

Francisco Rodríguez

Data & AI Engineer | LLMs | Agentic Systems | Data Engineering
firguezcerezo@gmail.com | LinkedIn | Portfolio | GitHub

Professional Summary

Engineer with a hybrid profile across **data engineering and AI**, a scientific background and a **product-oriented mindset**. I bring **end-to-end vision** from data to production, analytical rigor and effective communication with technical and business stakeholders, turning strategic needs into **robust, governable and impact-measurable solutions**.

Professional Experience

Gran Telescopio de Canarias (GRANTECAN), Data & AI Engineer

07/2025 - 07/2026

- Co-developed an **agentic talk-to-your-data system** (speech-to-text, LangGraph and RAG over Deltalake) that lets non-technical users query in natural language with no schema or variable-name knowledge and **receive graphical answers**, democratizing data access across the organization.
- Designed an agentic assistant (RAG over local LLMs) for the assessment and drafting of **technical and administrative tender specifications**, **reducing manual workload by ~50%** for the technical-administrative staff and bringing corporate context into the documentation process.
- Led the build of a SOP and playbooks platform as a **procedure sequencer-orchestrator for VUCA environments**: rigorous version control and execution traceability, removing reliance on tacit knowledge and outdated procedures scattered across emails and documents.
- Developed a React and Node.js web app for telescope time-usage analytics for management and operations: asynchronous cross-validation of incident logs and observation reports in a single tool, **replacing manual Excel-based coordination** between teams.
- Built PySpark ETL pipelines on the corporate Deltalake: ingested new sources to **break information silos** and promoted data from **raw to silver and gold layers** (medallion architecture), accelerating on-demand telemetry access and offloading the data team for both technical and non-technical users.
- Deployed the local inference infrastructure on a GPU cluster with **vLLM, SLURM and RHEL**, technical backbone for the agentic systems above and a key enabler of the on-premise approach.
- Member of the organization's pioneer on-premise AI team: led the study and validation of **compliance with the EU AI Act and AESIA**, shaping the **privacy and governance framework** underpinning the team's PoCs, MVPs and production systems.

Cabildo of La Palma, Data Scientist

01/2024 - 01/2025

- Led EDA with Python and R across 5+ air-quality datasets in Puerto Naos, uncovering a **Pearson correlation of ~0.6 between coastal CO2 levels and tidal cycles**; the finding drove the Emergency Department's post-eruption habitability protocol review.
- Designed interactive dashboards in Power BI, Tableau and ArcGIS Online (including a **Puerto Naos digital twin** used daily by Emergency for volcanic CO2 tracking) and automated weekly Python workflows, **saving ~12 h per month** for non-technical stakeholders.
- Standardized and integrated a **3M+ record legacy dataset** into PostgreSQL, extending the Cabildo's accessible meteorological archive from 2022 back to 2016 and ensuring continuity for the public portal.

Previous experience as private tutor in Physics, Chemistry and Mathematics.

Education

MSc in Data Science and Big Data, Open University of Catalonia (UOC)

2025

- Master's Thesis:** sales forecasting and audience segmentation system for Filarmonía de Madrid with real ticketing data (Random Forest, CNN, RNN, ARIMA, K-Means, DBSCAN, RFM).

BSc in Physics, University of La Laguna (ULL)

2021

- Bachelor's Thesis:** photometric calibration of the EMIR instrument (GTC, IAC) using Python and scientific documentation in LaTeX.

Skills

Programming: Python, SQL, Node.js, React, TypeScript, R

AI & Machine Learning: LangChain, LangGraph, vLLM, Ollama, RAG, agentic systems, prompt engineering, embeddings, vector databases, NLP, deep learning, PyTorch, TensorFlow, Scikit-learn

Data Engineering: PySpark, Deltalake (medallion architecture), ETL, data pipelines, Apache Kafka, Trino, PostgreSQL

MLOps & DevOps: Linux, SLURM, Docker, Kubernetes, CI/CD, Git, REST API, Nginx, Apache

Cloud & Visualization: AWS, Power BI, Tableau, ArcGIS Online

Templating & Markup: Jinja, LaTeX

Languages: Spanish (native), English (fluent)