



# Apresentação Case

Iuli Hardt

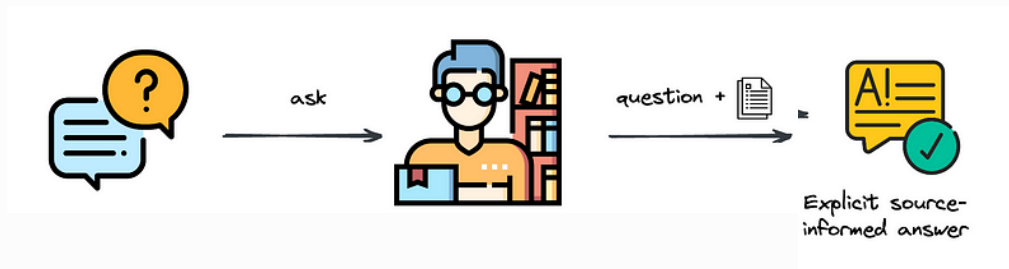
# O problema

A empresa de telecomunicações contratou a A3Data para avaliar o **cenário de churn elevado** dos seus clientes e, uma vez que estamos falando de um produto com custo elevado de setup (instalação), a empresa gostaria de uma estratégia para reduzir esse churn.

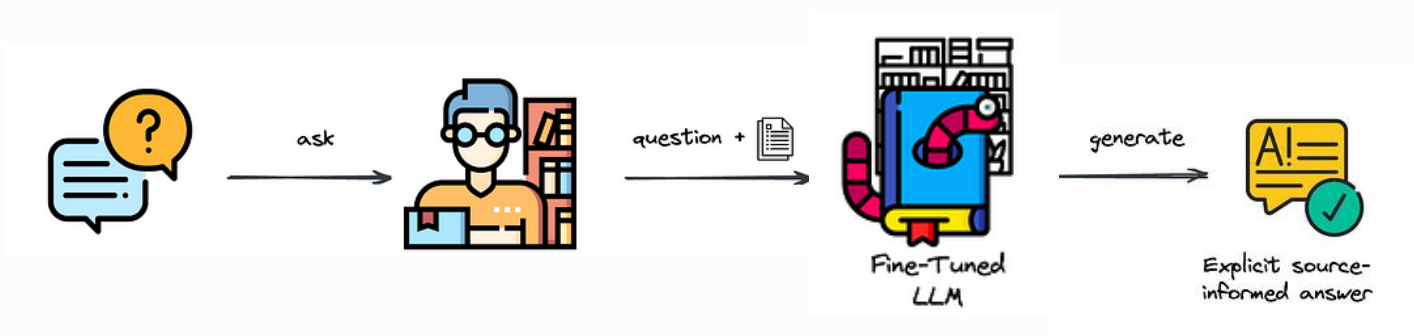
# Solução




# Solução



# Solução



# O processo

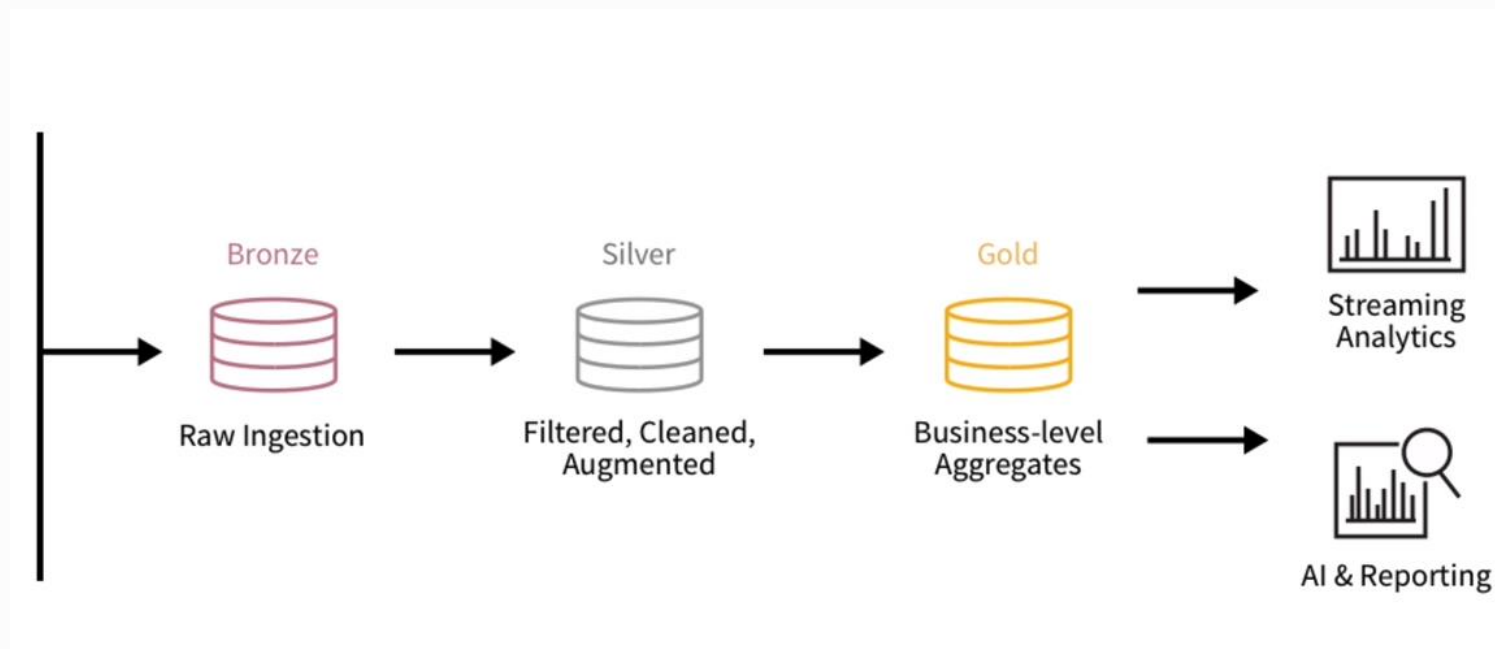
- 01 → Levantamento dos requisitos
  - 02 → Coleta de dados;
  - 03 → Higienização e tratamento de dados
  - 04 → Vetorização
  - 05 → Fine Tune
  - 06 → Testes e produção
- 

# Levantamento dos requisitos;

- Aplicação Q&A;
- Open Source;
- Dados .csv



# Coleta de dados;

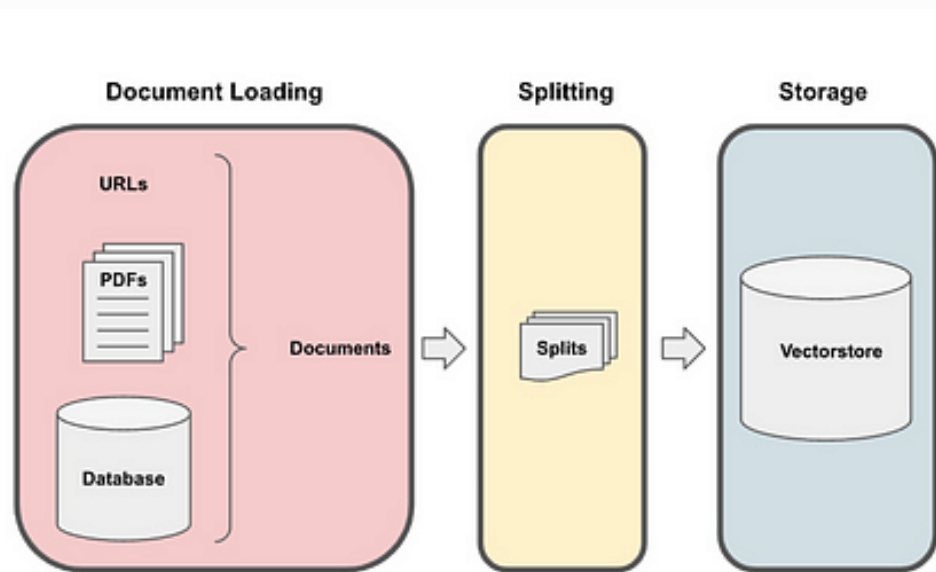




# Higienização e Tratamento;

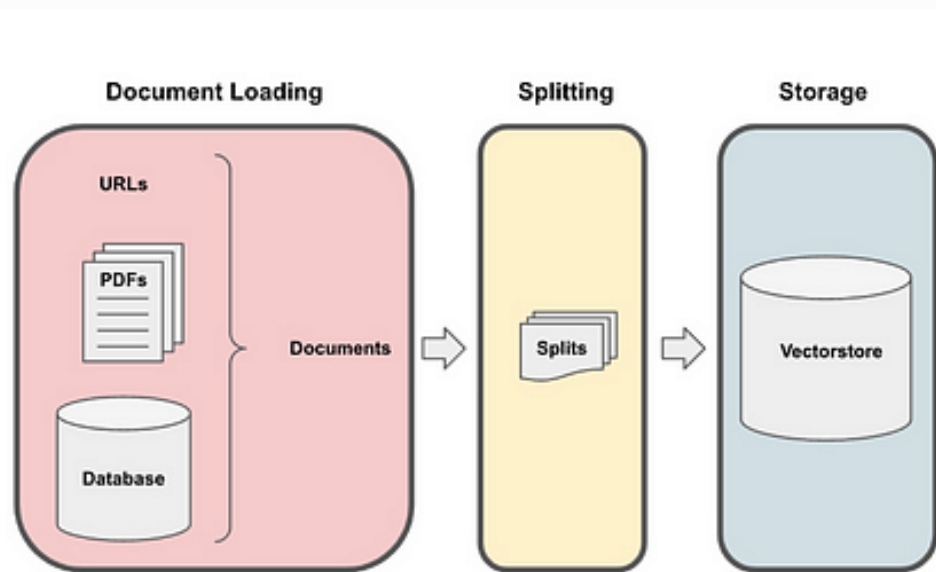


# Vetorização das informações;



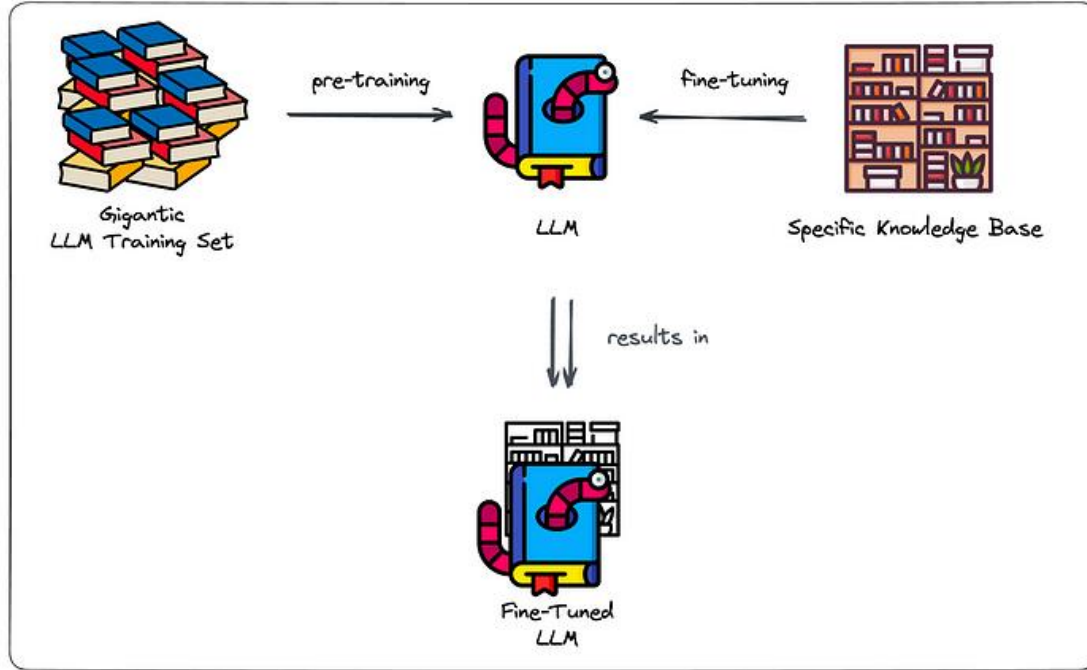
O processo de split de texto para vetorizar refere-se à etapa de dividir um documento de texto em unidades menores, como palavras ou n-gramas, para posteriormente convertê-las em representações numéricas, conhecidas como vetores. A vetorização é uma técnica essencial no processamento de linguagem natural (PLN) e é frequentemente utilizada em tarefas de aprendizado de máquina aplicadas a dados textuais. Após o texto ser dividido em unidades, cada uma dessas unidades é mapeada para um vetor numérico usando métodos como modelos de incorporação de palavras (Word Embeddings), onde palavras são representadas em espaços vetoriais, ou técnicas mais avançadas como embeddings contextuais, como os fornecidos por modelos como BERT. Essa representação vetorial permite que algoritmos de machine learning compreendam e processem informações textuais de maneira mais eficaz, sendo útil em tarefas como classificação de texto, agrupamento de documentos e análise de sentimentos.

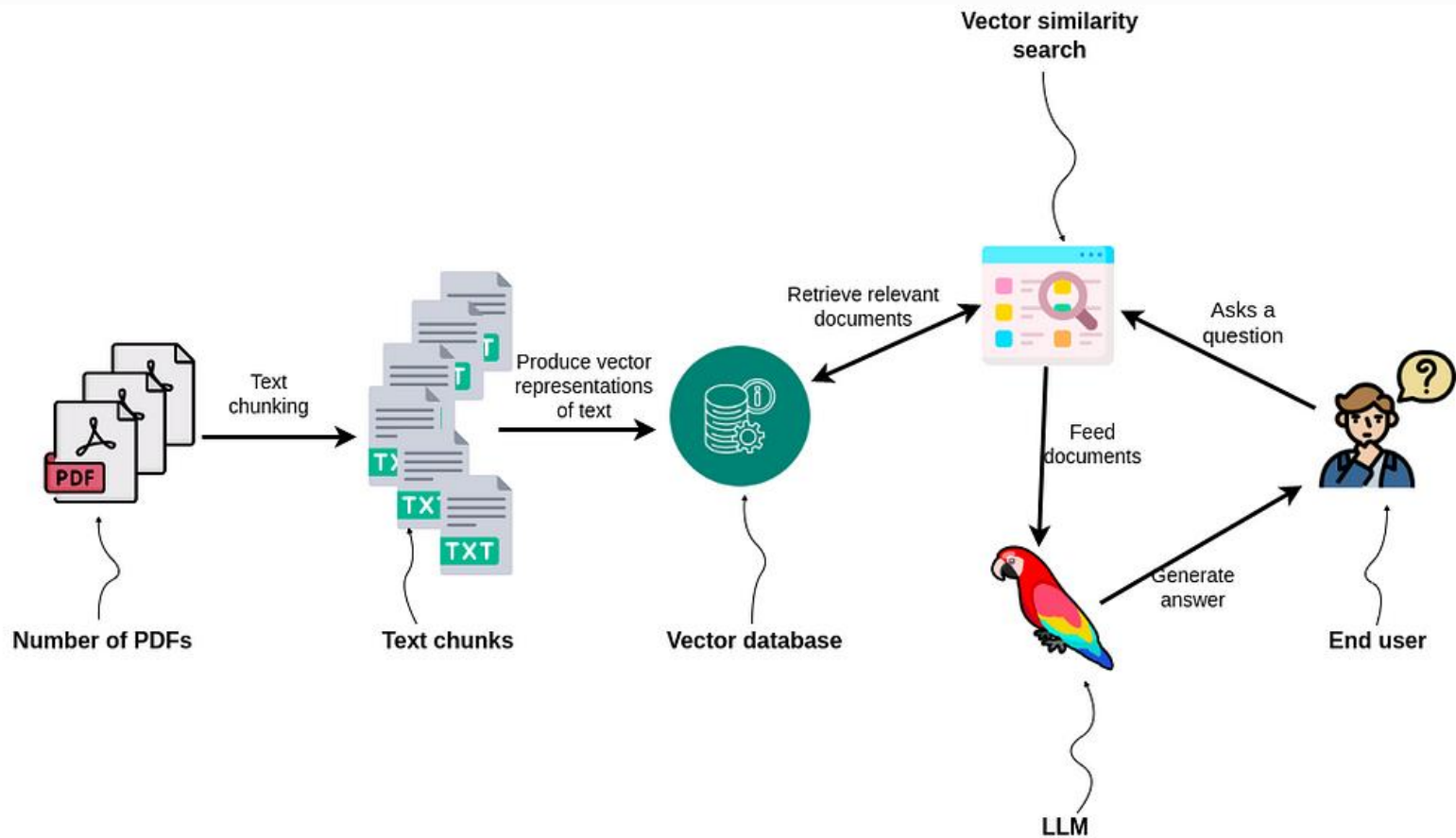
# Vetorização das informações;



O processo de split de texto para vetorizar refere-se à etapa de dividir um documento de texto em unidades menores, como palavras ou n-gramas, para posteriormente convertê-las em representações numéricas, conhecidas como vetores. A vetorização é uma técnica essencial no processamento de linguagem natural (PLN) e é frequentemente utilizada em tarefas de aprendizado de máquina aplicadas a dados textuais. Após o texto ser dividido em unidades, cada uma dessas unidades é mapeada para um vetor numérico usando métodos como modelos de incorporação de palavras (Word Embeddings), onde palavras são representadas em espaços vetoriais, ou técnicas mais avançadas como embeddings contextuais, como os fornecidos por modelos como BERT. Essa representação vetorial permite que algoritmos de machine learning compreendam e processem informações textuais de maneira mais eficaz, sendo útil em tarefas como classificação de texto, agrupamento de documentos e análise de sentimentos.

# Fine Tune de modelos offline;





# RoadMap

- FrontEnd;
- Serviço na nuvem;
- Acompanhamento e evolução do template prompt
- Uso de Outro modelo Llama2 para fine tune

