

LUCAS CRUZ

Engenheiro de Dados | Engenheiro de IA Generativa

+55 (21) 96552-9308 • dev.lucasacruz@gmail.com • linkedin.com/in/lucas-cruz • devzurc.github.io/portfolio

Curitiba, PR • Disponível para trabalho presencial, híbrido ou remoto

RESUMO PROFISSIONAL

Engenheiro de Dados e IA Generativa com mais de 5 anos de experiência construindo pipelines de dados de ponta a ponta, lakehouses em nuvem e soluções de automação com IA. Atuação comprovada em plataformas IoT, fintechs e agronegócio utilizando AWS, Azure e GCP. Especialista em LLMs, workflows de agentes com n8n e IA conversacional. Inglês fluente (C1) e perfil orientado a resultados e impacto de negócio.

COMPETENCIAS TÉCNICAS

IA Generativa & LLMs	Claude AI, Google AI Studio, Integração com LLMs, Agentes de IA, RAG, Engenharia de Prompt, Chatbots
Engenharia de Dados	Python, SQL, Apache Spark (PySpark), Pandas, dbt, Trino, REST APIs, Web Scraping, Parquet
Orquestração de Pipelines	Apache Airflow, n8n
Plataformas de Nuvem	AWS (S3, EC2, ECS, Lambda, IAM), Microsoft Azure, Google Cloud (GCP), Oracle Cloud (OCI)
Plataformas de Dados	Snowflake, Databricks, Amazon Redshift, Delta Lake, Oracle Autonomous Database
Bancos de Dados Relacionais	PostgreSQL, SQL Server, MySQL
Analytics & BI	Power BI, Qlik Sense, Apache Superset, Looker
DevOps & Infraestrutura	Docker, Kubernetes, Git, Pipelines CI/CD, Linux, Bash Scripting
Idiomas	Portugues (Nativo) • Ingles (C1 - Avancado) • Espanhol (B1 - Intermediario)

EXPERIÊNCIA PROFISSIONAL

Engenheiro de IA Generativa - TK Technologies | Curitiba, PR (Presencial · Contrato 3 meses) Fev 2026 - Mai 2026

- Desenvolveu pipeline de web scraping em Python coletando dados de preço e metadados de mais de 10 sites de peças automotivas, gerando inteligência competitiva de mercado em escala.
- Projetou workflows de automação com IA no n8n (ETL, RPA e agentes chatbot) integrando LLMs (Claude AI, Google AI Studio) e APIs do GCP (Sheets, Drive) para enriquecer dados e aumentar a produtividade do time.
- Desenvolveu chatbot de IA conversacional permitindo que usuários de negócio consultassem e refinaram datasets de preço por linguagem natural, reduzindo significativamente o tempo de obtenção de insights.
- Realizou deploy de serviços backend containerizados no Azure com Docker, padronizando ambientes e melhorando a confiabilidade das implantações em ~25%.

Engenheiro de Dados - Spacecom Monitoramento | Curitiba, PR (Presencial) Abr 2023 - Jul 2024

- Projetou e manteve um Data Lake House de Big Data na AWS S3 com Arquitetura Medallion (Bronze/Prata/Ouro) utilizando Delta Lake, Parquet, PySpark e Trino, suportando cargas de trabalho analíticas e de reportes.
- Suportou plataforma IoT de grande escala processando dados de mais de 95.000 dispositivos, com centenas de milhões de registros por dia em alta disponibilidade.
- Orquestrou e otimizou pipelines ETL com Apache Airflow, reduzindo falhas em ~30%; realizou deploy de serviços com Docker na AWS EC2, diminuindo problemas de implantação em ~25%.
- Identificou e corrigiu etapas inconsistentes nos pipelines, aumentando a acurácia dos dados em ~40%; entregou datasets consumidos por dashboards no Qlik Sense e Apache Superset.

Consultor de Engenharia de Dados - wDiscover | Curitiba, PR (Remoto) Jun 2022 - Mai 2024

- Entregou projetos completos de dados para clientes de fintechs, agronegócio e operações, integrando provedores de pagamento (PagSeguro, Stone, Cielo), API IoT John Deere, REST APIs e web scraping.
- Projetou Data Warehouses no SQL Server e Oracle Autonomous Database com modelagem dimensional; automatizou ETL com Airflow em Docker/Oracle Cloud, reduzindo esforço manual em ~35%.
- Construiu dashboards no Power BI para monitoramento de KPIs, métricas financeiras e desempenho operacional, habilitando decisões orientadas a dados para stakeholders.

Analista de Dados - Grupo Itaeté | Curitiba, PR (Híbrido) Jun 2020 - Fev 2023

- Projetou Data Warehouse no SQL Server e pipelines Python/SQL com Pandas para limpeza, transformação e enriquecimento de dados, integrando dispositivos IoT, PostgreSQL, APIs e arquivos CSV.
- Criou dashboards no Power BI e workflows automatizados de relatórios em Excel/VBA, reduzindo o tempo de processamento de dados em ~50% e o volume de trabalho manual operacional.

- Desenvolveu sistemas internos de controle de estoque e automação financeira, incluindo geração de notas fiscais e distribuição automática de relatórios semanais.

PROJETOS DE DESTAQUE

Plataforma de Inteligência de Preços - Mercado Automotivo · Python, n8n, LLMs, GCP APIs, Azure, Docker · *TK Technologies - Fev 2026*

- Coletou dados de mais de 10 sites automotivos com Python; orquestrou agentes de IA via n8n com Claude AI, Google AI Studio e GCP (Sheets, Drive) para automatizar análise de preços e gerar insights de compras competitivas. Deploy completo containerizado no Azure.

Pipeline de Dados em Nuvem com Stack Moderna · Snowflake, dbt, Apache Airflow, Docker · *Projeto Pessoal - Jan 2026*

- Pipeline analítico de ponta a ponta com Airflow, Snowflake e dbt; containerizou o dbt com Docker reduzindo erros de deploy em ~50%; transformações SQL modulares e testadas orquestradas por DAGs do Airflow.

Data Lakehouse IoT na AWS - Arquitetura Medallion · AWS S3, Apache Spark, Delta Lake, Airflow, Trino, Apache Superset · *Projeto de Cliente - Fev 2024*

- Projetou lakehouse Bronze/Prata/Ouro na AWS S3 com ingestão via APIs, bancos de dados e CSV; virtualizou a camada Ouro com Trino; entregou insights via dashboards no Apache Superset.

FORMACAO ACADEMICA

Tecnologo em Analise e Desenvolvimento de Sistemas - Universidade Dom Bosco, Curitiba, PR - Jun 2023

Ênfase em: Sistemas de Informacao, Business Analytics, Analise de Dados, Banco de Dados, Engenharia de Software