

PROFILE

Mathematics graduate specialized in Data Engineering, with hands-on experience building streaming pipelines (Kafka, Flink), Lakehouse architectures (Iceberg, Paimon, MinIO), orchestrated ETL workflows (Airflow, dbt), and analytical data modeling. Additional background in Machine Learning and Applied AI (TensorFlow, RAG systems, custom neural networks).

PROJECTS

- Real-Time Streaming Pipeline for Transaction Risk Analysis

Flink, Kafka, Paimon

 - Designed an event-driven Apache Flink 2.2 + Kafka pipeline for real-time transaction scoring and risk analysis.
 - Built an Apache Paimon Lakehouse on MinIO for continuous stream ingestion, integrated with Prometheus monitoring.
 - Integrated local Ollama LLMs to enrich flagged transaction events with dynamic contextual risk insights.
- Lakehouse & RAG Recommendation System

Spark, Airflow, Iceberg

 - Architected an Apache Iceberg Lakehouse (Airflow-orchestrated) for automated ETL batch data processing.
 - Implemented a collaborative filtering model using PySpark MLlib's Alternating Least Squares (ALS) algorithm.
 - Developed a Streamlit frontend with a RAG pipeline and pgvector similarity search for semantic queries.
- Analytics Engineering Pipeline (Airbnb Data)

dbt, Athena, Evidence

 - Engineered dbt transformation models on AWS Athena to convert raw data into staging and dimensional data marts.
 - Automated schema testing and data freshness assertions within dbt to maintain high data quality standards.
 - Built code-driven BI reports and interactive dashboards using Evidence for automated analytical reporting.
- ML & Neuro-Fuzzy Modeling

TensorFlow, Keras, Rust

 - Trained a Seq2Seq deep learning model with Bahdanau Attention mechanism for date text sequence normalization.
 - Implemented an Adaptive Neuro-Fuzzy Inference System (ANFIS) with Wang-Mendel automated rule extraction.
 - Rewrote the Wang-Mendel algorithm in Rust for high-performance inference; published as `wm_fuzzy` on `crates.io`.
- E-Commerce Platform on Microservices Architecture

Amazon EKS, Argo CD, GitHub Actions

 - Architected a scalable e-commerce platform composed of independent microservices deployed on Amazon EKS.
 - Configured traffic routing with AWS Load Balancer Controller and DNS management via Route 53.
 - Implemented a GitOps CI/CD pipeline using GitHub Actions for automated builds and Argo CD for progressive delivery.

TECHNICAL SKILLS

Data Engineering: Apache Spark, Kafka, Flink, Airflow, dbt, Iceberg, Paimon, MinIO, DuckDB, Athena
Databases: PostgreSQL, pgvector, SQL
Machine Learning: TensorFlow, Keras, Scikit-Learn, Neural Networks, Attention Mechanisms, NLP, RAG
Infrastructure & Tools: Linux, Docker, Docker Compose, Prometheus, Git, Amazon EKS, Argo CD.
Programming: Python, Julia, SQL, Bash, $\LaTeX$, Rust

EDUCATION

University of Campinas (UNICAMP)

M.Sc. in Applied Mathematics

Federal University of São Carlos (UFSCar)

B.Sc. in Mathematics

Campinas, Brazil

August 2024 – September 2026

São Carlos, Brazil

March 2020 – August 2024

CERTIFICATIONS

AWS Agent AI Practitioner

Amazon Web Services (AWS)

2026

AWS Certified Cloud Practitioner

Amazon Web Services (AWS)

2026

LANGUAGES

English: Advanced (Professional Working Proficiency)

French: Intermediary