



Flávia Gaia

CIENTISTA DE DADOS SÊNIOR | IA GENERATIVA E LLMS

+55 (38) 9 9898-7165

Brasília, Brasil

<https://linktr.ee/FlaviaGaia>

flaviaggpaula@gmail.com

RESUMO PROFISSIONAL

Cientista de Dados Sênior com sólida experiência no desenvolvimento de soluções de Inteligência Artificial aplicadas a problemas complexos de negócio. Atuação especializada em IA Generativa, LLMS, arquiteturas RAG e sistemas multiagentes, com forte foco em ambientes produtivos, governança, automação e confiabilidade. Experiência em projetos de auditoria, compliance, processamento de documentos, Big Data e arquitetura de dados, com desenvolvimento end to end em ambientes Databricks, cloud e plataformas corporativas.

EXPERIÊNCIA PROFISSIONAL

G4F ○ Cientista de Dados Sênior

- Estruturação de ambiente de dados baseado em Medallion Architecture em Azure e Databricks, utilizando Python e Jupyter Notebooks para o desenvolvimento de pipelines escaláveis com foco em CI/CD.
- Desenvolvimento de modelos de Machine Learning utilizando Scikit-learn, TensorFlow e PyTorch para análise de padrões e previsões institucionais, fundamentados em conceitos de Estatística Aplicada.
- Manipulação de grandes volumes de dados com Pandas e PySpark, integração de bancos de dados relacionais e não relacionais para batimento, conciliação e validação de bases via Palantir Foundry.
- Implementação de rotinas de qualidade e governança de dados, com a criação de dashboards de monitoramento e suporte à decisão utilizando Palantir Foundry e Databricks.

BBTS - Banco do Brasil Tecnologia e Serviços ○ Cientista de Dados Sênior Aplicado a Engenharia de IA

- Criação de sistema inteligente automatizado de calendário de pagamentos a partir de contratos em PDF.
- Extração de cláusulas financeiras com LLMS via LangChain e arquitetura multiagente com CrewAI.
- Implementação de agentes para auditoria, detecção de anomalias e supervisão do fluxo.
- Orquestração do pipeline em Databricks Workflows, com persistência em Delta Lake.

Compass UOL ○ Cientista de Dados Especialista em IA

- Extração de informações técnicas de PDFs com LLM.
- Criação de RAG de normativos e documentações.
- Desenvolvimento de interface em Streamlit para validação das extrações.
- Avaliação e versionamento de prompts e métricas com MLflow.
- Implementação sistema multi-agentes com CrewAI.

EVCOMX ○ Cientista de Dados Especialista em IA

- Extração inicial de informações técnicas com regex e evolução para LLMS via LangChain.
- Desenvolvimento de interface em Streamlit para validação das extrações.
- Avaliação e versionamento de prompts e métricas com MLflow.
- Implementação de Human in the Loop com CrewAI para feedback e melhoria contínua.

Ministério MDS ○ Analista de Dados

- Análise de grandes bases governamentais como Bolsa Família, BPC e Cadastro Único.
- Processamento distribuído em ambiente Big Data utilizando Hadoop e PySpark.
- Construção de indicadores analíticos com QlikSense e Power BI para apoio a políticas públicas.

**Presidência
da
República**

Estagiária de Ciência de Dados

- Coleta automatizada de notícias via web scraping com Python.
- Processamento textual com NLP e modelos em português do spaCy para NER e sumarização.
- Visualização analítica das informações em dashboards no Power BI.

FORMAÇÃO ACADÊMICA

**Universidade
de Brasília**
2024-2026

Mestrado em Computação Aplicada

Linha de pesquisa em Ciência de Dados

**Universidade
de São Paulo**
2022-2023

MBA em Inteligência Artificial e Big Data

Uso de CNN para detecção de COVID-19 em raio-x de pulmão

IESB
2019-2023

Bacharelado em Ciência de Dados e Inteligência Artificial

Aplicação de técnicas de reconhecimento de entidades nomeadas em textos jurídicos: um estudo de caso

CURSOS E CERTIFICAÇÕES

AWS

AWS AI Practitioner

Certificação técnica que valida o domínio em conceitos de IA Generativa e Machine Learning. Capacitado para identificar casos de uso de IA e utilizar serviços da AWS como Amazon Bedrock e SageMaker para construir soluções.

**DeepLearning.
AI**

Functions, Tools and Agents with LangChain

Especialização no uso de OpenAI Function Calling e ferramentas do LangChain para transformar LLMs em agentes operacionais.

LangChain

Introduction to LangGraph

LangChain

Introduction to LangSmith

**Data Science
Academy**

Formação Arquiteto de Dados 4.0

Experiência prática com modelagem de dados (SQL e NoSQL), processamento de fluxos em tempo real, pipelines de ETL/ELT e orquestração. Conhecimento em tecnologias como Spark, Kafka, Hadoop, além de arquiteturas em nuvem e governança de dados

Coursera

Certificado Profissional de Análise de Dados do Google

Processo completo de análise de dados utilizando BigQuery, R, SQL e Sheets

**Palantir
Technologies**

Speedrun: Data Science Fundamentals

Capacitação prática em ciência de dados aplicada, utilizando o Palantir Foundry

**Palantir
Technologies**

Foundry & AIP Builder Foundations

Certificação técnica que valida o domínio em conceitos de IA Generativa e Machine Learning. Capacitado para identificar casos de uso de IA, selecionar modelos de base (FMs) e utilizar serviços da AWS como Amazon Bedrock e SageMaker para construir soluções.