

BIUD Talks



BANCO VETORIAL

Marcelo Luk - Dev

Cronograma do Programa

 22/04

Wanderson

Inteligência Artificial para Dev

 29/04

Praxedes

Testes Unitários

 06/05

Luk

Arquitetura de Software

 13/05

Regis

Integração de Sistemas

 20/05

Rogério

Root Cause Analysis (RCA)

 27/05

Gustavo

Metodologia Ágil

 03/06

Luk

CR – Code Review

Cronograma do Programa

10/06	Thiago	IA na prática: Claude Code
17/06	Guilherme	Segurança de dados no ambiente corporativo e pessoal
24/06	Jean	Machine Learning
01/07	Santiago	CNV - Comunicação não Violenta no ambiente de trabalho
08/07	Rogério	Saúde Mental
15/07	Guilherme	DevOps e Infraestrutura
22/07	Gustavo	Ambidestria Organizacional

Cronograma do Programa

29/07

Santiago

Mindset Ágil

05/08

Jean

RAG

12/08

????

Pode ser você! Envie seu tema para o Rogério.

Estrutura de cada Sessão

Apresentação

20-30 minutos

Conteúdo e expertise do palestrante

Debate & Discussão

30 minutos

Dúvidas, insights e networking

 **Moderador: Rogério**

Responsável por mediar o debate e garantir engajamento de todos os participantes

BANCO VETORIAL



Um banco vetorial é um banco de dados feito para guardar listas de números chamadas vetores.

Esses números vêm de textos, fotos ou áudios transformados por inteligência artificial.

Em vez de buscar palavras exatas, ele acha dados parecidos pelo significado, usando contas de proximidade

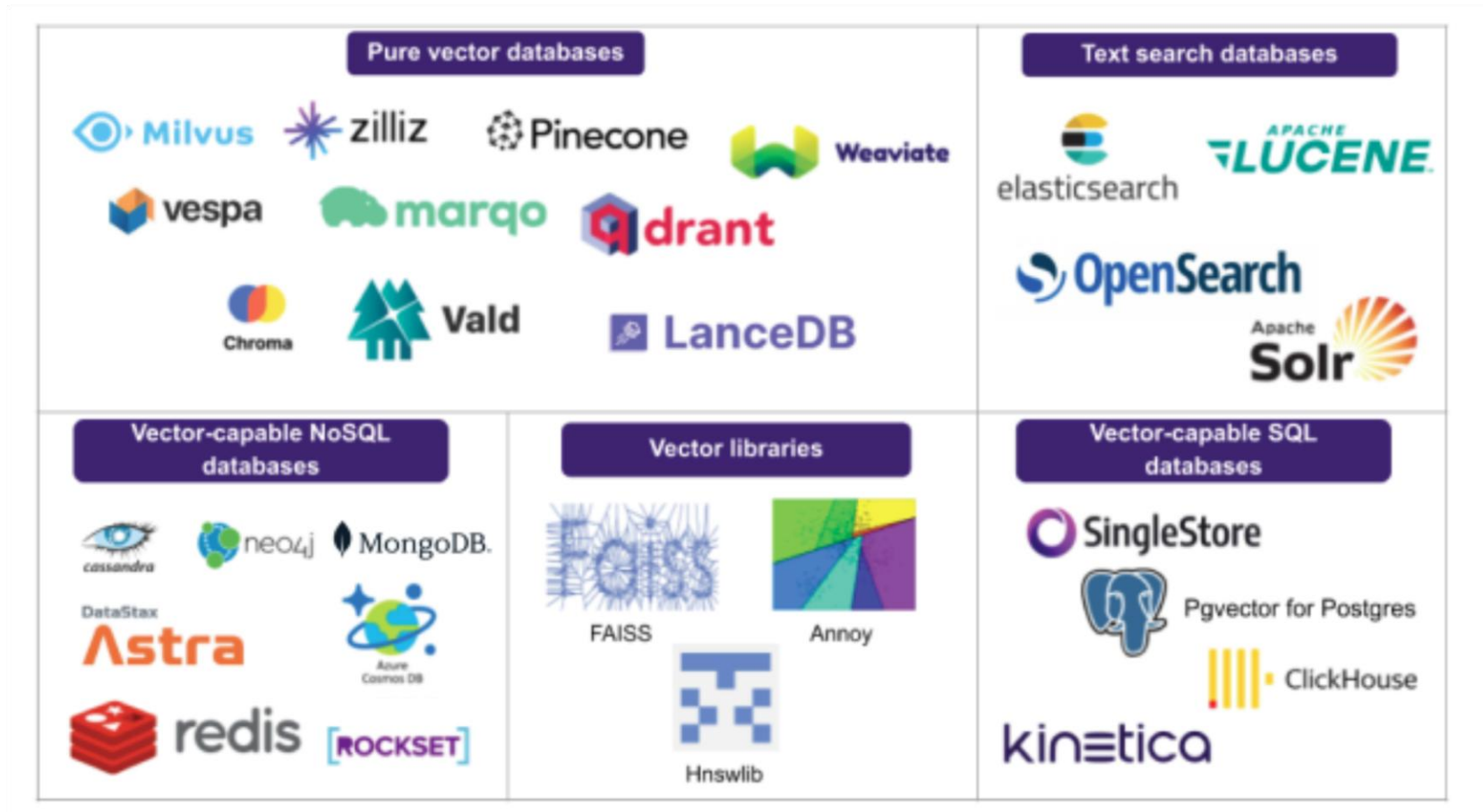
Busca Semântica: Achar documentos pelo assunto, mesmo sem usar as mesmas palavras.

IA Generativa (RAG): Dar uma base de dados externa para o ChatGPT ou outros robôs consultarem fatos reais.

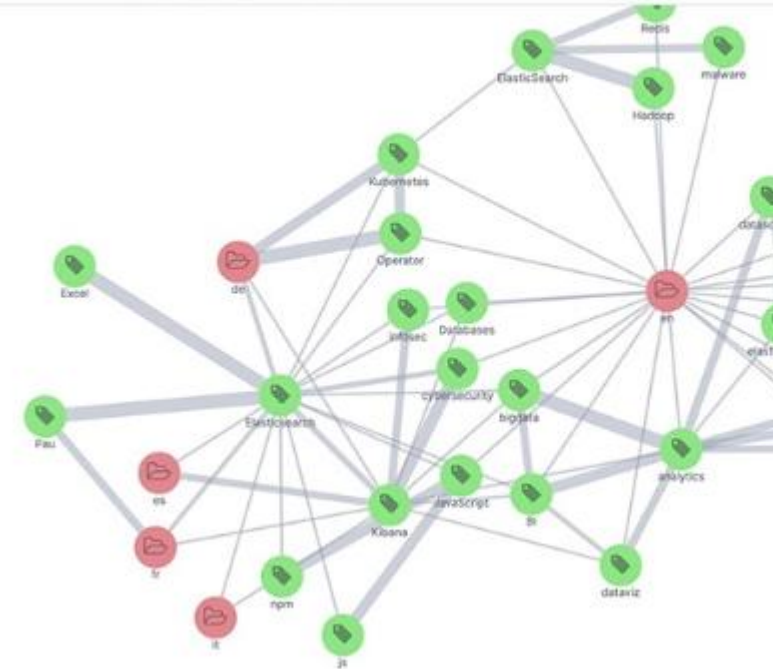
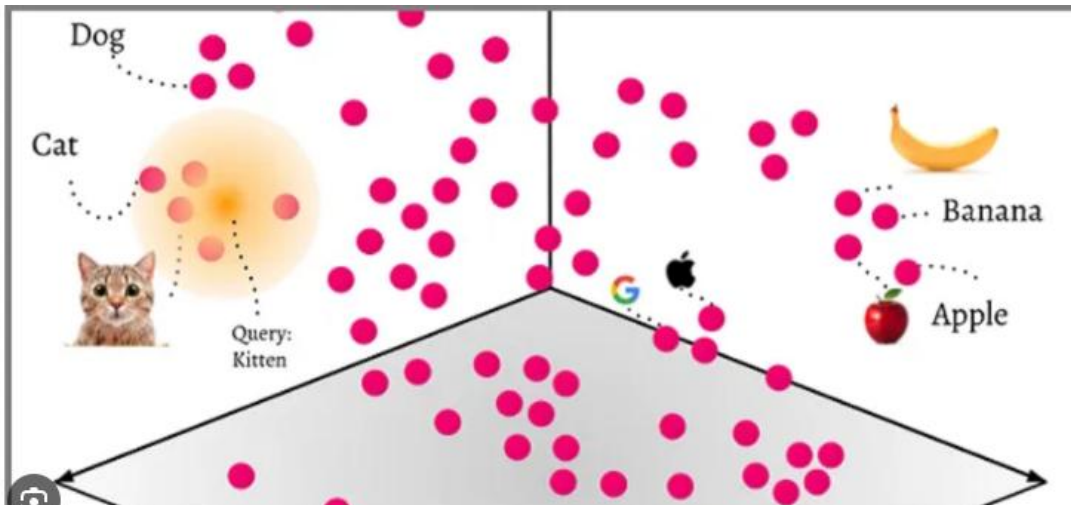
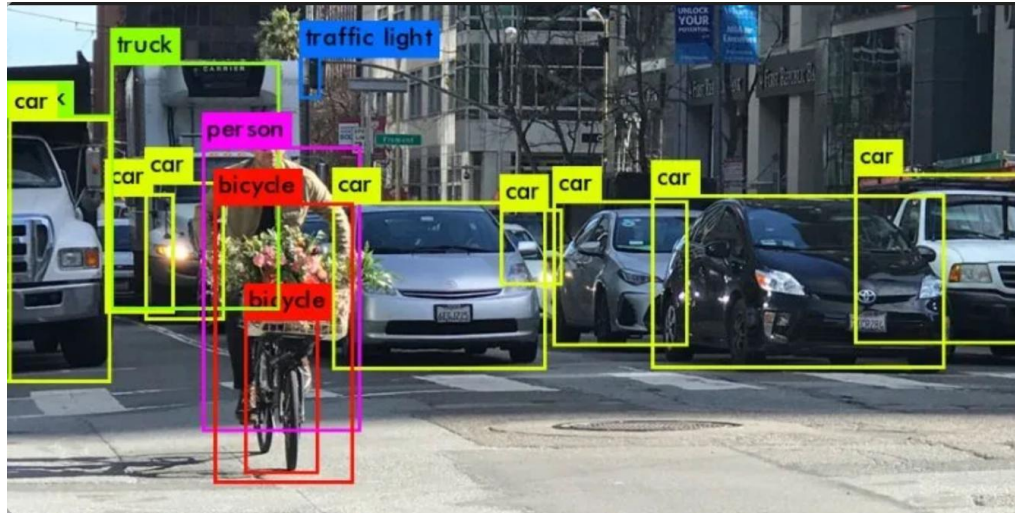
Recomendações: Sugerir produtos ou músicas parecidas com o seu gosto.

Busca por Imagem: Achar fotos iguais a partir de uma imagem de exemplo.

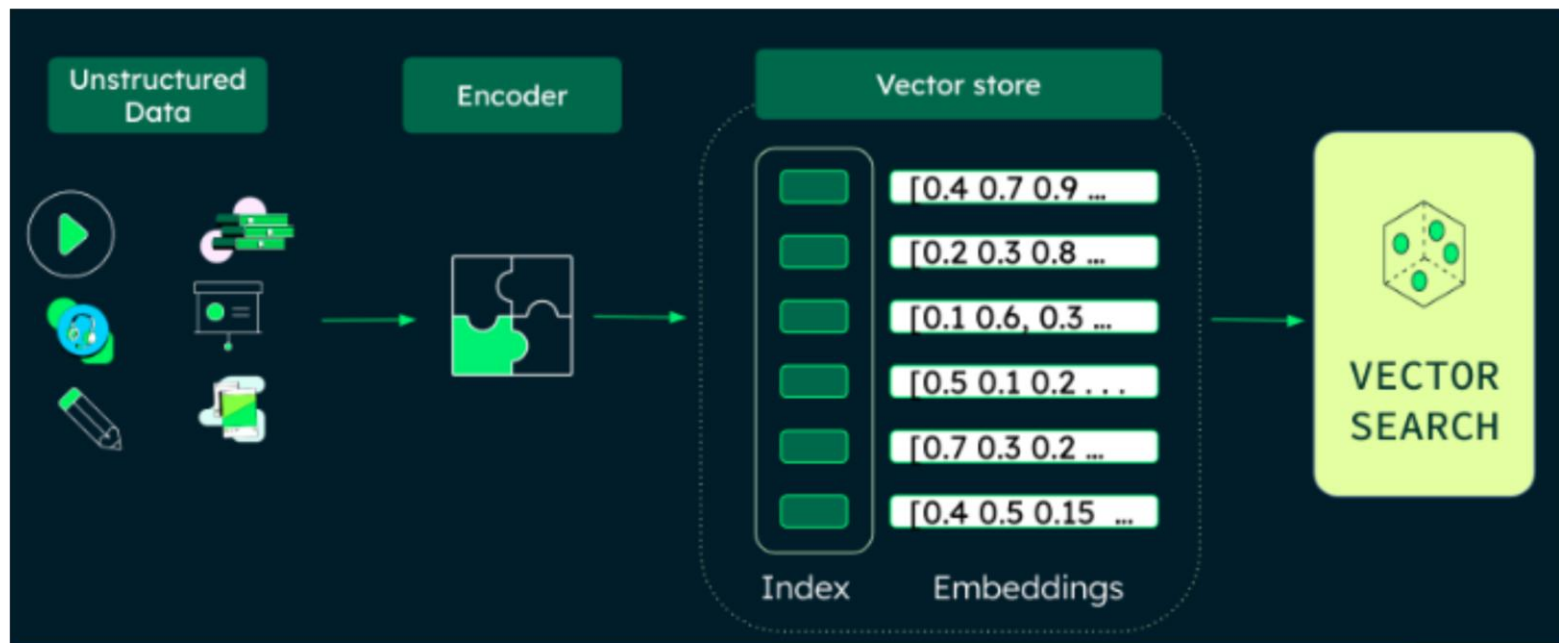
BANCOS VETORIAIS



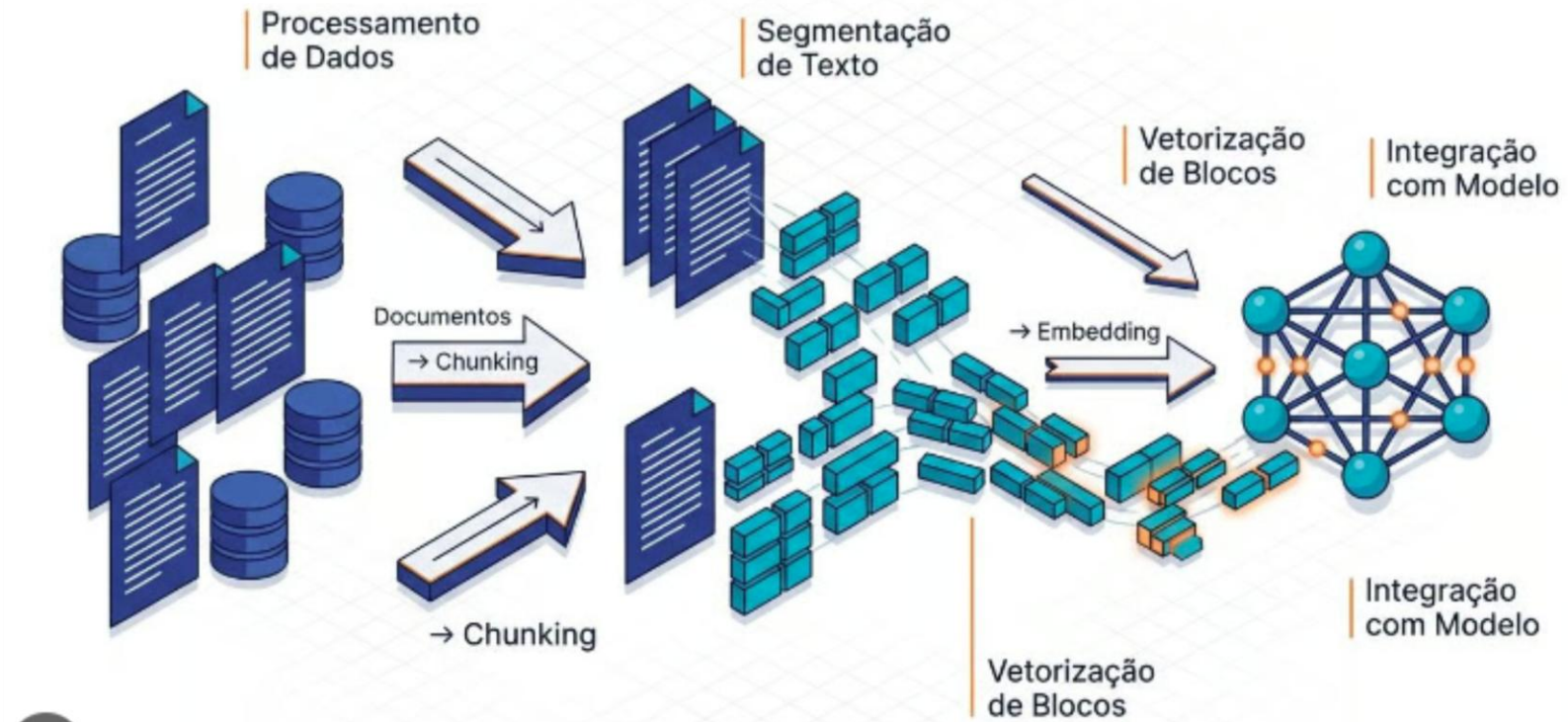
EXEMPLOS



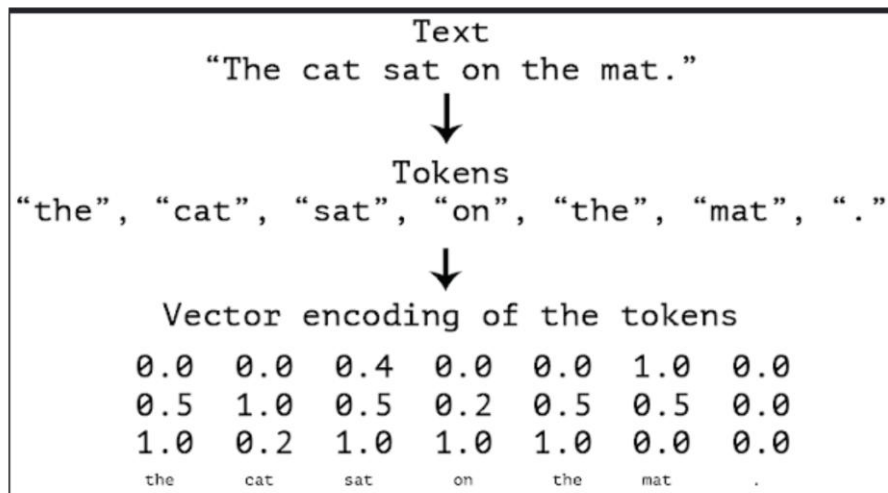
Como funciona



Como funciona?



CHUNK



Configuration

☐ Custom settings



llama-text-embed-v2
NVIDIA - HOSTED

State-of-the-art text embedding model.

 Inference API



multilingual-e5-large
MICROSOFT - HOSTED

Top performing text embedding model
from Microsoft research.

 Inference API



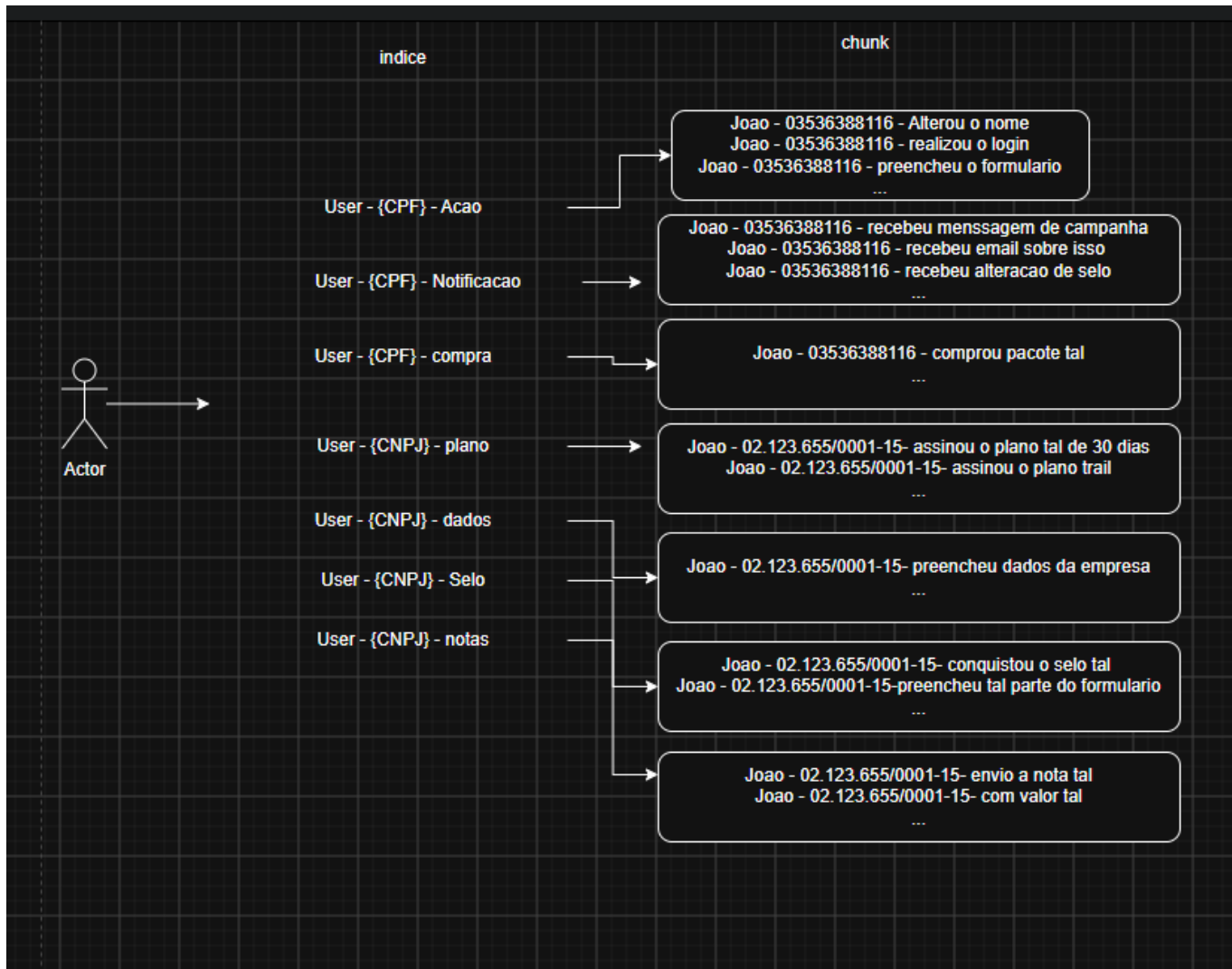
pinecone-sparse-english-v0
PINECONE - HOSTED

Sparse vector model for keyword-style
search.

 Inference API



Segmentação por índices





PLAY

Abrindo o Debate



💬 Questões para reflexão:

1. Qual foi o principal aprendizado para você?
2. Como você pode aplicar isso em seu dia a dia?
3. Quais foram as dúvidas geradas?
4. Que conexões você vê com outros tópicos?

Obrigado!

Até o próximo episódio da BIUD Talks

💡 Inovando juntos | 🚀 Crescendo juntos