Guilherme Yuji Marcolongo Adania

Engenharia de Dados | Análise de Dados | BI | Automação
São Paulo, Brasil | +55 (11) 99334-5225 | gyuji1904@gmail.com | linkedin.com/in/guilherme-yuji | github.com/GuiYuuji

RESUMO PROFISSIONAL

Profissional de Dados e BI com experiência prática em engenharia de dados, qualidade de dados, automação em Python, SQL, APIs REST, Big Data e governança de dados em ambientes corporativos. Experiência em projetar e entregar produtos de dados internos de ponta a ponta — de integrações via API e motores de validação a sistemas automatizados de monitoramento e soluções de reporting — além de construir sistemas full-stack para clientes externos como desenvolvedor independente. Perfil analítico com mentalidade prática de Engenharia de Dados, focado em soluções escaláveis, precisas e sustentáveis.

EXPERIÊNCIA PROFISSIONAL

Engineering Brasil — Analista de Dados Jr | Projeto Vivo | Mar 2026 – Atual

- Lidero a iniciativa da Engineering Brasil de criação de agentes de IA e soluções de automação para a Vivo, utilizando Python, LangGraph e arquiteturas multi-agente para automatizar processos de negócio (iniciativa em estágio inicial).
- Atuo na frente de Curadoria de Dados do projeto Vivo, apoiando organização de dados, metadados, catalogação, governança e processos de qualidade.
- Projetei e construí um sistema de validação de briefings de metadados integrado ao Informatica EDC via APIs REST, com requisições HTTP paralelas e thread pooling. Valida ativos, bancos, tabelas, campos, descrições, nomes de negócio e demais atributos de metadados em uma única execução.
- **Impacto:** briefings acima de 2.000 linhas eram inviáveis ou levavam horas no processo legado; briefings acima de 15.000 linhas passaram a ser concluídos em minutos — cerca de 30x mais rápido em validações únicas e 85x em ciclos contínuos de validação.
- Construí um sistema automatizado de monitoramento cobrindo ~70 projetos multi-tabela, que detecta desvios de volumetria, atrasos de carga e quebras de SLA, substituindo uma checagem diária totalmente manual feita por dois analistas.
- **Impacto:** ~60 incidentes detectados nos primeiros 4 meses de operação; tempo diário de validação reduzido de 3–4 horas para menos de 1 hora (~70%).
- Expandi o sistema de monitoramento com validação visual automatizada de dashboards para capturar falhas silenciosas de atualização — dashboards que atualizam sem erro, mas renderizam em branco ou corrompidos. Roda em horários fixos ao longo do dia, exporta o estado atual do dashboard em PDF, converte para PNG e compara com uma imagem de referência usando OpenCV para sinalizar anomalias. Criado após um dashboard quebrado reportado pelo cliente; desde então já sinalizou 2 falhas reais.
- Realizo análise, validação, monitoramento e troubleshooting de Qualidade de Dados em processos corporativos, investigando inconsistências em registros, metadados e documentação técnica.
- Atuo com catalogação, rastreabilidade, taxonomia e governança de dados para melhorar a organização e a confiabilidade dos ativos de dados.
- Colaboro com stakeholders técnicos e de negócio para padronizar processos, documentar requisitos e melhorar a qualidade de dados entre equipes.
- Trabalho com arquitetura medalhão (bronze / prata / ouro) e organização de domínios de dados na plataforma do projeto.
- Opero cargas de trabalho de dados em máquinas virtuais GCP no ambiente corporativo.
- **Stack corporativo:** Python, SQL, Databricks, PySpark, Hive, Teradata, DBeaver, Power BI, Informatica EDC, APIs REST, GCP, Azure DevOps.

Cellere Soluções em TIC — Analista de Dados | Projeto Vivo | Mai 2025 – Dez 2025

- Desenvolvi automações em Python que reduziram trabalho manual recorrente dos analistas e aumentaram a eficiência operacional.
- Criei e mantive sistemas internos e integrações de dados entre plataformas por meio de APIs REST e fluxos automatizados.
- Realizei análise e curadoria de dados em ambientes Big Data, validando tabelas, campos, estruturas e metadados em escala.
- Escrevi consultas SQL em ambientes distribuídos Hive/Tez sobre tabelas particionadas para validação, reconciliação e análise cruzada de fontes.
- Revisei briefings técnicos de dados e traduzi requisitos entre times técnicos e de negócio, apoiando governança e documentação de dados.

Desenvolvedor Independente — Desenvolvimento de Software Freelance | 2022 – Atual

- Entreguei sistemas sob medida de ponta a ponta para aproximadamente 15 clientes de pequeno e médio porte, incluindo oficinas automotivas, óticas e empresas de serviços.
- Conduzi o ciclo completo: levantamento de requisitos com clientes não técnicos, design do sistema, desenvolvimento, deploy e suporte pós-entrega.
- Construí sistemas web e ferramentas internas de gestão com PHP, Next.js, React e TypeScript, containerizados com Docker e publicados em ambientes VPS.

COMPETÊNCIAS TÉCNICAS

Programação & Automação: Python, SQL, JavaScript, TypeScript, PHP, APIs REST, scripting, automação de processos, processamento paralelo / thread pooling, Selenium, OpenCV, openpyxl

Engenharia de Dados: ETL/ELT, pipelines de dados, transformação de dados, integração de dados, qualidade de dados, QA/QC, arquitetura medalhão, modelagem de dados, particionamento, troubleshooting, análise de causa raiz

Plataformas & Processamento: Databricks, PySpark, pandas, Hive/Tez, Teradata, Big Data, GCP, Informatica EDC, PostgreSQL, Supabase

Formatos de Dados: Parquet, JSON, CSV, XML

BI & Reporting: Power BI, desenvolvimento de dashboards, visualização de dados, rotinas recorrentes de reporting

IA & Agentes: LangGraph, arquiteturas multi-agente, APIs OpenAI, automação baseada em LLM

Web & Frameworks: Next.js, React, Flask, FastAPI, Streamlit, HTML/CSS

DevOps & Ferramentas: Docker, Git, Azure DevOps, deploy em VPS, DBeaver, Figma

Governança de Dados: Qualidade de dados, gestão de metadados, catalogação, taxonomia, rastreabilidade, governança, documentação técnica

FORMAÇÃO ACADÊMICA

Bacharelado em Análise e Desenvolvimento de Sistemas — Universidade Presbiteriana Mackenzie | Fev 2026 – Dez 2027

Técnico em Desenvolvimento de Sistemas — ETEC | Fev 2020 – Dez 2022

Anteriormente: Análise e Desenvolvimento de Sistemas — FATEC (Fev 2023 – Dez 2025, transferência para o Mackenzie)

CERTIFICAÇÕES & CURSOS

CS50x: Introduction to Computer Science — Harvard University | Mar 2026

Data Engineering Foundations Professional Certificate — Astronomer | Mar 2026

OpenAI e Python: APIs e Chatbots — Alura | Nov 2025

Engenharia de Dados com Databricks — Alura | Jul 2025

Data Science — Alura | Jun 2025

Machine Learning na Prática — Alura | Mai 2025

IDIOMAS

Português — Nativo | Inglês — C2 (EF SET)