Thiago Bellotto Rosa

Lead Data Engineer

California, US · thiago.br@live.com · thiagobellotto.com · linkedin.com/in/thiago-bellotto · github.com/thiagobellotto

Summary

Lead Data Engineer with **9 years** of experience, including about five in data engineering, mostly for US healthcare. Started in audit, M&A due diligence and data science at EY. Leads client-facing workstreams from discovery and architecture to production, building cloud data platforms and pipelines on **AWS**, **Azure**, **Google Cloud**, **Snowflake** and **Databricks**.

Experience

SoftServe · Big Data practice · Remote, US

Lead Data Engineer · Lead Software Engineer (L4) — Jan 2026 – Present

- Led architecture for a health-plan platform replacing manual, partner-dependent **Medicare Advantage risk adjustment**, from discovery through design and implementation planning.
- Designed a Databricks platform with clinical, retrieval, coding and operations domains, and secure partner data exchange on Azure (ADLS, SFTP, Delta Sharing).
- Defined workflow orchestration, domain data models and operational audit trails (**Kafka**, **FastAPI**, Kubernetes, **PostgreSQL**); coordinated the engineering workstream with stakeholders.
- Delivered a Snowflake clinical proof of concept: cross-account data sharing and replication, and extraction of audit evidence for coding QA.

Senior Data Engineer · Senior Software Engineer (L3) — Aug 2024 – Dec 2025

- Technical and workstream lead for a **Regrouping Engine** that consolidates chase requests by user-defined rules, cutting processing from several hours to **under a minute** for most workloads.
- Designed the database, infrastructure and central warehouse for a health-sector social app serving rare-disease patients in the US and Canada (Aurora PostgreSQL, **Redshift**, **dbt**).
- Built ML recommendation pipelines on **AWS Personalize** and headless BI on Metabase, deployed with ECS/Docker, **Terraform** and CodePipeline.
- Prototyped a Snowflake clinical platform with streams, tasks and external stages for event-driven processing of clinical documents.

Data Engineer · Software Engineer (L2) — Jun 2022 – Jul 2024

- Designed ingestion of **FHIR**, C-CDA, medical notes and PDFs into Snowflake, normalized to the **OMOP** model, with file-arrival notifications and secure views (S3, SNS, Flyway, Textract, Comprehend Medical).
- Created a dbt framework with reusable staging and BI layers from Azure Storage and SQL Server into Snowflake, deployed with Docker and Azure Scheduler.
- Engineered low-latency medical-record retrieval on **AWS Lambda**, Step Functions, **DynamoDB** and **OpenSearch**.
- Delivered data modeling, data-quality tooling and synthetic test data, with Terraform and GitHub Actions CI/CD.

027Capital · Strategic advisory and analytics firm · Remote, Brazil

Data Engineer — Aug 2021 – Jun 2022

- Designed and built the company's serverless Google Cloud platform from scratch (Cloud Run, Cloud Functions, VMs, Cloud Build CI/CD).
- Integrated client CRM/ERP systems, databases and files through **Airflow** pipelines to keep BI dashboards refreshed.
- Developed FastAPI services with input validation and SQL-injection/XSS protection; automated churn, MRR, LTV and cohort metrics.
- Wrote architecture proposals and estimates, and coached the team on Docker and cloud practices.

EY (Ernst & Young) · Brazil

Data Scientist · Business Modeling & Data Analytics Jan 2021 – Aug 2021

- Owned the practice's data workflows from multi-source ingestion and ETL through validation, modeling and presentation.
- Built a soccer-player database from web sources (**Python**, Selenium, BeautifulSoup) and modeled market value with scikit-learn and PyCaret; built **Power BI** dashboards.

Senior Consultant · Financial Due Diligence May 2019 – Dec 2020

- Financial due diligence for corporate and private-equity buyers in healthcare, education and agribusiness; promoted from Consultant to Senior.
- Reconciled accounting, billing, receivables and payer data for healthcare groups to assess payer concentration, denials, normalized **EBITDA**, net debt and working capital.
- Automated AP/AR aging and multi-file consolidation (Python, Power Query, Alteryx) adopted across engagements.

Assurance Consultant · Retail & Consumer Products Oct 2017 – May 2019

- External audits: walkthroughs, tests of controls, substantive testing and analytical reviews; on one of the first teams to adopt EY's **audit analytics** tools.

Skills

Cloud & platforms: AWS (Glue, DMS, Lambda, Step Functions, Redshift, Personalize), Azure (ADLS), Google Cloud, Snowflake, Databricks

Data: SQL, Python, dbt, Airflow, Kafka, Iceberg/Hudi, Delta Sharing, PostgreSQL, DuckDB, FHIR, OMOP

Engineering: Terraform, Pulumi, Docker, Kubernetes, FastAPI, GitHub Actions, CodePipeline

Analytics: Power BI, Metabase, scikit-learn, financial modeling

Education

Economics · Federal University of Pelotas (Brazil) 2013 – 2018