

GERARDO VITALE ERRICO

Madrid, Spain • +34 656 836 229 • gerardo.vitale.errico@gmail.com • [linkedin.com/in/gerardo-vitale-errico](https://www.linkedin.com/in/gerardo-vitale-errico)

SUMMARY

Data Engineer with 4+ years of experience designing, building, and optimizing scalable data platforms across e-commerce, healthcare, energy, and technology domains. Strong background in Python, Spark, SQL, dbt, SQLMesh, and cloud-native architectures. Proven track record improving data pipeline reliability, implementing robust DataOps practices, and delivering high-quality, business-critical data products. Certified Google Cloud Professional Data Engineer with an MSc in Big Data & Analytics. Looking for challenging roles where data engineering drives measurable business impact.

EXPERIENCE

Axpo

Madrid, Community of Madrid, Spain.

Data Engineer

December 2025 - Present.

Asset Optimization and Logistics Division - EU Energy Markets

- Maintain and evolve a Kubernetes-based data platform that supports internal data services and critical business processes within Axpo's Asset Optimization & Logistics division.
- Build and enhance backend and data platform components using Python, FastAPI, PostgreSQL, and Airflow to support EU energy market operations.
- Contribute to the scalability, resilience, and day-to-day operation of the platform by improving core infrastructure, service integrations, and engineering practices.
- Collaborate on the design of an in-house end-to-end data lineage solution aimed at enabling row-level traceability for debugging, auditing, and data quality investigations.

Thoughtworks

Madrid, Community of Madrid, Spain.

Data Engineer

March 2023 - December 2025.

Top Global Airline, United Kingdom

- Modernized the data serving platform by implementing SQLMesh (open-source dbt-like DataOps framework), Snowflake, AWS Glue, and SYNQ, enabling scalable, maintainable, and auditable data workflows.
- Designed and delivered a Terraform-based initialize kit to standardize and accelerate the deployment of data product infrastructure, supporting CI/CD best practices for fast and reliable releases.
- Developed modular SQLMesh pipelines with SQL model transformations, adhering to medallion architecture principles and incorporating robust data quality checks through audits, ensuring reliable data assets.
- Engineered and implemented Slowly Changing Dimension (SCD) models to accurately capture and track changes in source systems, improving historical data integrity and auditability.

Major Healthcare Private Provider, United Kingdom

- Developed a Proof of Concept (PoC) for a scalable data platform using Microsoft Fabric, Azure DevOps and Microsoft Purview, enabling efficient data management, governance and analytics solutions.
- Architected and implemented seamless data pipelines using PySpark, Delta, Notebooks, and Lakehouses within the Microsoft Fabric environment, following ETL pattern and feeding PowerBI dashboards with critical business insights. Utilized Semantic Models to optimize data consumption effectively across teams, domains and organization.
- Engineered a comprehensive test strategy encompassing unit, integration, and End-to-End testing for the data pipelines, fully automated for continuous integration and deployment, ensuring robust data accuracy and system reliability.

- Utilized Great Expectations to implement robust data quality checks within the data pipeline, ensuring adherence to business rules and guaranteeing high-quality, trustworthy data and metrics.
- Proposed and initiated a roadmap to implement medallion architecture for the PoC domain within the Microsoft Fabric environment, designed to be scalable at the enterprise level, supporting future organizational growth and improving data governance.

Europe's Leading e-Commerce, Germany

- Enhanced the efficiency of data pipelines and improved the data serving process by redesigning aggregation jobs and metrics calculations into a new, streamlined pipeline, significantly boosting performance.
- Conducted comprehensive Exploratory Data Analysis (EDA) on diverse datasets, uncovering actionable insights that guided the development of targeted Data Products. Ensured these products were aligned with business objectives, driving value extraction and strategic decision-making.
- Addressed and mitigated data quality issues, ensuring the accuracy, consistency, and trustworthiness of metrics generated across the platform, enhancing the reliability of business intelligence.
- Collaborated closely with stakeholders to ensure alignment of project goals and deliverables with business needs, fostering clear communication and a shared vision for successful outcomes.

Junior Data Engineer

March 2022 - March 2023.

Europe's Leading e-Commerce, Germany

- Built and maintained a comprehensive data platform from scratch, delivering Data Products served as MicroStrategy dashboards to support critical business monitoring and decision-making.
- Designed and implemented robust Extract, Load, Transform (ELT) pipelines to integrate diverse raw data from external transactional systems and event-driven services. Leveraged CI/CD pipelines hosted on Kubernetes to ensure platform reliability and efficient deployment of data pipelines.
- Implemented Data Quality checks using PyDeequ to ensure the accuracy of Data Products and the metrics they generate. Integrated LightStep and Grafana for real-time data quality reporting and alerting based on predefined SLA thresholds, ensuring adherence to data quality standards.
- Developed a comprehensive testing strategy, focusing on unit and integration tests to ensure the reliability and robustness of data pipelines throughout the development lifecycle.
- Regularly hosted demos and showcases to communicate project progress and results to stakeholders, improving collaboration and fostering team engagement. Managed progress through Agile methodologies, including pair programming, leading to improved code quality and team efficiency.

Telefonica

Madrid, Community of Madrid, Spain.

September 2021 - February 2022.

Junior Backend Developer

- Developed TV Open Platform components as containerized microservices, following Agile methodologies, Test-Driven Development (TDD), SOLID principles, and Clean Architecture.
- Built and deployed a dockerized Kafka consumer service in Python, integrated with a SQL Server database, leveraging a CI/CD pipeline in Jenkins for seamless deployment and continuous delivery.
- Enhanced service monitoring by integrating Prometheus to expose consumer metrics, allowing real-time performance tracking and visualization through Grafana.
- Collaborated with a multidisciplinary, cross-functional team of Architects, Engineers, QAs, DevOps, Program Managers, Technical Program Managers, and Developers to analyze use cases and ensure successful project delivery.

EDUCATION & CERTIFICATIONS

Professional Data Engineer Certification

Google Cloud Platform, [Credential](#). Dec 2024

Master in Big Data and Analytics

Universidad Europea, Madrid, Spain. Oct 2021 - Oct 2022.

Data Analytics Bootcamp

Ironhack, Madrid, Spain. Oct 2020 – Dec 2020.

BA in Mechanical Engineering

UNEXPO, Barquisimeto, Lara, Venezuela. Feb 2018.

LANGUAGES

- **Spanish:** Native
- **English:** Fluent
- **Italian:** Intermediate