

GUILHERME FERREIRA

Staff Data Platform Engineer

Databricks | AWS | Spark/PySpark | Kafka | dbt | Terraform

Sao Paulo, Brazil | +55 11 99645-1902 | guilherme_s.ferreira@outlook.com

guilhermeferreira.dev | github.com/guilhermesferr | linkedin.com/in/guilherme-ferreira-844052192

PROFESSIONAL PROFILE

Staff Data Platform Engineer with end-to-end ownership of data platforms, from architecture and ingestion to production operations. Technical lead for Martelo's Databricks platform built from scratch, with experience owning Snowflake and AWS data platforms, batch and streaming pipelines, and controlled cloud migrations in financial services. Focused on platform reliability, governance, reusable data products, and cost and performance engineering.

CORE TECHNOLOGIES

Databricks, AWS, Spark/PySpark, Kafka, dbt, Terraform, Snowflake, Delta Lake, Airflow, AWS DMS, Python, Docker, Kubernetes, CI/CD.

PROFESSIONAL EXPERIENCE

Staff Data Platform Engineer | Martelo

Oct 2025 - Present

Remote | Florida, United States

- Led the technical architecture and built the company's Databricks data platform from scratch, owning implementation and ongoing production operations.
- Owned ingestion, batch and streaming pipelines, the analytics warehouse, data-quality controls, observability, and data delivery for analytics and production ML.
- Managed governance, user access, AWS integration, infrastructure operations, and recovery, establishing reusable platform standards and clear ownership across teams.
- Right-sized compute and memory in Databricks and AWS, balancing workload performance, reliability, scaling needs, and infrastructure costs.
- Reduced data-quality incidents by 40% and met a 99.9% model-serving SLO through quality gates, monitoring, and controlled delivery and recovery practices.

Staff Data Platform Engineer | Insi (formerly Meta)

Jan 2025 - Oct 2025

Fiserv client engagement | Hybrid | Brazil

- Owned the Fiserv data platform across Snowflake, AWS, GCP, and on-premises systems, developing batch and streaming pipelines with PySpark, Kafka, and dbt.
- Removed ingestion and processing bottlenecks to improve real-time processing speed by 33% and batch processing speed by 26%.
- Delivered a controlled AWS migration for 40% of high-volume financial-operations data. Implemented and operated AWS DMS for the migration of part of Fiserv's card transaction data from on-premises infrastructure to the cloud, with reconciliation and dual runs supporting validation.
- Implemented and managed infrastructure as code with Terraform, supporting repeatable environments, reviewed changes, and ongoing platform operations.
- Separately optimized workloads retained on premises under PCI DSS constraints, improving processing speed by 17%. Owned orchestration, data quality, observability, and production recovery across the engagement.

PROFESSIONAL EXPERIENCE CONTINUED

Staff Data & ML Engineer | Plural Solutions

Nov 2023 - Dec 2024

Hybrid | Brazil

- Built and owned the company's data platform and data solutions, establishing the foundations for analytics and production ML.
- Designed ingestion, data preparation, and transformation workflows with Python, Spark/PySpark, and AWS, creating reusable datasets for downstream applications and analytics.
- Managed data workflows, quality controls, monitoring, and operational recovery, working with engineering, data science, product, and business teams on reliable data delivery.
- Developed the data and feature pipelines supporting ML workflows, using AWS, SageMaker, and MLflow for reproducible delivery, deployment, and monitoring.

Professional development | Apr 2023 - Nov 2023

Focused study in data-platform architecture, cloud and hybrid systems, batch and streaming design, orchestration, governance, observability, and production delivery.

Senior Data Engineer | Mercado Libre / Mercado Pago

Jul 2021 - Mar 2023

Remote | Sao Paulo, Brazil

- Built and operated Snowflake and PySpark data products and Kafka streaming pipelines for Mercado Pago workloads, including PIX, reconciliation, CDB, executive analytics, and regulatory reporting.
- Owned ingestion from on-premises databases, cloud storage, APIs, and streaming sources, including near-real-time delivery with Snowpipe and Kafka.
- Reduced data-quality incidents by 22% and cut regulatory-report production time by 25% by strengthening validation, observability, recovery, and the reporting workflow.

Data Analyst | Itau Unibanco

Oct 2020 - Jul 2021

On-site | Sao Paulo, Brazil

- Built ETL and analytical workflows to identify and correct customer-data inconsistencies, improving the reliability of banking datasets.
- Analyzed Snowflake query plans and ingestion patterns to identify bottlenecks and provide actionable performance recommendations to engineering and business stakeholders.

Data/Analytics Intern | Itau Unibanco

Nov 2019 - Oct 2020

On-site | Sao Paulo, Brazil

- Supported data preparation, ETL development, and workflow automation for banking analytics and data-engineering delivery.
- Worked with senior engineers on Snowflake performance analysis and reliable analytical datasets, developing a foundation in Python, data quality, and production data workflows.

PLATFORM EXPERTISE

Platform engineering: Terraform, Python, Docker, Kubernetes, CI/CD; infrastructure as code, cloud and hybrid operations, compute sizing, and cost optimization.

Architecture and reliability: Lakehouse and warehouse design, data products and contracts, governance, user access, lineage, data quality, observability, SLIs/SLOs, and recovery.

Additional experience: GCP, Flink, Trino, DataHub, OpenMetadata; production ML data pipelines, MLflow, and SageMaker.

EDUCATION AND SELECTED CREDENTIALS

B.Sc. Computer Engineering | Mackenzie Presbyterian University | 2021

Udacity: Data Architect; Data Engineering with AWS; Data Streaming; Data Product Manager. Additional training in production ML and MLOps through Udacity and DeepLearning.AI.