

RODRIGO CALDERON FLORES

Data Analyst | Applied AI & Automation

rocalderonflores@gmail.com • Lima, Peru • [linkedin.com/in/rodrigocalderonf/](https://www.linkedin.com/in/rodrigocalderonf/) • rodrigocalderon.pages.dev • English C2 • Spanish native • French intermediate

PROFESSIONAL SUMMARY

Data analyst with seven years across industrial manufacturing, real-time economic data, and research funded by the U.S. government and the pharmaceutical industry. A physician by training, with a master's in biostatistics that turned clinical reasoning into quantitative method under research standards where data quality is not optional. Works embedded in the business: learning every rule and every exception, using analysis to find what the operation cannot see, then building the data models, pipelines, infrastructure, and AI automation that act on it.

SKILLS

Languages & Query: Python, SQL, R, Stata, TypeScript, JavaScript

Analytics & Modeling: pandas, NumPy, statsmodels, SciPy, scikit-learn, time-series and lead-lag analysis, demand forecasting, econometric and macro-financial analysis, statistical inference, experimental and validation study design, classification performance evaluation (sensitivity, specificity, ROC/AUC), quantitative backtesting

Data Platforms & Pipelines: BigQuery, Oracle, PostgreSQL / Supabase (row-level security), Firestore (NoSQL), REDCap, Google Cloud Platform (GCP), AWS (S3, SageMaker, Lambda), Firebase, dbt (dimensional modeling, semantic layer and metric definition), ETL / ELT design, Apache Airflow / Cloud Composer, batch and event-driven pipelines, data quality validation and pipeline observability

Visualization & BI: Power BI, Tableau, Plotly, matplotlib, dashboard and KPI design, executive and institutional data storytelling

AI & LLM Engineering: Anthropic Claude API, OpenAI, Google Gemini, vision and multimodal models, computer vision (YOLO, RF-DETR, OpenCV, pose estimation and tracking), on-device inference (llama.cpp, whisper.cpp), MCP server development, Claude Code, retrieval-augmented generation (RAG), prompt engineering with schema-constrained structured output, large-scale classification of unstructured data with LLMs

Automation & Agents: n8n workflow and agent orchestration, WhatsApp Cloud API, web scraping and automated data extraction, document and transaction processing automation, email marketing automation

Software & Infrastructure: React, Next.js, Vite, Fastify / Node.js, three.js / WebXR, REST API design and integration, microservices, serverless architectures (Firebase Cloud Functions, Cloudflare Workers), Docker, CI/CD with automated testing (Vitest, Playwright), Git, legacy system migration and modernization

PROFESSIONAL EXPERIENCE

Data Analyst | AI Solutions Engineer

2026 – Present

Industrias Electroquímicas S.A. (IEQSA) — Lima, Peru | Reporting to the CEO and executive committee

Largest zinc-processing industrial complex in South America; ~400 employees, 60+ years of operation, exporting to 25+ countries.

- Integrate the company's Oracle production database — hundreds of millions of records — into the new ERP, preserving critical business logic through the migration from GeneXus 9 (2000) to a modern TypeScript architecture with microservices, Docker, and CI/CD.
- Build demand forecasting models combining SQL queries against the ERP with external market sources, delivering inputs for decision-making across the seven company-wide modules the ERP serves: Sales, Maintenance, Warehouse, Safety, Logistics, HR, and Finance.
- Automate freight-cost quotation, cutting turnaround from 15–18 minutes to under one minute across 20–30 daily quotations — a 94% reduction that removes 5–8 hours of manual work per day, roughly 1,500 hours a year. Extended the same approach across international metal price tracking, order entry, electronic invoicing, HR and shipping-terminal API integrations, and automated extraction and analysis of competitor pricing.
- Built a company-specific MCP server encoding business rules, architecture standards, validators, and mandatory controls, enabling a 20-person internal Systems Engineering team to work autonomously with Claude Code at consistent quality and traceability.

Data & AI Lead | Process Consultant

2026 – Present

E & J Matthei Maderas del Perú S.A. — Madre de Dios, Peru | Reporting to executive management

One of Peru's five largest timber exporters; three plants across the Peruvian Amazon, operating under SERFOR's national traceability regime.

- Model the company's timber traceability chain from standing forest to dispatch, working from a full understanding of the operation and of the regulator's public and private systems, defining the rules that govern lot assignment, volume calculation, and source attribution, to replace manual spreadsheet record-keeping with an automated system.
- Test those rules against ~30,000 records, disproving several the operating team held as true, including a species key already in production that merged two species the regulator keeps apart, affecting 74% of incoming timber.
- Build the ingestion engine replacing manual data entry at goods receipt: a streaming HTML extractor for the regulator's API-less public portal on Cloudflare Workers, plus an authenticated client for its private REST operations book.

- Design a multi-model AI reading pipeline for handwritten guides, with model routing and failover, dual parallel readings cross-checked field by field, and automated escalation of unverifiable fields to human review; validation arithmetic runs independently of the model, reaching 100% accuracy on every verifiable field against the official record.
- Implement the traceability modules in React and TypeScript over Firebase with Python Cloud Functions: automatic lot classification, plant coding, purchased-lumber registry, and per-contract volume tracking.

Data Analytics & Economics Research Lead

2024 – 2026

Truflation — Startup / B2B SaaS, real-time economic data

- Designed and executed the first systematic benchmark between Truflation's real-time inflation data and the U.S. Bureau of Labor Statistics CPI, identifying an average lead of ~45 days in trend changes — the analysis that repositioned the flagship product from an alternative inflation measure to a leading economic indicator.
- Built forecast-accuracy benchmarking pipelines in Python (pandas, NumPy) comparing institutional CPI predictions against BLS prints, with rolling-window error ranking and consistency analysis used in published research and client-facing content.
- Developed analytical models quantifying temporal relationships between inflation and asset classes including U.S. Treasury yields, the DXY index, and Bitcoin, and implemented a state-aware trend-detection algorithm with hysteresis to classify inflation regimes.
- Designed a deterministic asset allocation strategy driven by inflation signals, validated through leveraged backtesting with liquidation modelling and Sharpe/Sortino metrics. Adapted to on-chain infrastructure with the founders of QuantAMM; the partnership took first place at Chainlink Demo Day, SmartCon 2025 (New York).
- Led the collaboration with Argentina's Ministry of Economy to launch Truflation Argentina, designing an automated pipeline that ingests, structures, and AI-categorizes Argentina's national SEPA price feed — roughly 12 million daily records reported by 3,600 retailers across 70,000+ products — published as open data on blockchain infrastructure for immutability and traceability.
- Produced dashboards, reports, and animated data visualizations (Plotly, matplotlib, MoviePy) for a global audience of 200,000+ subscribers and delivered webinars to institutional clients including banks, hedge funds, and portfolio managers.

Project Manager — Clinical Research Data

2021 – Present

Institute of Tropical Medicine "Alexander von Humboldt" — UPCH

- Coordinate an ongoing AI-based cough-monitoring study (IDRC / AI4PEP Grant 109981-001) run with McGill University (Canada), the University of Texas Medical Branch (United States), and Heidelberg University (Germany), building the longitudinal audio and clinical database that will train a convolutional neural network to classify five respiratory infections from cough sound.
- Own data architecture and the analysis plan for that study: REDCap instrument design with field-level validation rules, staged quality control and source data verification, training and validation splits, model evaluation by AUC, sensitivity and specificity, algorithmic bias assessment stratified by sex and demographic subgroup, and geospatial density mapping across two Lima districts.
- Study coordinator and AI/data analyst on a UPCH-InnovaLab proposal to the TDR/WHO call on digital solutions for infectious diseases (CA26-0034; US\$48K, 12 months; Oct 2026): drone imagery and computer vision (RF-DETR vs. YOLO11) to map Aedes breeding sites across ~700 households in San Juan de Lurigancho, Lima. Trained the pilot YOLO11s detector (mAP@0.5 of 0.74–0.75 on water tanks).
- Managed data operations for four international diagnostic trials funded by the U.S. National Institutes of Health and Abbott Pharmaceuticals, spanning multi-country sites in Peru and Uganda.
- Performed statistical analysis of trial results and translated findings into publications and regulatory deliverables, contributing to eight peer-reviewed papers.

Studies: AI cough monitor (IDRC 109981-001) • NOD-PedFEND (NIH 1R01AI152159-01) • FEND-TB (NIH 1U01AI152084-01) • Panbio™ Rapid Panel, Abbott.

Methodological & Statistical Advisor

2020 – 2021

Novo Nordisk Pharmaceutical

- Designed statistical analysis frameworks, data validation strategies, and multivariable modeling plans (regression, causal inference, risk stratification, model validation) for industry-sponsored research on real-world clinical datasets; both studies reached peer-reviewed publication.
- Developed mortality risk models for hospitalized COVID-19 patients, evaluating clinical predictors while controlling for confounding and bias, and analysed multi-factor risk patterns and physician decision-making in type 2 diabetes from patient-level data.

Data & Operations Strategy Advisor

2020 – 2021

INABIF, National Comprehensive Family Welfare Program — Ministry of Vulnerable Populations, Peru

- Advised executive leadership on data-driven COVID-19 response and operational strategy across government-run shelters serving vulnerable populations nationwide.
- Designed and automated information workflows between shelters and central management, improving how operational data was collected and consolidated.

Guest Lecturer — Biostatistics and Research

2024 – Present

School of Medicine, Universidad Peruana Cayetano Heredia

- Invited back each year to teach biostatistics and research methodology to medical students, covering applied data analysis, statistical inference, regression modeling, and study design, and guiding students from research question to structured analytical approach.

Research Fellow — PARACAS Program

2019 – 2021

Institute of Tropical Medicine "Alexander von Humboldt" — UPCH | U.S. NIH Grant D43 TW009763

- Worked with real-world clinical and epidemiological datasets under a competitive NIH training grant: data cleaning, validation, regression modeling, and analytical frameworks to quantify associations in complex health data.

SELECTED INDEPENDENT PROJECTS (2026)

- Prevtor.** My own computer-vision product for industrial safety and production. It watches a plant's existing cameras and flags missing protective equipment and people under suspended loads, with WhatsApp alerts and a 3D plant view; no facial recognition, faces blurred. Audited 108 cameras and 7 recorders before defining what could be detected, chose RF-DETR over YOLO11 after a same-data benchmark (59/59 vs. 56/59 steel straps; Apache licence), and cut false "no helmet" alerts from 98 to 5 by combining pose, segmentation, and colour.
- Sable CRM.** Designed and built in one week a multi-tenant real-estate CRM around WhatsApp — inbox, sales funnel, visit scheduling, unit inventory, quotes, and campaigns — with an AI assistant (Claude) that answers, scores, and hands off leads. Tenant and role isolation enforced in Postgres row-level security, per-company AI cost budgets, and 105 unit and 18 end-to-end tests. Next.js, Supabase, WhatsApp Cloud API. Working demo.
- Offline voice virtual patient.** Clinical-interview simulator for a tuberculosis case that runs fully on-device (whisper.cpp, llama.cpp with a 35B mixture-of-experts model, Piper), with ~250 ms latency. A deterministic case director releases each hidden finding only when asked, cutting leakage of unasked facts from 5/6 to 0/12. Prototype.
- Tuberculous granuloma in 3D.** Parametric model of a TB granuloma and TB lungs spanning five orders of magnitude (25 cm to 3 µm), exporting 3D-printable STL, a WebXR viewer, and a morphometry report, plus a CT-to-STL pipeline. Prototype.
- Applied AI for Productivity.** Designed a 7-module, 31.5-hour course for non-technical teams: prompting, ChatGPT, Claude and Gemini, custom assistants, automation, role-based cases, and AI ethics.

EDUCATION

MSc in Biostatistics and Research — Ranked 2nd in graduating class

2019 – 2020

Universidad Peruana Cayetano Heredia (UPCH)

M.D., Doctor of Medicine — Ranked 1st in graduating class (Valedictorian)

2011 – 2018

Universidad Peruana Cayetano Heredia (UPCH)

SELECTED PEER-REVIEWED PUBLICATIONS

First or named author on 8 peer-reviewed publications. Selected below; full list available on request.

- Calderon-Flores R, Caceres-Cardenas G, Ali K, De Vos M, Emperador D, Cáceres T, et al. Diagnostic performance of four lateral flow immunoassays for COVID-19 antibodies in Peruvian population. *PLOS Global Public Health*. Jun 2023;3(6). doi:10.1371/journal.pgph.0001555
- Wanda F, Calderon-Flores R, Held F, et al. Acceptability and feasibility of oral swabs for tuberculosis diagnosis in young children: a qualitative study from Uganda and Peru. *Frontiers in Tuberculosis*. Jan 2026. doi:10.3389/ftubr.2026.1726647

SELECTED SPEAKING ENGAGEMENTS

- Chainlink House, Ethereum Devconnect 2025 (Argentina)** — Invited by Chainlink Labs after the QuantAMM strategy took first place at Chainlink Demo Day, SmartCon 2025; panelist for Truflation on real-time economic data.
- Internal Research Project Competition, Ministry of Health of Peru, Dec 2023** — Evaluation panel member; technical assessment of research projects from hospitals nationwide.

AWARDS & SCHOLARSHIPS

- Emerging Leaders in Global Health Scholarship (2022)** — Perelman School of Medicine, University of Pennsylvania.
- Contenta Award, Valedictorian (2019)** — ranked 1st across all eight years of Human Medicine, UPCH.
- Volunteer Physician, Manchay Alto Health Post (2024–Present)** — pro bono primary care in an underserved community.