

Leonardo Borges Silva Braga

Ciência de Dados e IA | Oracle Certified Generative AI & Data Science Professional

(61) 9 9294-4998

leoborgesprofissional@gmail.com

Brasília, DF

[linkedin.com/in/leonardo-borges1/](https://www.linkedin.com/in/leonardo-borges1/)

github.com/Leo-bsb

<https://leonardo-braga.dev>

Resumo profissional:

Cientista de Dados formado em Ciência de Dados e IA, com atuação prática em Engenharia de Dados e um interesse central em Governança, Segurança e Ética no tratamento de dados, alinhado à LGPD. Domino SQL, Python e ferramentas de integração e analytics (SAP/Pentaho, SAS), com histórico de otimização de pipelines e melhoria na qualidade de dados em ambientes enterprise. Criei o Datalock Studio, uma biblioteca e plataforma de anonimização de dados com k-anonimato, pseudonimização e criptografia, publicada e em processo de submissão na Microsoft Store, além de projetos de Machine Learning como sistemas de recomendação híbridos. Busco uma oportunidade para aplicar essas competências no desenvolvimento de sistemas de dados robustos, seguros e eticamente responsáveis.

Experiências:

- **Estagiário em Ciência de Dados | IZE Gestão Empresarial (01/2026 - 05/2026) - Híbrido**
 - **Dashboards e Integrações de sistemas:** Desenvolvi dashboards e KPIs atualizados em tempo real, utilizando Streamlit, Airflow e EvolutionAPI para integração com WhatsApp.
 - **Refatoração de Codebase:** Auditei a qualidade dos programas internos desenvolvidos e corriji bugs e más práticas.
 - **Alimentação de base de dados:** Desenvolvi automações fluxos de alimentação em banco de dados e em planilhas com Apps Script.
- **Estagiário em Data Migration | First Decision (04/2025 - 01/2026) - Híbrido**
 - **Engenharia de Dados e Automação:** Atuei no desenvolvimento e sustentação de pipelines de ETL complexos, garantindo a integridade e higienização de grandes volumes de dados para clientes enterprise (Natura, Neugebauer).
 - **Eficiência Operacional:** Otimizei fluxos de carga e transformação de dados, resultando em uma entrega 40% antecipada em relação ao cronograma oficial do projeto.
 - **Qualidade de Dados:** Implementei rotinas de validação e tratamento de erros (Data Quality), reduzindo inconsistências e assegurando a confiabilidade dos relatórios finais.
- **Estagiário e membro consultor - Empresa Júnior DataÍ Tecnologia EJ (07/2024 - 04/2025) - Remoto**
 - Desenvolvi dashboards e análises interativas para clientes, utilizando SAS, Python, SQL e Power BI.
 - Estruturei bancos de dados SQL para facilitar acesso e integração de dados.

Tecnologias e Conhecimentos:

- Linguagens: Python, SQL, SAS.
- Engenharia de Dados e ETL: Pentaho, SAP Data Services.
- Pipeline e Deploy: Docker, CI/CD, Airflow.
- Banco de Dados: MySQL e PostgreSQL.
- Cloud e Infraestrutura: Oracle Cloud Infrastructure (OCI)
- Governança e Privacidade de Dados: (LGPD, K-anonimato, pseudonimização, criptografia, etc.)

Certificações:

Oracle Cloud Infrastructure (OCI) – 2025

- Professional: **Generative AI, Data Science**
- Foundations: **OCI AI, Data Platform, Cloud Infrastructure**

Databricks – 2025

- Accreditations: **Fundamentals, Generative AI Fundamentals**

Projetos:

Datalock Studio: Biblioteca Python de Anonimização e tratamento de dados + Plataforma No-Code (LGPD)

- **Biblioteca:** projetei e implementei uma biblioteca Python focada em privacidade de dados, tratamentos e segurança informacional com mais de 30 funções públicas, hash determinístico, pseudonimização reversível, cálculo de k-anonimato e risco de reidentificação, geração de dados sintéticos e um formato de arquivo binário próprio com criptografia autenticada e verificação de integridade.
- **Plataforma (Site e Software):** construí do zero, uma interface no-code full-stack com 14 tipos de transformação de dados, arquitetura de motor duplo (JavaScript no navegador / Python real no backend), integração com bancos de dados relacionais e automação.
- **Distribuição:** empacotei como aplicativo nativo para Windows, em processo de submissão na Microsoft Store, além da versão web e executável para instalação publicado em produção.
Python | Polars | FastApi | SQLAlchemy | cryptography

Veja o projeto: datalock-studio.tech / <https://github.com/py-datalock/datalock>

- **Sistema Híbrido personalizado de Recomendação (Deep Learning & Filtragem Colaborativa)**
 - Sistema híbrido combinando filtragem colaborativa e redes neurais para recomendações com **80% de diversidade**.
 - Desenvolvimento focado em personalização precisa e descoberta de conteúdos menos populares.Fatoração de Matriz | Deep Learning | PyTorch | Gradio | Polars

Veja o projeto

Educação:

Bacharelado em Ciência de Dados e Inteligência Artificial

IESB | 02/2023 - 10/2026

Disciplinas relevantes:

Aprendizagem de Máquina, Mineração de Dados, Estatística Aplicada, Análise Exploratória e Visualização, Modelagem e Inferência Estatística, Processamento de Dados Massivos.

Artigos Técnicos (Medium):

- [Como Construí um Sistema de Predição de Tempo de Entrega que Atinge 82% de Acurácia.](#)
- [Como Construí um Sistema de Recomendação Híbrido que Descobre Livros Escondidos Combinando filtragem colaborativa, análise de conteúdo e redes neurais para recomendações personalizadas com 80% de diversidade.](#)