RODRIGO CALDERON FLORES

Data Analyst | Applied AI & Automation

rocalderonflores@gmail.com • Lima, Perú • linkedin.com/in/rodrigocalderonf
• rodrigocalderon.pages.dev • Inglés C2 · Español nativo · Francés intermedio

RESUMEN PROFESIONAL

Analista de datos con siete años de experiencia en manufactura industrial, datos económicos en tiempo real e investigación financiada por el gobierno de Estados Unidos y por la industria farmacéutica. Médico de formación, con maestría en bioestadística, acostumbrado a trabajar con datos bajo estándares donde un error tiene consecuencias reales. Siempre empiezo investigando a fondo el negocio, cada regla y cada excepción, para encontrar lo que todavía no se está midiendo y construir end-to-end los modelos de datos, pipelines, infraestructura y automatizaciones con IA necesarios para tomar acción.

HABILIDADES TÉCNICAS

Lenguajes y Consulta: Python, SQL, R, Stata, TypeScript, JavaScript

Analítica y Modelamiento: pandas, NumPy, statsmodels, SciPy, scikit-learn, series de tiempo y relaciones lead–lag, forecasting de demanda, análisis econométrico y macrofinanciero, inferencia estadística, diseño experimental y de validación, desempeño de clasificadores (sensibilidad, especificidad, ROC/AUC), backtesting cuantitativo

Plataformas de Datos y Pipelines: BigQuery, Oracle, PostgreSQL / Supabase (row-level security), Firestore (NoSQL), REDCap, Google Cloud Platform (GCP), AWS (S3, SageMaker, Lambda), Firebase, dbt (modelamiento dimensional, capa semántica y definición de métricas), diseño ETL / ELT, Apache Airflow / Cloud Composer, pipelines batch y event-driven, validación de calidad de datos y observabilidad de pipelines

Visualización y BI: Power BI, Tableau, Plotly, matplotlib, diseño de dashboards y KPIs, data storytelling para audiencias ejecutivas e institucionales

IA, LLM y Automatización: Anthropic Claude API, OpenAI, Google Gemini, modelos de visión y multimodales, visión por computadora (YOLO, RF-DETR), inferencia local (llama.cpp, whisper.cpp), desarrollo de servidores MCP, Claude Code, RAG, prompt engineering con salida estructurada por esquema, clasificación a gran escala de datos no estructurados con LLMs, orquestación de workflows y agentes con n8n, WhatsApp Cloud API, web scraping y extracción automatizada de datos, automatización documental y transaccional

Software e Infraestructura: React, Next.js, Vite, Fastify / Node.js, three.js / WebXR, diseño e integración de APIs REST, microservicios, arquitecturas serverless (Firebase Cloud Functions, Cloudflare Workers), Docker, CI/CD con testing automatizado (Vitest, Playwright), Git, migración y modernización de sistemas legacy

EXPERIENCIA PROFESIONAL

Data Analyst | AI Solutions Engineer **2026 – Actualidad**

Industrias Electroquímicas S.A. (IEQSA) — Lima, Perú | Reporta al CEO y al comité de gerencia

Mayor complejo industrial de procesamiento de zinc de Sudamérica; ~400 empleados, más de 60 años de operación, exporta a más de 25 países.

- Integración de la base de datos Oracle de producción — cientos de millones de registros — al nuevo ERP, preservando la lógica de negocio crítica durante la migración desde GeneXus 9 (2000) a una arquitectura moderna en TypeScript con microservicios, Docker y CI/CD.
- Construcción de modelos de forecasting de demanda combinando consultas SQL sobre el ERP con fuentes de mercado externas, entregando insumos para la toma de decisiones en los siete módulos transversales que cubre el ERP: Ventas, Mantenimiento, Almacén, Seguridad, Logística, RR.HH. y Finanzas.
- Automatización de la cotización de fletes, de 15–18 minutos a menos de uno sobre 20–30 cotizaciones diarias — 94% de reducción, que elimina de 5 a 8 horas de trabajo manual al día, cerca de 1,500 horas al año. Extensión del enfoque al seguimiento de precios internacionales de metales, ingreso de órdenes, facturación electrónica, integraciones con APIs de RR.HH. y terminales portuarios, y análisis automatizado de precios de la competencia.
- Desarrollo de un servidor MCP propio que codifica reglas de negocio, estándares de arquitectura, validadores y controles obligatorios, habilitando a un equipo interno de 20 personas de Ingeniería de Sistemas para trabajar de forma autónoma con Claude Code con calidad y trazabilidad consistentes.

Data & AI Lead | Consultor de Procesos **2026 – Actualidad**

E & J Matthei Maderas del Perú S.A. — Madre de Dios, Perú | Reporta a la gerencia

Uno de los cinco mayores exportadores de madera del Perú; tres plantas en la Amazonía peruana, bajo el régimen nacional de trazabilidad de SERFOR.

- Modelamiento de la cadena de trazabilidad de la madera desde el árbol en pie hasta el despacho, a partir del entendimiento completo de la operación y de los sistemas del regulador, definiendo las reglas que gobiernan la asignación de lotes, el cálculo de volúmenes y la atribución de origen, y reemplazando el registro manual en hojas de cálculo por un sistema automatizado.

- Contraste de esas reglas contra ~30,000 registros, refutando varias que el equipo operativo daba por ciertas, entre ellas una clave de especies ya en producción que fusionaba dos especies que el regulador mantiene separadas, afectando al 74% de la madera que ingresa.
- Construcción del motor de ingesta que reemplaza el ingreso manual de datos en recepción: un extractor HTML por streaming sobre Cloudflare Workers para el portal público del regulador, que no tiene API, más un cliente autenticado para su libro de operaciones REST privado.
- Diseño de un pipeline de lectura con IA multi-modelo para guías manuscritas: ruteo de modelos y failover, doble lectura paralela contrastada campo por campo y escalamiento automático a revisión humana; la validación aritmética corre independiente del modelo, con 100% de exactitud en todo campo verificable.
- Implementación de los módulos de trazabilidad en React y TypeScript sobre Firebase con Cloud Functions en Python: clasificación automática de lotes, codificación de planta, registro de madera comprada y seguimiento de volúmenes por contrato.

Data Analytics & Economics Research Lead **2024 – 2026**
Truflation — Startup | SaaS B2B de datos económicos en tiempo real

- Diseño y ejecución del primer benchmark sistemático entre los datos de inflación en tiempo real de Truflation y el CPI de la Oficina de Estadísticas Laborales de EE.UU., identificando un adelanto promedio de ~45 días en los cambios de tendencia — el análisis que reposicionó el producto insignia de medida alternativa de inflación a indicador económico líder.
- Construcción de pipelines en Python (pandas, NumPy) para benchmarking de precisión de pronósticos, comparando las predicciones de CPI de instituciones contra las publicaciones del BLS, con ranking de error por ventanas móviles y análisis de consistencia usados en investigación publicada y contenido para clientes.
- Desarrollo de modelos analíticos que cuantifican relaciones temporales entre la inflación y distintas clases de activos, incluidos los rendimientos del Tesoro de EE.UU., el índice DXY y Bitcoin, e implementación de un algoritmo de detección de tendencia con histéresis para clasificar regímenes inflacionarios.
- Diseño de una estrategia determinística de asignación de activos basada en señales de inflación, validada mediante backtesting apalancado con modelamiento de liquidación y métricas Sharpe/Sortino, adaptada a infraestructura on-chain junto a los fundadores de QuantAMM; la alianza obtuvo el primer lugar en el Chainlink Demo Day, SmartCon 2025 (Nueva York).
- Liderazgo de la colaboración con el Ministerio de Economía de Argentina para lanzar Truflation Argentina, con un pipeline que ingiere, estructura y categoriza con IA el feed nacional de precios SEPA — ~12 millones de registros diarios de 3,600 comercios sobre más de 70,000 productos — publicados como datos abiertos en blockchain.
- Producción de dashboards, reportes y visualizaciones de datos animadas (Plotly, matplotlib, MoviePy) para una audiencia global de más de 200,000 suscriptores, y dictado de webinars a clientes institucionales como bancos, hedge funds y gestores de portafolio.

Project Manager — Datos de Investigación Clínica **2021 – Actualidad**
Instituto de Medicina Tropical "Alexander von Humboldt" — UPCH

- Coordinación de un estudio en curso de monitoreo de tos con IA (IDRC / AI4PEP, grant 109981-001) junto a McGill University (Canadá), la University of Texas Medical Branch (Estados Unidos) y Heidelberg University (Alemania), construyendo la base de datos longitudinal de audio y clínica con la que se entrenará una red neuronal convolucional para clasificar cinco infecciones respiratorias a partir del sonido de la tos.
- Responsable de la arquitectura de datos y del plan de análisis: diseño de instrumentos en REDCap con validación a nivel de campo, control de calidad por etapas, particiones de entrenamiento y validación, evaluación por AUC, sensibilidad y especificidad, evaluación de sesgo algorítmico por sexo y subgrupo demográfico, y mapeo geoespacial de densidad en dos distritos de Lima.
- Coordinador del estudio y analista de IA y datos en la propuesta UPCH–InnovaLab a la convocatoria TDR/OMS sobre soluciones digitales para enfermedades infecciosas (CA26-0034; US$48K, 12 meses; oct. 2026): imágenes de drones y visión por computadora (RF-DETR vs. YOLO11) para mapear criaderos de Aedes en ~700 viviendas de San Juan de Lurigancho, Lima. Entrenamiento del detector piloto YOLO11s (mAP@0.5 de 0.74–0.75 en tanques de agua).
- Gestión de las operaciones de datos de cuatro ensayos diagnósticos internacionales financiados por los Institutos Nacionales de Salud de EE.UU. (NIH) y Abbott Pharmaceuticals, con sedes en Perú y Uganda.
- Análisis estadístico de los resultados y traducción de los hallazgos en publicaciones y entregables regulatorios, contribuyendo a ocho artículos arbitrados.

Estudios: monitor de tos con IA (IDRC 109981-001) • NOD-PedFEND (NIH 1R01AI152159-01) • FEND-TB (NIH 1U01AI152084-01) • Panel Rápido Panbio™, Abbott.

Asesor Metodológico y Estadístico **2020 – 2021**
Novo Nordisk Pharmaceutical

- Diseño de marcos de análisis estadístico, estrategias de validación de datos y planes de modelamiento multivariable (regresión, inferencia causal, estratificación de riesgo y validación de modelos) para investigación patrocinada por la industria sobre datos clínicos de mundo real; ambos estudios llegaron a publicación arbitrada.
- Desarrollo de modelos de riesgo de mortalidad para pacientes hospitalizados con COVID-19, evaluando predictores clínicos y controlando confusión y sesgo, y análisis de patrones de riesgo multifactorial y la toma de decisiones médicas en diabetes tipo 2 a partir de datos a nivel de paciente.

Asesor de Estrategia de Datos y Operaciones 2020 – 2021
INABIF, Programa Integral Nacional para el Bienestar Familiar — Ministerio de la Mujer y Poblaciones Vulnerables, Perú

- Asesoría a la Dirección Ejecutiva en la respuesta al COVID-19 basada en datos y en la estrategia operativa de los albergues estatales que atienden poblaciones vulnerables a nivel nacional.
- Diseño y automatización de flujos de información y comunicación entre albergues, mejorando la recolección, consolidación y circulación de datos operativos entre las sedes y la administración central.

Docente Invitado — Bioestadística e Investigación 2024 – Actualidad
Facultad de Medicina, Universidad Peruana Cayetano Heredia

- Invitado cada año a dictar bioestadística y metodología de la investigación a estudiantes de medicina: análisis de datos aplicado, inferencia estadística, modelamiento de regresión y diseño de estudios.

Research Fellow — Programa PARACAS 2019 – 2021
Instituto de Medicina Tropical "Alexander von Humboldt" — UPCH | Grant NIH D43 TW009763 (EE.UU.)

- Manejo de datos clínicos y epidemiológicos de mundo real bajo un grant competitivo de formación del NIH: limpieza y validación de datos, modelamiento de regresión y marcos analíticos para cuantificar asociaciones en datos de salud complejos.

PROYECTOS INDEPENDIENTES SELECCIONADOS (2026)

- **Sable CRM.** Diseño y construcción en una semana de un CRM inmobiliario multiempresa alrededor de WhatsApp — bandeja, embudo, agenda de visitas, inventario de unidades, cotizaciones y campañas — con un asistente de IA (Claude, con OpenRouter como respaldo) que responde, califica y deriva leads. Aislamiento entre empresas y roles garantizado con row-level security en Postgres, presupuesto de costo de IA por empresa, y 105 pruebas unitarias y 18 de extremo a extremo como control previo a staging. Next.js, Supabase, WhatsApp Cloud API. Demo funcional.
- **Paciente virtual por voz, sin conexión.** Simulador de entrevista clínica para un caso de tuberculosis que corre íntegramente en local (whisper.cpp, llama.cpp con un modelo mixture-of-experts de 35B, Piper), con ~250 ms de latencia de punta a punta. Un director de caso determinista libera los hallazgos ocultos solo cuando se preguntan, reduciendo la filtración de datos no preguntados de 5/6 a 0/12 y manteniendo 12/12 de los preguntados. Prototipo.
- **Granuloma tuberculoso en 3D.** Modelo paramétrico de un granuloma y de pulmones con tuberculosis que abarca cinco órdenes de magnitud (de 25 cm a 3 µm) y exporta piezas STL imprimibles, un visor WebXR y un reporte de morfometría trazado a la literatura, más un pipeline de tomografía (DICOM/NIfTI) a STL. Prototipo.
- **IA aplicada a la productividad.** Diseño de un curso de 7 módulos y 31.5 horas para equipos no técnicos: prompting, ChatGPT, Claude y Gemini, asistentes propios, automatización, casos por rol y ética de la IA.

FORMACIÓN ACADÉMICA

Maestría en Bioestadística e Investigación — Segundo puesto de la promoción 2019 – 2020
Universidad Peruana Cayetano Heredia (UPCH)

Médico Cirujano — Primer puesto de la promoción 2011 – 2018
Universidad Peruana Cayetano Heredia (UPCH)

PUBLICACIONES ARBITRADAS SELECCIONADAS

Primer autor o autor nombrado en 8 publicaciones arbitradas. A continuación una selección; lista completa disponible a solicitud.

1. Calderon-Flores R, Caceres-Cardenas G, Alí 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
2. 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

PONENCIAS SELECCIONADAS

- **Chainlink House, Ethereum Devconnect 2025 (Argentina)** — Invitado por Chainlink Labs tras el primer lugar de la estrategia QuantAMM en el Chainlink Demo Day, SmartCon 2025. Panelista por Truflation sobre datos económicos en tiempo real y análisis on-chain.
- **Ministerio de Salud del Perú, dic. 2023** — Jurado evaluador del Concurso Interno de Proyectos de Investigación, con hospitales de todo el país.

PREMIOS Y BECAS

- **Beca Emerging Leaders in Global Health (2022)** — Perelman School of Medicine, University of Pennsylvania.
- **Premio Contenta (2019)** — primer puesto en los ocho años de Medicina Humana, UPCH.
- **Médico voluntario, Posta Manchay Alto (2024–Actualidad)** — atención primaria pro bono en una comunidad desatendida.