MATEUS CARNIEL BRAMBILLA
Data Engineer | Databricks | Applied AI
Brazil | Remote | English B2
mateuscarnielbrambilla@gmail.com | +55 54 99915-6120
LinkedIn | GitHub | Credly

PROFESSIONAL PROFILE

Data Engineer with production experience in the Databricks ecosystem, working daily with PySpark, SQL, Delta Lake, Unity Catalog, and Asset Bundles. Builds and maintains batch pipelines, platform automation, and applied AI solutions. Has worked across 20+ Databricks jobs, migrated an approximately 20-notebook pipeline to Lakeflow Pipelines/DLT reducing runtime from about 30 to 8 minutes, and reduced LLM output cost by 75% in a production pipeline.

ACADEMIC BACKGROUND

- **Bachelor's Degree in Software Engineering — Univates** | 2025 – Expected Completion: 12/2028

CERTIFICATIONS AND COMPLEMENTARY TRAINING

- Google Cloud - Associate Cloud Engineer;
- Astronomer — Apache Airflow 3 Fundamentals (2026);
- Cisco Networking Academy — CCNA 1 & CCNA 2;
- Linux Fundamentals (Red Hat Training);
- Advanced SQL (HackerRank).

PROFESSIONAL EXPERIENCE

Tramontina Multi S.A. — Developer (Data Engineering & Backend) | Nov/2025 – Present

- **Data Engineering & Databricks:** Built and maintained 20+ batch jobs with PySpark, SQL, Delta Lake, and Unity Catalog, contributing to the evolution of production pipelines and tables in Databricks.
- **Lakeflow Pipelines / DLT:** Independently migrated an approximately 20-notebook pipeline to serverless Lakeflow Pipelines/DLT, reducing runtime from about 30 to 8 minutes.
- **Platform Tooling & GitOps:** Built a Python library with the Databricks SDK to export Genie Spaces as human-readable, parameterized metadata, normalize generated identifiers, and version/promote configurations across workspaces through GitLab.
- **AI Governance:** Created approximately 8 Metric Views in Unity Catalog to constrain Genie Spaces to pre-defined metrics and dimensions, providing standardized queries and governed context for agents.
- **Applied AI:** Reduced LLM output cost by 75% in a cataloging pipeline by optimizing model context/metadata and moving to a more efficient model; later implemented a primary/fallback strategy through Databricks AI Gateway

with exponential backoff and jitter.

Tramontina Multi S.A. — Industrial Processes Trainee | Apr/2025 – Nov/2025

- **Data Engineering Foundations:** Deepened practical experience with PySpark, Databricks, data analysis, and Git while working with notebooks and Bronze/Silver architectures during the transition into the current role.
- **Data Science:** Developed time-series PoCs with Prophet and LSTM for demand forecasting.
- **Engineering Practices:** Gained experience with version control, code review, and CI/CD workflows based on the GitLab templates used by the team.

FEATURED TECHNICAL PROJECTS

- **WaterScript — EY AI & Data Challenge 2026:** Built a multimodal data pipeline with Python and Snowflake/Snowpark, integrating Landsat imagery through STAC and TerraClimate data, with geospatial feature engineering and XGBoost/Optuna modeling using GroupKFold validation.
- **Multi-Agent Orchestrator (PoC):** Built a supervisor for routing requests between specialized Genie Spaces, exploring domain separation and agent composition within the Databricks ecosystem.

TECHNICAL SKILLS

- **Data Engineering:** Python, PySpark, SQL, Delta Lake, Databricks (Lakeflow Pipelines/DLT, Unity Catalog, Metric Views, SQL Warehouse), Apache Airflow 3, Medallion Architecture, Dimensional Modeling.
- **Applied AI & Platform:** Databricks Genie, Databricks SDK, MLflow, LangChain, LLM APIs, Databricks AI Gateway, GitOps.
- **Backend & Databases:** Flask, FastAPI, Pydantic, SQLAlchemy, PostgreSQL, SQL Server, MongoDB.
- **DevOps & Tooling:** Databricks Asset Bundles, Jobs/Workflows, Git/GitLab, Docker, Linux; Google Cloud Associate Cloud Engineer certified.

AWARDS AND RECOGNITIONS

- **Global Nominee — NASA Space Apps Challenge (2025):** Winner of the local stage (Lajeado) and global nominee for the AirMood project (satellite data processing for public health).