. **Objetivo Geral:** Facilitar o acesso e a análise de dados de processos judiciais para apoiar decisões estratégicas no setor jurídico, através de um sistema de coleta e mineração de dados.
2. **Objetivos Específicos:**
 - Criar uma estrutura automatizada de coleta de dados públicos com Python e web scraping.
 - Armazenar dados em um banco NoSQL (MongoDB ou Cassandra) e em um banco de grafos (Neo4j) para consultas relacionais.
 - Minerar informações relevantes, como leis aplicadas e prazos, para identificação de padrões.
 - Desenvolver um dashboard interativo com Streamlit para visualização de dados em tempo real, usando Redis como cache em memória.

Setor de Atuação e Problemas de Negócio

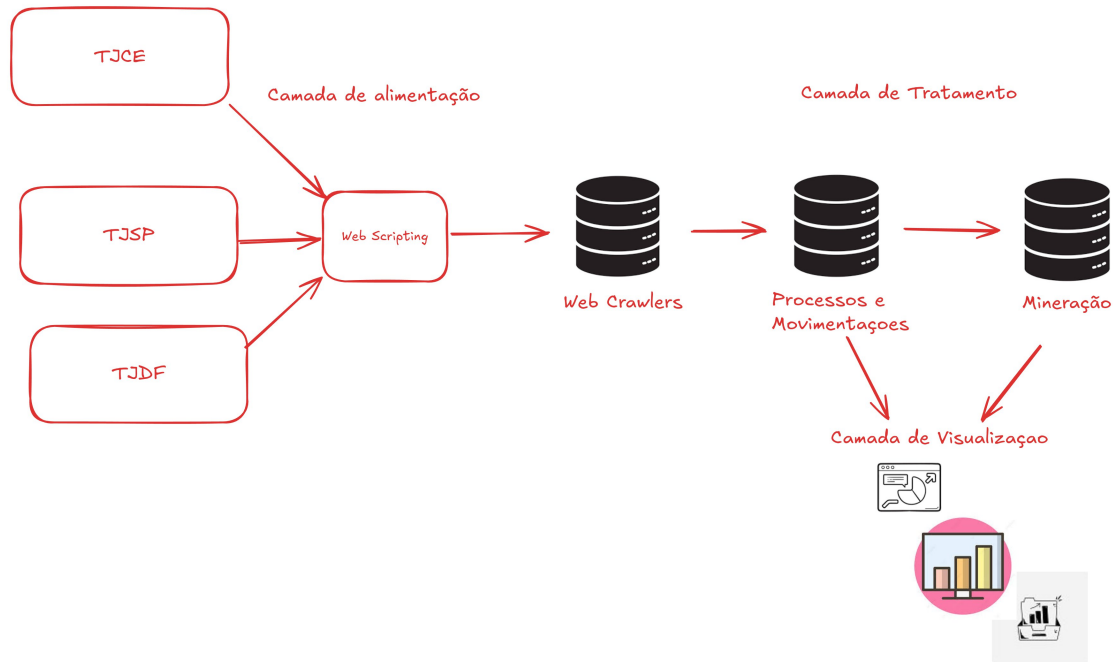
- **Setor:** Jurídico e de análise de dados.
- **Problemas:**
 - **Problema 1:** Necessidade de otimização de tempo para acessar e interpretar informações judiciais.
 - **Problema 2:** Identificação de padrões em processos para avaliar a eficácia de decisões judiciais.
 - **Problema 3:** Necessidade de visualizações claras e diretas para facilitar decisões estratégicas baseadas em dados.

Rastreamento dos Problemas de Negócio

Os problemas e perguntas de negócio associados são descritos abaixo, destacando as decisões estratégicas que podem ser otimizadas com as informações extraídas:

Problema de Negócio (PN)	Pergunta (P)	Tomada de Decisão (TD)
PN1: Identificação de padrões legais	P1: Quais são as leis mais frequentemente aplicadas?	TD1: Usar padrões para identificar decisões comuns
PN2: Análise de movimentações processuais	P2: Qual o tempo médio de cada fase processual?	TD2: Avaliar a celeridade e identificar gargalos
PN3: Efetividade das decisões judiciais	P3: Quais termos correlacionam com decisões positivas?	TD3: Usar análise de linguagem para prever resultados

Arquitetura da Pipeline de Dados



Conforme a simulação de contexto de banco de dados apresentada no diagrama, a pipeline de dados proposta é estruturada nas seguintes camadas:

- **Camada de Alimentação:** Responsável pela coleta dos dados dos tribunais estaduais (TJCE, TJSP e TJDF). Nesta etapa, são utilizados scripts de web scraping para capturar informações brutas de processos e movimentações judiciais.
- **Web Scripting e Web Crawlers:** Utilizados para a extração e a organização dos dados. O web scripting realiza a coleta, enquanto os web crawlers navegam nas fontes judiciais e armazenam os dados extraídos no sistema de armazenamento para posterior processamento.

- **Camada de Tratamento:** Esta fase envolve o pré-processamento dos dados, transformando-os em informações padronizadas e estruturadas para análise. Aqui, os dados passam por processos de limpeza e transformação, sendo organizados conforme atributos como movimentações de processos e detalhes específicos de cada caso judicial.
- **Mineração de Dados:** Após o tratamento, aplicam-se técnicas de mineração de dados para identificar padrões e insights relevantes para o contexto jurídico. Essa fase permite a análise preditiva e a geração de relatórios analíticos que apoiam a tomada de decisões.
- **Camada de Visualização:** Utilizando dashboards e gráficos interativos, esta camada exibe as informações processadas e analisadas de maneira visual, facilitando a compreensão e a comunicação dos resultados aos clientes. A visualização inclui tendências de movimentações judiciais e estatísticas de processos por tribunal, auxiliando advogados, pesquisadores e o público interessado em acompanhar a evolução dos casos.

Considerações Finais

A pipeline proposta para a **DataJustice Analytics** proporciona um fluxo de dados eficiente, desde a coleta até a visualização de insights. A arquitetura apresentada, apoiada por técnicas de mineração de dados e visualização, permite uma análise abrangente e detalhada de processos judiciais, contribuindo para uma maior transparência e eficiência na gestão de dados jurídicos.