MATEUS CARNIEL BRAMBILLA
Data Engineer | Databricks | Applied AI
Rio Grande do Sul, Brasil | Remoto | Inglês B2
mateuscarnielbrambilla@gmail.com | +55 54 99915-6120
LinkedIn | GitHub | Credly

PERFIL PROFISSIONAL

Engenheiro de Dados com experiência de produção no ecossistema Databricks, atuando diariamente com PySpark, SQL, Delta Lake, Unity Catalog e Asset Bundles. Desenvolve e mantém pipelines batch, automações de plataforma e soluções de IA aplicada. Já atuou em mais de 20 jobs Databricks, migrou um pipeline de aproximadamente 20 notebooks para Lakeflow Pipelines/DLT reduzindo o runtime de cerca de 30 para 8 minutos e reduziu em 75% o custo de output de um pipeline baseado em LLM.

FORMAÇÃO ACADÊMICA

- **Bacharelado em Engenharia de Software — Univates** | 2025 – Previsão de Conclusão: 12/2028

CERTIFICAÇÕES E FORMAÇÃO COMPLEMENTAR

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

EXPERIÊNCIA PROFISSIONAL

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

- **Data Engineering & Databricks:** Construí e mantive mais de 20 jobs batch com PySpark, SQL, Delta Lake e Unity Catalog, atuando na evolução de pipelines e tabelas em ambientes Databricks.
- **Lakeflow Pipelines / DLT:** Migrei integralmente um pipeline de aproximadamente 20 notebooks para Lakeflow Pipelines/DLT em serverless, reduzindo o runtime de cerca de 30 para 8 minutos.
- **Platform Tooling & GitOps:** Desenvolvi uma biblioteca Python com Databricks SDK para exportar Genie Spaces como metadados legíveis e parametrizáveis, normalizar identificadores gerados automaticamente e versionar/promover configurações entre workspaces via GitLab.
- **Governança para IA:** Criei cerca de 8 Metric Views no Unity Catalog para restringir o contexto de Genie Spaces a métricas e dimensões pré-definidas, fornecendo consultas padronizadas e contexto governado aos agentes.
- **Applied AI:** Reduzi em 75% o custo de output de um pipeline de catalogação ao otimizar os metadados enviados ao modelo e migrar para uma alternativa

mais eficiente; posteriormente implementei estratégia de modelo primário/fallback no Databricks AI Gateway com exponential backoff e jitter.

Tramontina Multi S.A. — Trainee de Processos Industriais | Abr/2025 – Nov/2025

- **Formação em Engenharia de Dados:** Aprofundei PySpark, Databricks, análise de dados e Git, trabalhando com notebooks e arquiteturas Bronze/Silver durante a transição para a função atual.
- **Data Science:** Desenvolvi PoCs de séries temporais com Prophet e LSTM para previsão de demanda.
- **Práticas de Engenharia:** Ganhei experiência com versionamento, code review e fluxos de CI/CD baseados nos templates GitLab utilizados pela equipe.

PROJETOS TÉCNICOS EM DESTAQUE

- **WaterScript — EY AI & Data Challenge 2026:** Pipeline multimodal com Python e Snowflake/Snowpark, integrando imagens Landsat via STAC e dados TerraClimate; inclui feature engineering geoespacial e modelagem com XGBoost/Optuna e validação GroupKFold.
- **Orquestrador Multi-Agente (PoC):** Desenvolvi um supervisor para roteamento de consultas entre Genie Spaces especialistas, explorando separação de domínios e composição de agentes no ecossistema Databricks.

HABILIDADES TÉCNICAS

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

PRÊMIOS E RECONHECIMENTOS

- **Global Nominee — NASA Space Apps Challenge (2025):** Vencedor da etapa local (Lajeado) e indicado global pelo projeto AirMood (processamento de dados de satélite para saúde pública).