

Gilson Fonsaca

Senior Data Engineer | Data Architect | Multi-Cloud Specialist

Limeira – SP, Brazil

LINKEDIN

EMAIL

PROFESSIONAL SUMMARY

Experienced data professional with a strong background in **Data Engineering, Data Architecture, and Database Administration (DBA)**, working in high-volume, complex, and mission-critical environments. Proven track record in corporate data modeling across **conceptual, logical, and physical stages** for transactional (OLTP), analytical (OLAP), and multidimensional systems.

Specialized in **performance tuning** and **optimization of relational databases** (PostgreSQL, SQL Server, Oracle) in both on-premises and multi-cloud environments (AWS, GCP, Microsoft Azure). Adept at identifying and resolving complex performance bottlenecks, optimizing queries, partitioning large tables, and designing high-availability/disaster recovery strategies.

Extensive experience building end-to-end data pipelines focused on the delivery of robust **Data Products**, implementing rigorous **data quality, data governance, and metadata management** (data lineage and cataloging) strategies using tools like OpenMetadata and Azure Dataverse.

Proficient in designing advanced **semantic and ontological layers** (OWL/Turtle, SPARQL-over-SQL semantic virtualization, Apache Jena Fuseki). Highly skilled in building modern data pipelines (batch and streaming) using Python, Apache Airflow, Apache Spark, and cloud-native services (AWS Glue/EMR, GCP Composer/Dataproc, Azure ADF/Databricks), as well as integrating event-driven architectures with Kafka and RabbitMQ.

Expertise in modern cloud storage architectures (Data Lakehouse) utilizing **Trino, Dremio, Apache Iceberg, AWS S3 / S3 Tables, GCP Cloud Storage / BigQuery, and Azure ADLS Gen2 / Synapse**, ensuring architectural portability using Docker and containerized ecosystems.

TECHNOLOGIES & SKILLS

- **Data Engineering & Architecture:** Modern Data Lakehouse, Enterprise Data Warehouse (EDW), Conceptual, Logical, & Physical Corporate Data Modeling, Dimensional Modeling, ETL/ELT Pipelines, Data Products, Analytics Engineering.
- **Databases & Tuning:** PostgreSQL (Amazon RDS, Google Cloud SQL), SQL Server, Oracle Database (Exadata, RAC), MySQL, OLTP, OLAP, Query/SQL Tuning, Indexing, Partitioning, Replication, High Availability (HA) & Disaster Recovery (DR).

- **Cloud Platforms:**
 - **AWS:** S3, S3 Tables, Glue, MWAA, EMR, MSK, Amazon MQ, Kinesis, Athena, Redshift.
 - **GCP:** Cloud Storage (GCS), BigQuery, Dataplex, BigLake, Cloud Composer, Dataproc, Pub/Sub.
 - **Azure:** ADLS Gen2, Microsoft Fabric, Synapse Analytics, Databricks, Data Factory (ADF), Event Hubs, Service Bus.
- **Orchestration & Streaming:** Apache Airflow, Apache Spark, Apache Kafka, RabbitMQ, Apache Iceberg, Trino, Dremio.
- **Semantic Web & Knowledge Graphs:** Ontological Modeling, OWL (Web Ontology Language), SPARQL, SPARQL-over-SQL, Semantic Virtualization, Turtle Format (.ttl), Apache Jena Fuseki.
- **Languages & APIs:** Python, PL/SQL, PHP, REST APIs, Microservices, API Integration.
- **Governance, Observability & Quality:** Metadata Management, OpenMetadata, Azure Dataverse, Data Quality, Observability, Monitoring.
- **DevOps, Infrastructure & Virtualization:** Docker, Kubernetes, CI/CD, Infrastructure as Code (IaC), VMware.
- **AI & Analytics:** Machine Learning, Artificial Intelligence (AI), Data Science support, Data Analytics, Dashboarding.

PROFESSIONAL HISTORY

Database Specialist & Data Architect @ **RNP – National Education and Research Network**

July 2020 – Present * **Core Responsibilities:** * Design and implement modern data platform architectures across multi-cloud (AWS and GCP) and on-premises environments. * Architect, build, and orchestrate scalable data pipelines using **Apache Airflow (AWS MWAA, GCP Cloud Composer)** and **Apache Spark (AWS Glue/EMR, GCP Dataproc)**. * Modernize legacy analytical storage into **Data Lakehouse** architectures utilizing **Trino, Dremio, Apache Iceberg, AWS S3 / S3 Tables**, and **GCP Cloud Storage**. * Deliver corporate data models (conceptual, logical, and physical) for transactional and analytical systems. * Design and virtualize semantic metadata/knowledge graphs using OWL (.ttl) ontologies, SPARQL, and **Apache Jena Fuseki** servers. * Implement data governance, data catalogs, and metadata lineage tracking using **OpenMetadata**. * **Key Achievements & Projects:** * **Internet Brasil (Digital Inclusion Program):** Designed and deployed a scalable hybrid data architecture utilizing **PostgreSQL (RDS)** for transactional flows and **Google BigQuery** for high-performance large-scale analytics. Developed automated ELT pipelines for beneficiary data ingestion, optimizing query efficiency and reducing cloud operational costs. * **SISU (Unified Selection System):** Re-engineered the database backend and infrastructure configuration for a high-concurrency national application. Managed query optimization, index tuning, and partitioning on Oracle and SQL Server systems, ensuring stability and zero downtime during peak national registration access. * **Adapta Brasil (MCTI Climate Portal):** Engineered ingestion workflows and indicators database modeling, successfully unifying socio-environmental data from disparate scientific and governmental sources. * **Sinapse (MCTI):** Spearheaded semantic interoperability by virtualizing distributed relational databases into semantic layers, allowing complex federated queries across distinct government platforms using SPARQL. * **SIB (ANS**

Beneficiary System): Modernized and tuned bulk XML/JSON data exchange interfaces between healthcare providers and ANS databases, resolving systemic performance bottlenecks.

DBA SQL Server / Oracle / MySQL & Cloud Data Engineer @ Kumulus Data Services

November 2018 – July 2020 * **Core Responsibilities:** * Provided database administration, design, and performance optimization services to multiple enterprise clients at a Microsoft Azure Gold Partner consultancy. * Supported database platforms spanning Oracle Database, Microsoft SQL Server, MySQL, and PostgreSQL. * Designed cloud migration strategies and built modern data lake architectures in Microsoft Azure. * **Key Achievements & Projects:** * **Cloud Database Migration:** Successfully migrated legacy on-premises databases to Azure-managed databases (**Azure SQL Database, Azure Database for PostgreSQL/MySQL**), planning and executing with minimal business disruption. * **Azure Data Engineering:** Created and automated serverless ETL/ELT pipelines using **Azure Data Factory (ADF)** and **Azure Databricks**, ingesting and structuring client data in **Azure Data Lake Storage Gen2 (ADLS Gen2)**. * **Performance Rescue:** Diagnosed and resolved critical SQL Server and Oracle database bottlenecks (e.g., locking, high CPU utilization, memory pressure) through SQL tuning, query refactoring, index reorganization, and parameter tuning.

Data Specialist / PL-SQL & Python Developer @ Raízen S/A

January 2018 – November 2018 * **Core Responsibilities:** * Developed database solutions, stored procedures, and integrations for fiscal and tax calculations within the energy enterprise (ethanol, oil, and gas). * Automated data transformation processes and synchronized cross-system integrations using Python and Oracle PL/SQL. * **Key Achievements & Projects:** * **Tax Processing Speed-up:** Optimized long-running PL/SQL packages, functions, and triggers, reducing processing windows for daily fiscal operations. * **Python Automation:** Replaced manual file-handling processes with automated Python scripts, securing data integrity and ensuring tax compliance with government systems.

Oracle DBA / Developer / BI & VMware Admin @ Starplast Ind. E Com Ltda

July 2007 – November 2018 * **Core Responsibilities:** * Managed the company's entire core database (Oracle DBA), BI platform, virtual infrastructure, and internal software developments. * Developed custom applications in PL/SQL and PHP to support ERP JD Edwards integrations. * **Key Achievements & Projects:** * **VMware Virtualization:** Led the migration and virtualization of the entire bare-metal physical server infrastructure to a VMware ESXi private cloud environment, increasing hardware resource utilization by 70% and ensuring disaster recovery. * **Oracle JDE ERP Tuning:** Maintained and tuned the Oracle database infrastructure supporting the JD Edwards ERP, eliminating bottleneck issues during monthly financial closings. * **End-to-End Business Intelligence:** Modeled the corporate DW and designed, built, and maintained the company's BI platform utilizing **Oracle BI Enterprise Edition (OBIEE)**, creating reports and dashboards for executive decision-making. * **B2B Platform Development:** Programmed and launched the company's first B2B sales portal from scratch, designing the relational database model and creating APIs to sync transaction data directly with JDE.

PROJECTS & OPEN-SOURCE TOOLS

Productivity Tools

- **NoCloud Todo** — Privacy-focused, offline-first local task manager.
- **NoCloud SQL** — Client-side SQL query viewer, formatter, automatic business rules generator and query explainer with AI operating entirely in the browser using open formats.
- **NoCloud MD** — Browser-based offline-first Markdown live preview tool.
- **Speed Dial Page** — Customizable browser start page and quick-access dashboard.

Gnome Extensions

- **Magnific Launcher** — Search-focused application launcher for the Gnome shell.
 - **Water Reminder Extension** — Hydration tracker widget integrated into the Gnome shell panel.
 - **Compact Launcher** — Space-efficient, lightweight application menu.
 - **Movie Screensaver** — Gnome extension to play video files as desktop screensavers.
-

EDUCATION

Postgraduate Education

- **Postgraduate Certificate (Lato Sensu) in Data Science** — *Centro Universitário União das Américas Descomplica* (2024)
- **Database Engineering** — *Unicamp* (2018)

Graduation

- **Technician in Data Processing** (emphasis on Systems Development) — *Anhanguera University Center* (2002)
-

CERTIFICATIONS & TRAINING

AI & Agentic Engineering (Vanderbilt University)

- **AI Agent Developer** (2026)
- **Claude Code: Software Engineering with Generative AI Agents** (2026)
- **ChatGPT Advanced Data Analysis** (2026)
- **OpenAI GPTs: Creating Your Own Custom AI Assistants** (2026)
- **Prompt Engineering for ChatGPT** (2026)
- **AI Agents and Agentic AI with Python & Generative AI** (2026)
- **AI Agents and Agentic AI Architecture in Python** (2026)

AI Fundamentals & Productivity (Google)

- **AI for App Building** (2026)
- **AI for Data Analysis** (2026)
- **AI for Content Creation** (2026)
- **AI for Writing and Communicating** (2026)
- **AI for Research and Insights** (2026)
- **AI for Brainstorming and Planning** (2026)
- **AI Fundamentals** (2026)

Core Databases & Frameworks

- **Data Science Training Program with Technological Residency** — *Foundation for the Center for Research and Development in Telecommunications (CPQD) / PUC Campinas* (2023)
- **SCRUM Foundation Professional Certificate (SFPC)** — *Scrum Foundation* (2011)
- **MCTS Microsoft SQL Server Certification** — *Green Treinamentos / São Paulo – SP* (2008)
- **Oracle DBA Administrator** — *Senac / Limeira - SP* (2007)

Additional Training & Courses

- **Google Cloud Fundamentals:** Big Data and Machine Learning
- **Microsoft Azure:** Fundamentals
- **Introduction to R:** *Profa. Luciane Alcoforado and the Statistics is with R! team*
- **Specialization:** Open Source + Azure for Data and Analytics

LANGUAGES

- **Portuguese:** Native (Brazil)
- **English:** Intermediate

PERSONAL INFORMATION

- **Nationality:** Brazilian
- **Born:** July 4, 1977