

Artificial Intelligence for Aviation Data Analysis



Equipe



Marcelo Xavier Guterres



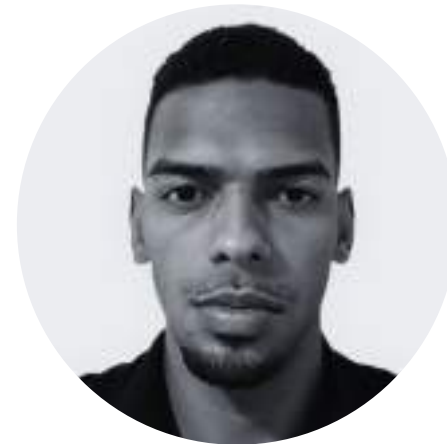
Flávio Mendes Neto



Dimas Betioli Ribeiro



**Jean Phelipe de
Oliveira Lima**



**Guilherme Trindade
Tolentino Bernardo**



**João Basílio
Tarelho Szenczuk**



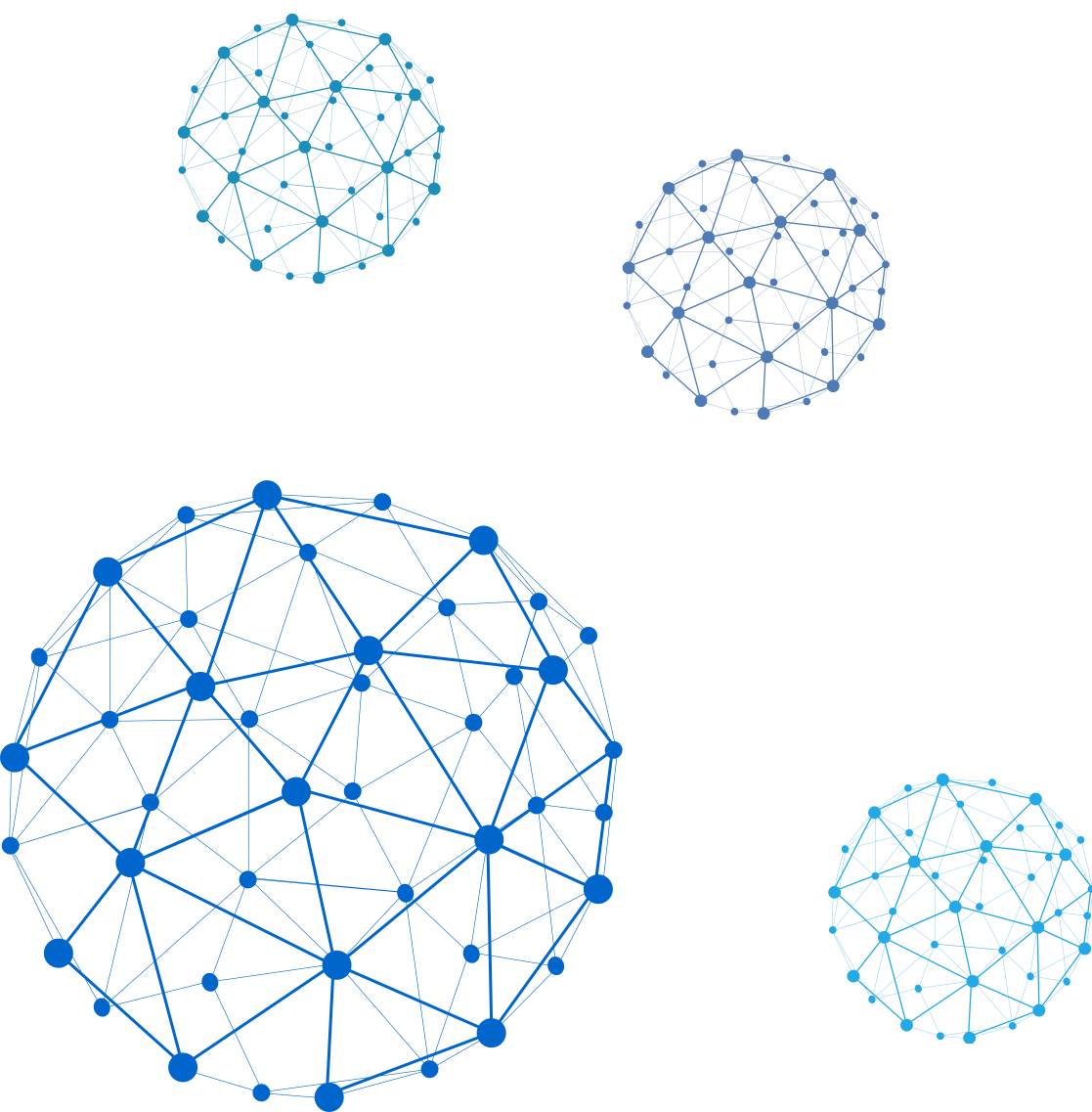
**Vitor Lucas Kohls
Correa**

Agradecimentos

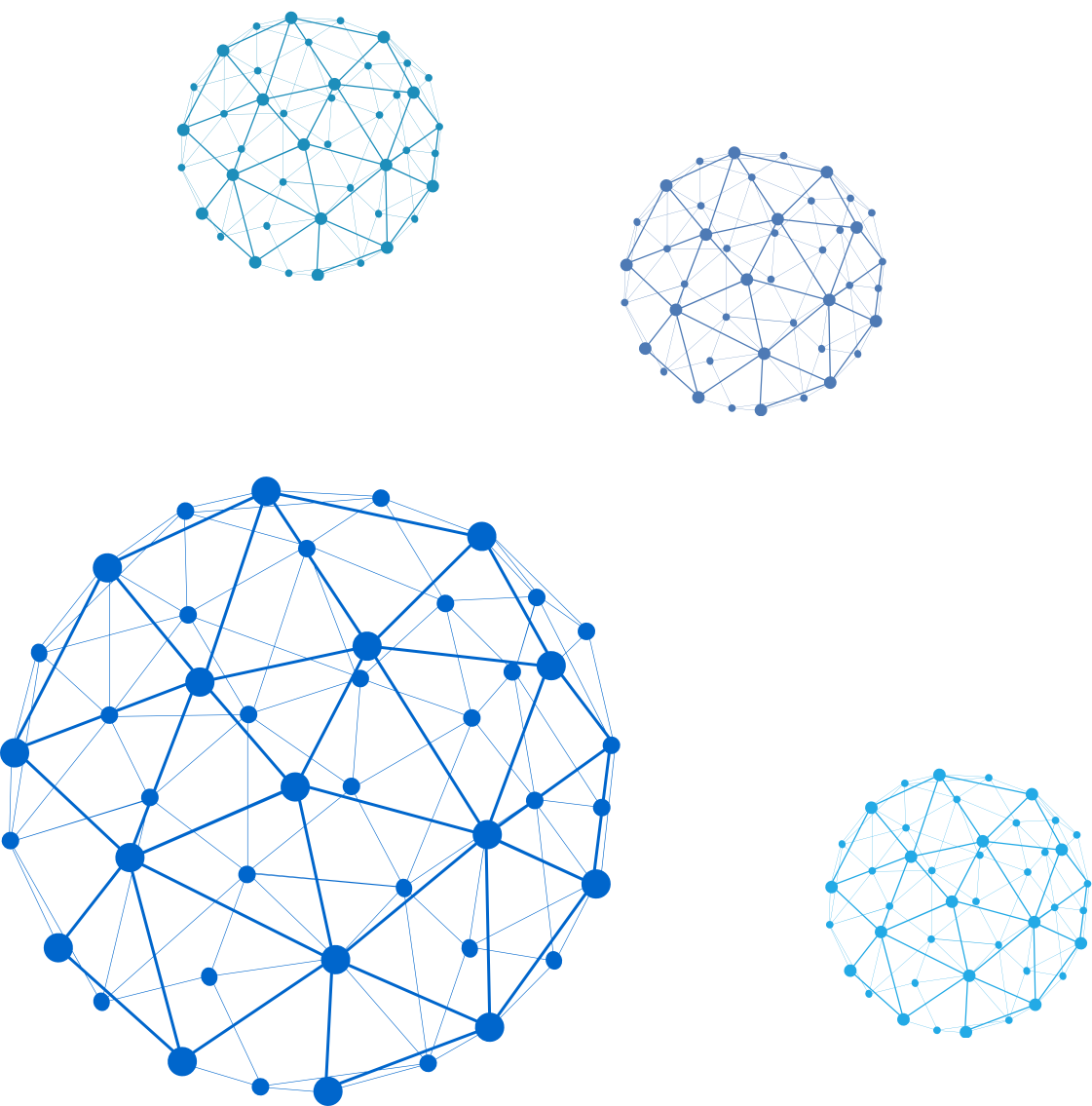
MINISTÉRIO DE
PORTOS E
AEROPORTOS



Departamento de Controle
do Espaço Aéreo



Motivação

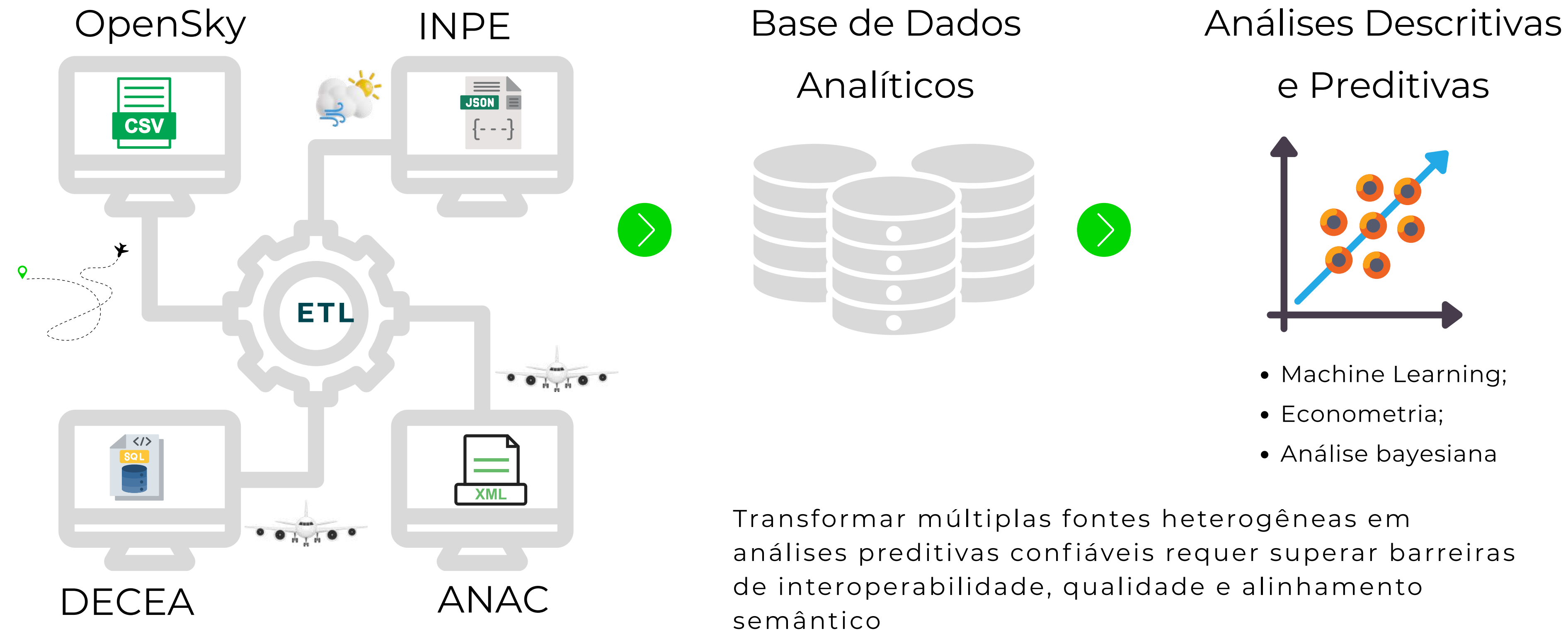


Resultado esperado - que chatbot seja capaz de responder perguntas como:

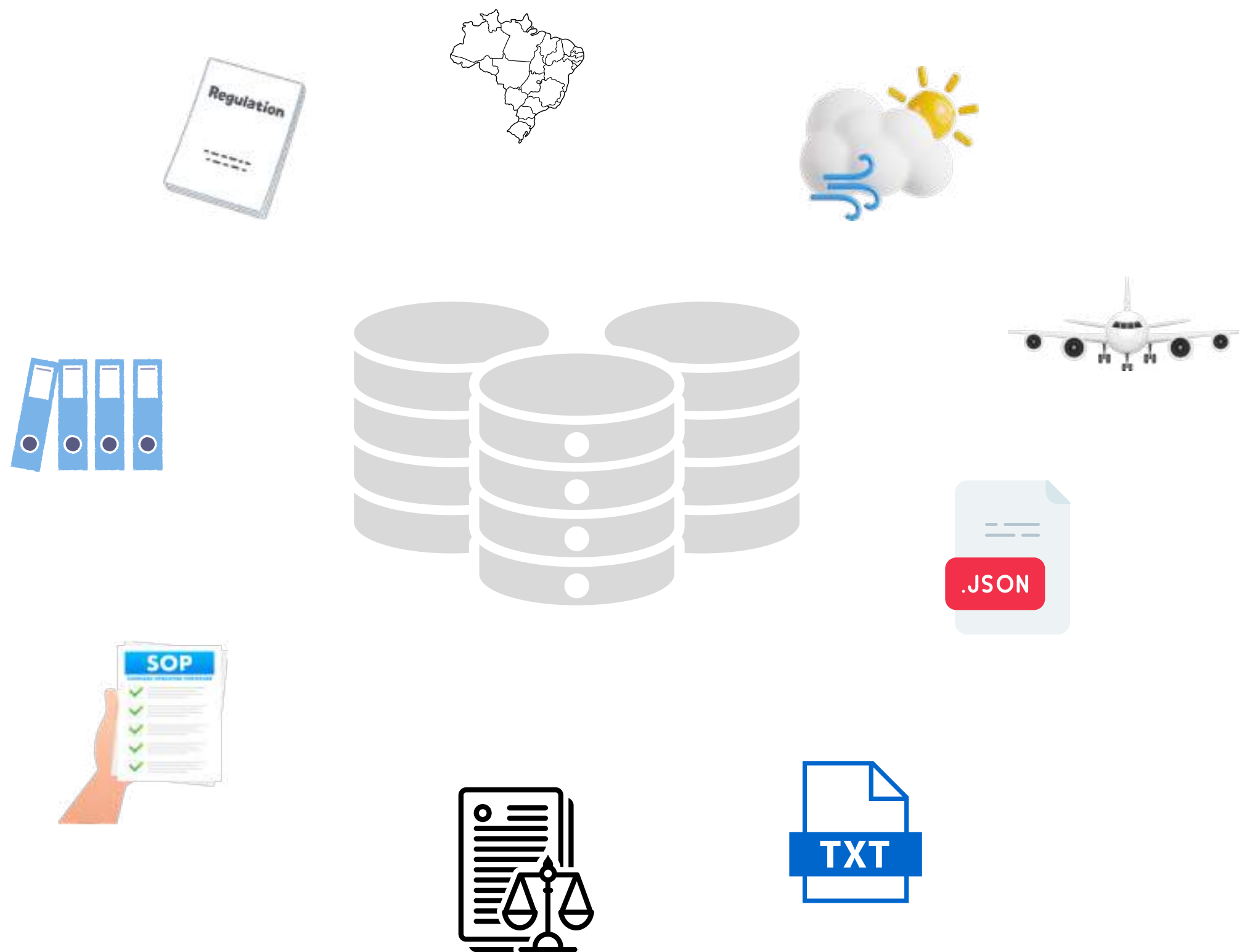
Por que o LATAM 5411 atrasou?

- Consulta o RDF → atraso de 45min por motivo X;
- Consulta o KB → cláusula da ICA 100-12 sobre nevoeiro em SBGR;
- Responde: “**Segundo o histórico de voo, o atraso de 45min foi causado por condições meteorológicas (visibilidade reduzida). De acordo com a ICA 100-12, tais condições exigem aumento no espaçamento mínimo entre pousos.**”

Base de Dados Analíticos

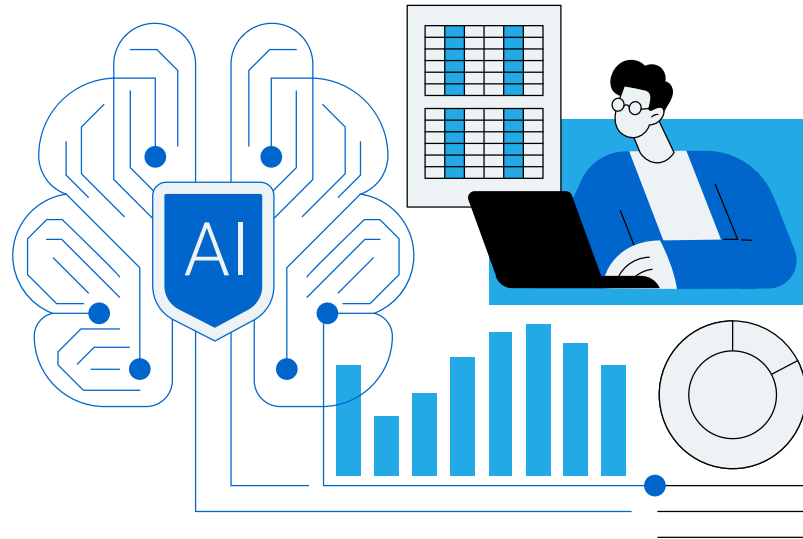


Base de Conhecimento Semântica



Essa base não trata apenas de dados numéricos, mas de **informações contextuais e semânticas** sobre o setor.

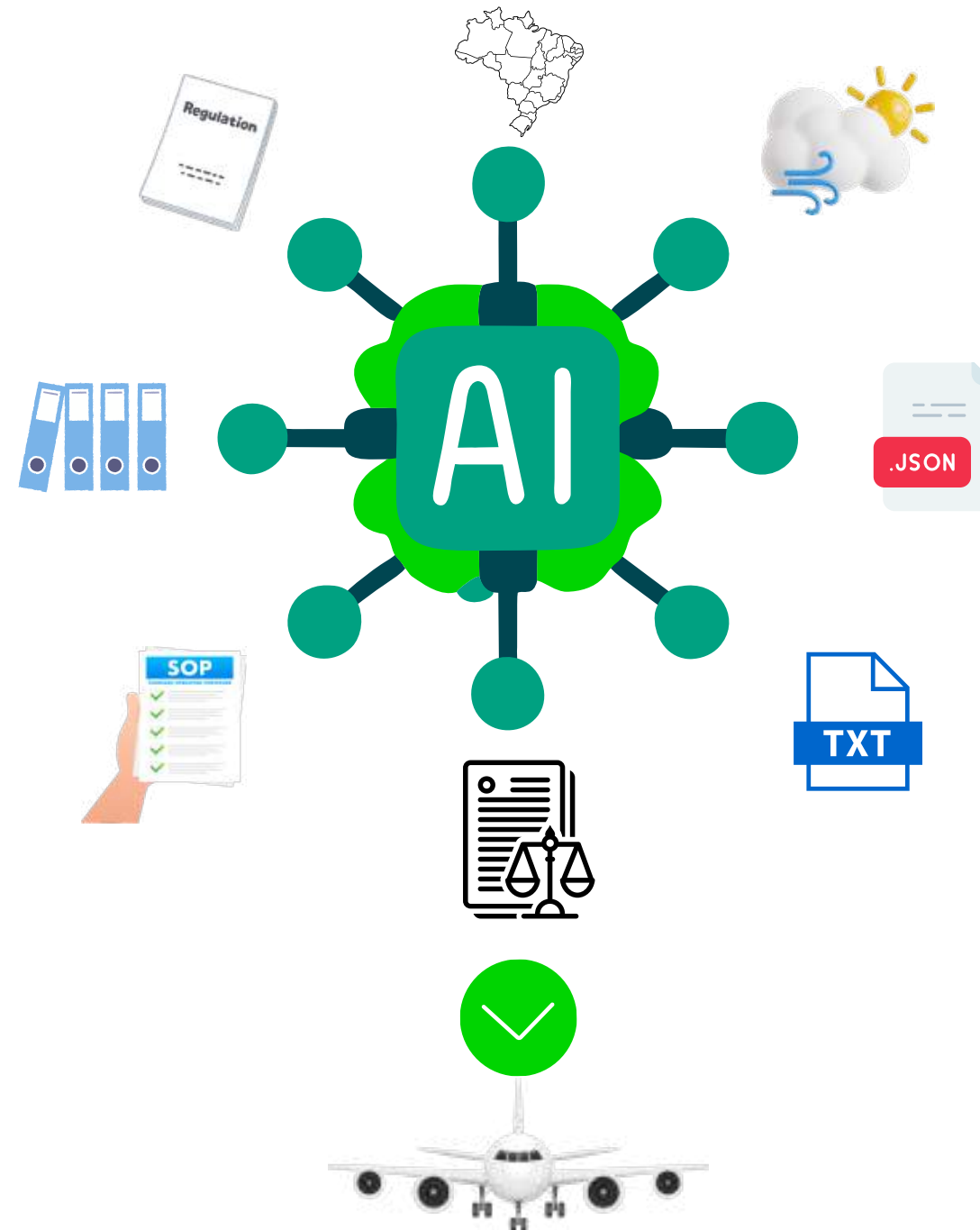
Problema



Como harmonizar e transformar dados fragmentados, complexos e distribuídos, oriundos de múltiplas fontes com formatos, semânticas e níveis de qualidade distintos , em uma **base analítica universal, confiável e interoperável**, que sirva de alicerce para o desenvolvimento de uma **inteligência artificial** especializada e efetiva na gestão do transporte aéreo brasileiro, considerando a **alta sensibilidade regulatória**, a **diversidade operacional** e a **constante atualização normativa do setor**?

Objetivo

Base de dados Analíticos



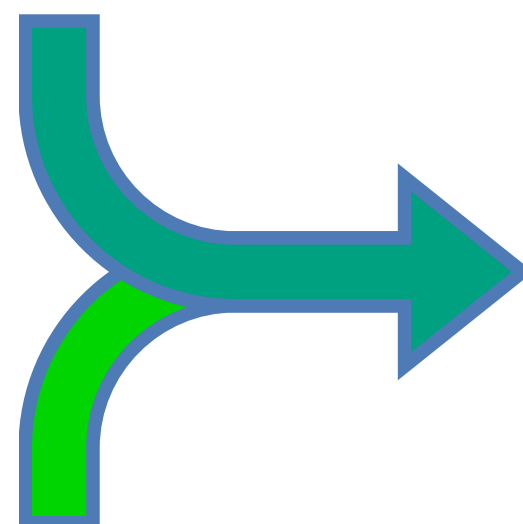
Base de Conhecimentos



Como usar uma IA para dar contexto, explicabilidade e suporte inteligente à decisão e gestão na aviação civil?

1º TRILHA — ONTOLOGIA + GRAFO DE CONHECIMENTO ANALÍTICO

Objetivo: transformar dados operacionais fragmentados (ANAC, DECEA, OpenSky, etc.) em um grafo semanticamente estruturado e consultável.



2º TRILHA — BASE DE CONHECIMENTO SEMÂNTICA

Objetivo: criar um repositório semântico e vetorial de conhecimento normativo e contextual sobre aviação civil (ICA, RBAC, manuais, IATA, EUROCONTROL, etc.) para alimentar modelos de linguagem (LLMs).

PONTO DE CONVERGÊNCIA: INTEGRAÇÃO VIA RAG

O QUE O GRAFO RDF RESPONDE:

- "QUEM": LATAM VOO 5411 "
- O QUÊ": ATRASO DE 45 MINUTOS
- "QUANDO": 17/07/2025 ÀS 13:30
- "ONDE": SBGR → SBPA

O QUE A BASE VETORIAL RESPONDE:

- "POR QUÊ": Condições meteorológicas (nevoeiro)
- "REGRA": ICA 100-12, seção 4.2.3
- "CONTEXTO": "Visibilidade < 600m exige espaçamento mínimo de 3nm"

O que o LLM irá entregar:

"O voo LATAM 5411 (SBGR→SBPA) teve atraso de 45min devido a nevoeiro em Guarulhos. Segundo a ICA 100-12, visibilidade reduzida exige aumento no espaçamento entre aeronaves, impactando a capacidade do aeroporto e gerando atrasos operacionais."

Convergência

ONTOLOGIA + GRAFO
DE CONHECIMENTO
ANALÍTICO

BASE DE
CONHECIMENTO
SEMÂNTICA

Trilha 01

Trilha 02





Conceito	Semântica	Ontologia
 Definição	É o significado de um termo, conceito ou estrutura	É uma estrutura formal que representa conceitos, relações e regras
 O que é	Uma propriedade geral da linguagem	Um modelo de conhecimento de um domínio
 Como funciona	Dá sentido ou interpretação a um dado, campo ou texto	Organiza esse conhecimento em classes, propriedades e relações
 Exemplo	Saber que aero_dep = “aeroporto”	Criar uma classe Aeroporto com codigo_icao, nome etc.
 Papel no Airdata	Traduz nomes e valores para o que eles realmente significam	Modela tudo isso em um grafo estruturado e consultável

Dados de diversas fontes:

- ✈️ ANAC chama de **aero_dep**,
- ✈️ DECEA chama de **id_airport**,
- 🛫 OpenSky usa **id_iata**.

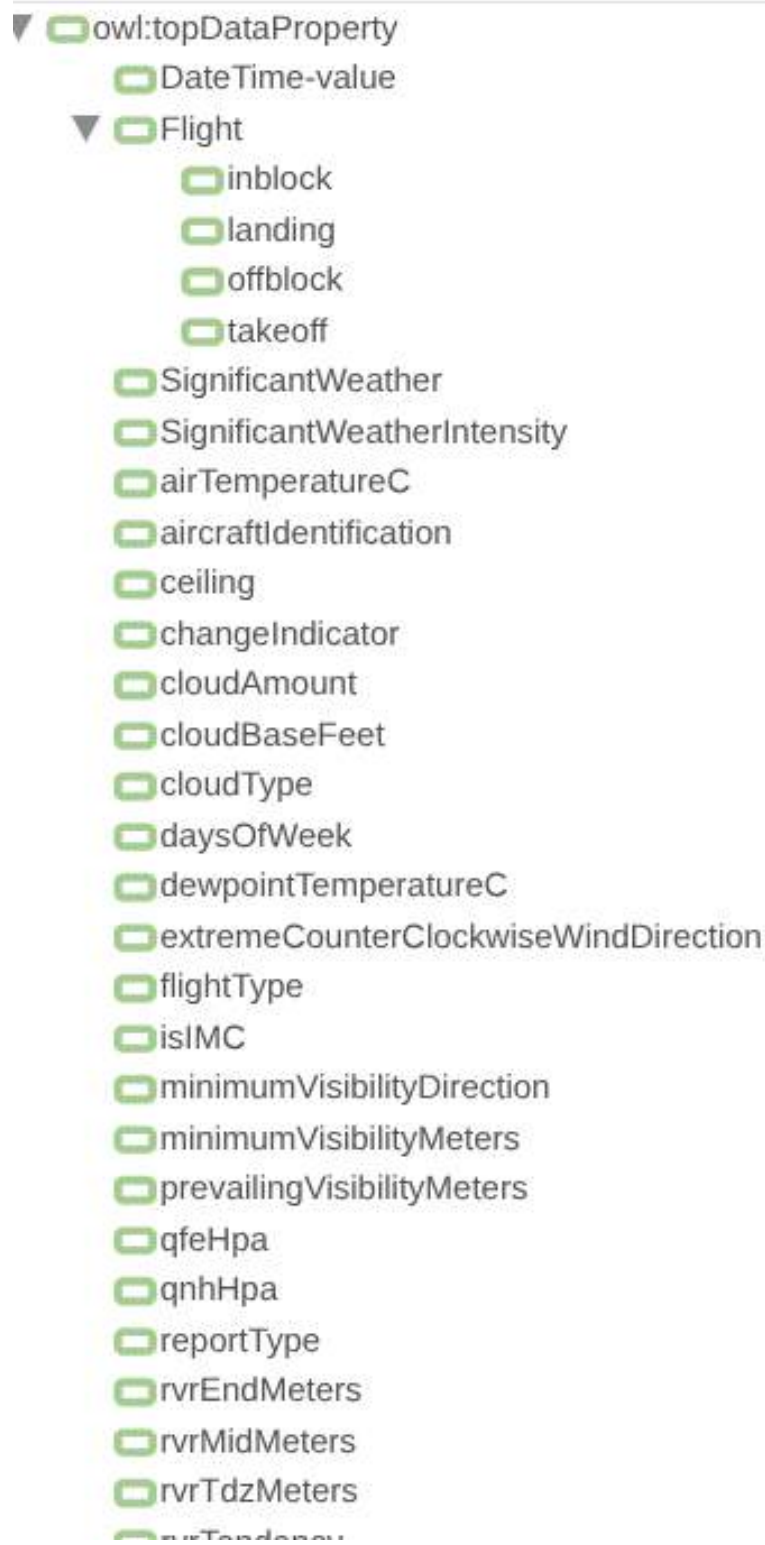
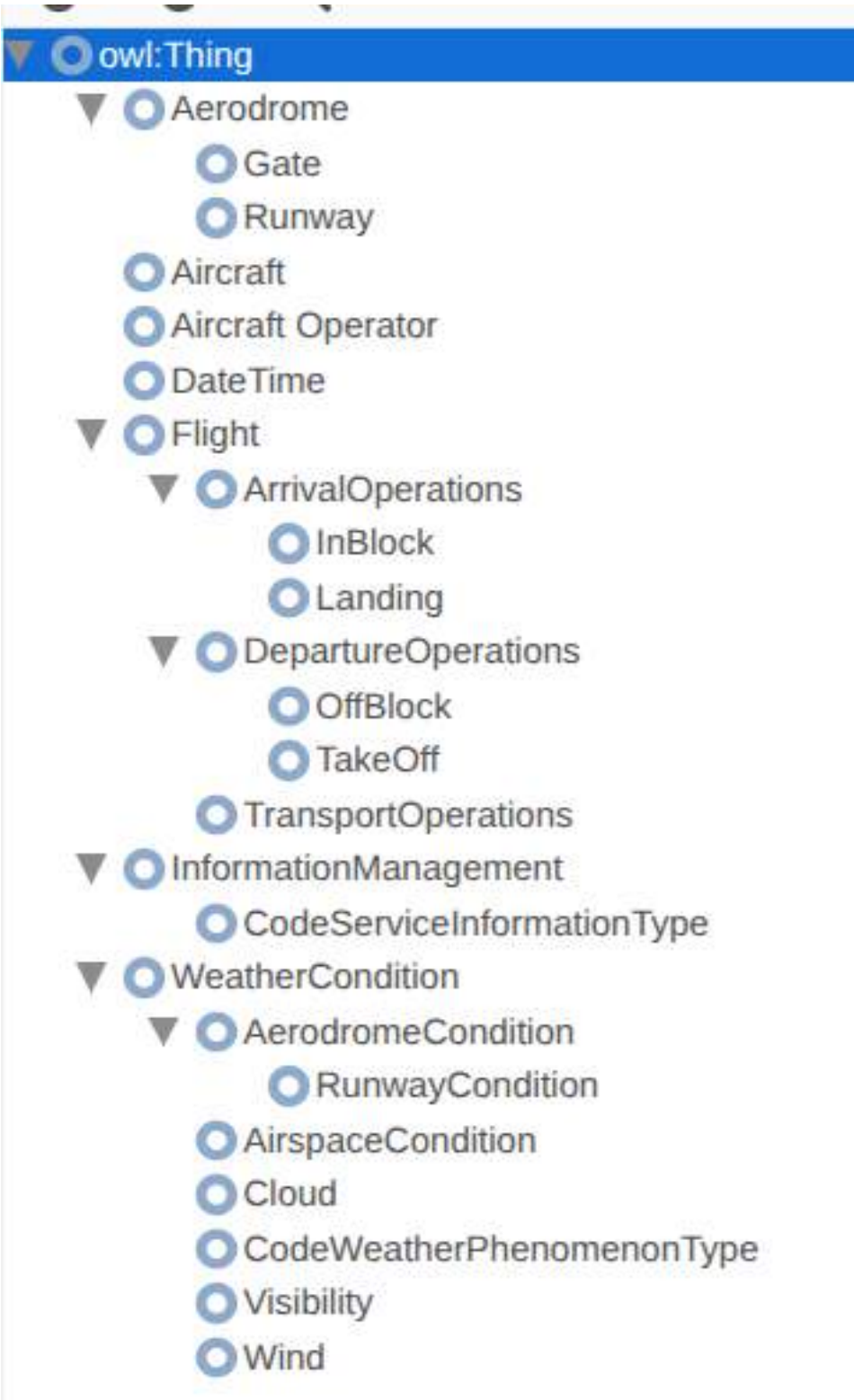
👉 A semântica expressa: **tudo isso quer dizer 'aeroporto'**.

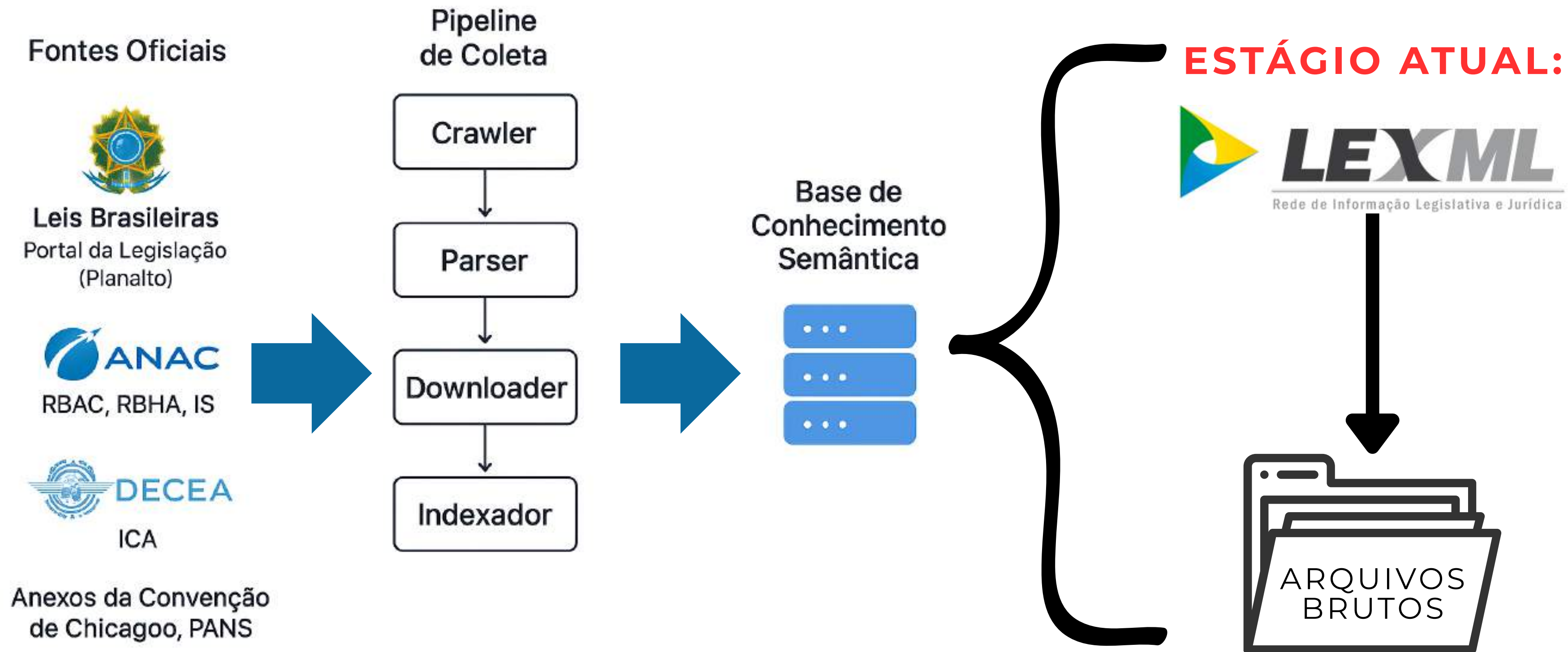
👉 A ontologia expressa: “**Aeroporto é uma entidade que tem propriedades codigo_icao, nome, e está relacionada a voos, localizações etc.**”

LLMs precisam de uma **Camada Semântica** para extrair o máximo de valor a partir de dados armazenados.



1. LLMs não “entendem” dados estruturados nativamente.
2. A semântica dá contexto e significado real aos dados
3. Facilita inferências e explicações mais ricas.
4. Permite integração com fontes múltiplas e heterogêneas.
5. É a base para **RAG (Retrieval-Augmented Generation)** eficiente.
6. Gera rastreabilidade e transparência.





Equipe



Marcelo Xavier Guterres



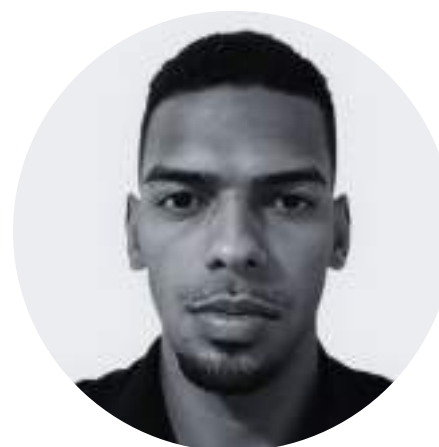
Flávio Mendes Neto



Dimas Betioli Ribeiro



**João Basílio
Tarelho Szenczuk**



**Guilherme Trindade
Tolentino Bernardo**



**Jean Phelipe de
Oliveira Lima**



**Vitor Lucas Kohls
Correa**