

Carlos Perez Molero

Forward Deployed AI Engineer & Tech Lead

perez.moleroc@gmail.com | +54 9 11 2251 0584 | Buenos Aires, Argentina |

linkedin.com/in/carlos-perez-molero | github.com/engcarlosperezmolero | charlymolero.com

PERFIL PROFESIONAL

Forward Deployed AI Engineer y Tech Lead con 7+ años de experiencia construyendo software y sistemas de datos en producción. Me integro con los equipos del cliente, hago el discovery, defino la arquitectura de la solución de punta a punta y lidero la entrega hasta el despliegue y el handoff. Recientemente enfocado en agentes de AI, aplicaciones con LLMs, sistemas RAG, evaluación de LLMs y despliegues cloud, integrando OpenAI, Claude y Gemini en workflows reales de usuarios. Formación universitaria en Ingeniería Química, valiosa para aplicaciones de AI en industria, energía y procesos.

EXPERIENCIA

HUMAI Institute

02/2025 - Present

Forward Deployed AI Engineer & Tech Lead · Contractor

- Nexo directo con los clientes: relevamiento de requerimientos, reuniones semanales, demos, handoff y entrevistas a ingenieros para staff augmentation
- Coordiné equipos de expertos externos en AI para producir cursos de vanguardia: entrevisté y negocié con los instructores, escribí guiones y guías docentes, y controlé la calidad de cada clase grabada
- Construí aplicaciones full-stack con FastAPI, PostgreSQL (Cloud SQL), GCP (Cloud Run, GCS), GitHub Actions y Docker
- Implementé RBAC seguro, autenticación basada en cookies y scraping en tiempo real con Playwright
- Diseñé pipelines ETL con VertexAI (basados en LLM y RAG) para procesamiento y análisis de datos
- Desarrollé frontend React con SSE, app companion Kotlin y backend usando FastAPI con CI/CD
- Dicté más de 100 clases sobre Python, automatización, web scraping, bases de datos, CI/CD y Git

AI, Python, FastAPI, React, Next.js, TypeScript, PostgreSQL, GCP, AWS, Docker, Playwright, VertexAI, Langfuse, Kotlin, GitHub Actions

Conversational AI Interview Platform

Plataforma de entrevistas conversacionales potenciada por LLMs para investigación cualitativa: un agente de AI conduce entrevistas en profundidad por texto o voz, decide cuándo repreguntar y consolida los hallazgos. Llevé el discovery, las demos quincenales y el handoff directamente con el equipo de investigación, y elegí los modelos con evals LLM-as-judge sobre entrevistas reales. Construí el sistema completo end-to-end en AWS.

AI, Next.js, TypeScript, Claude, AWS Bedrock, PostgreSQL, AWS

Parental Monitoring Platform

Aplicación de monitoreo parental potenciada por LLMs que scrapea plataformas de mensajería usando whatsmeow y procesa conversaciones a través de múltiples capas de LLMs para generar métricas sobre el estado emocional y alertas. Trabajé directamente con los fundadores desde la beta hasta el funding. Construí el sistema completo end-to-end: backend, frontend, scraper, servicio de análisis LLM, alertas y observabilidad con Langfuse en GCP.

AI, Go, whatsmeow, OpenAI, Claude, Langfuse, GCP, React, FastAPI

Music Industry Management App

Aplicación impulsada por AI para managers musicales que usa LLMs para generar y ejecutar planes de lanzamiento de artistas. Trabajé directamente con el fundador, lideré un equipo de 3 ingenieros y definí las decisiones de modelo y arquitectura: migré de Django a FastAPI sin downtime y reduje 88% el costo de LLM por plan.

AI, React, FastAPI, Django, OpenAI, PostgreSQL

Government Process Automation

Sistema de automatización de trámites a distancia para registros gubernamentales, que reemplazó la carga manual del equipo de operaciones del cliente. Relevé requerimientos con las operadoras y el PM, dirigí al técnico del cliente y puse en producción un pipeline de agentes self-healing que recupera sesiones caídas, retoma desde el paso que falló y nunca presenta un trámite dos veces. Construido con Playwright, backend API, dashboard y procesamiento por lotes.

AI, Playwright, FastAPI, React, PostgreSQL

Banking Data Migration

Sr. Data Engineer migrando un proceso diario de conciliaciones desde un servidor on-prem IBM AS400 a la nube usando AWS, Databricks, Airflow, Spark, Kafka y Kubernetes. Lideré la adopción de AI del equipo de 10 personas: skills de Claude Code, hooks como guardrails y servidores MCP seguros hacia on-prem y herramientas de observabilidad. Optimicé ETL y DAGs reduciendo tiempos en más del 50%.

AI, AWS, Databricks, Airflow, Spark, Kafka, Kubernetes, Python

Techunting
12/2023 - 02/2025

Desarrollador Sr. Full-stack Python · Contractor

- Desarrollé el backend en AWS para un sistema de brokerage de super yates usando ECS y FastAPI, diseñando una base de datos PostgreSQL con 20 tablas y construyendo 34 endpoints de API
- Implementé un motor de búsqueda con OpenSearch y gestioné imágenes con S3 y Cloudinary, con triggers de Postgres para logging y webhooks para uso interno y de clientes
- Construí pipelines ETL complejos para ingestar y unificar datos de data lakes MySQL de clientes en PostgreSQL, asegurando integración de datos fluida
- Diseñé workflows de generación de PDFs con templates y macros Jinja2, y construí extractores de datos de PDFs potenciados por LLMs usando OpenAI API para estructurar datos no estructurados
- Creé suites de tests comprehensivas con Pytest para testing unitario y de integración con más del 80% de cobertura
- Lideré un equipo pequeño para entregar features, gestionar releases e implementar workflows de CI/CD con GitHub Actions y Docker
- Apoyé el desarrollo de UI con React y TypeScript para colaboración full-stack

AI, Python, FastAPI, React, TypeScript, PostgreSQL, OpenSearch, AWS, Docker, GitHub Actions, Jinja2, Pytest, S3, Cloudinary, OpenAI

Ernst & Young (EY)
01/2022 - 12/2023

Desarrollador Ssr. Backend Python · Contractor

- Diseñé y mantuve 9 microservicios con FastAPI y MongoDB
- Implementé despliegue continuo usando Kubernetes a través de Azure DevOps
- Ejecuté migraciones de datos con PySpark, Pandas y Pymongo en Azure Databricks
- Desarrollé microservicios backend con ASP.NET (C#)

Python, FastAPI, MongoDB, Kubernetes, Azure, PySpark, Pandas, Databricks, ASP.NET, C#

Equifax
03/2021 - 12/2021

Desarrollador de Soluciones de Datos · Tiempo completo

- Construí aplicaciones de datos usando Python, HTML, CSS, JavaScript y SQL
- Gestioné un equipo de desarrollo de 15 miembros en GitHub
- Migré procesos ETL de servidor local a GCP

Python, JavaScript, SQL, GCP, HTML, CSS

Nerd Interactive
07/2020 - 02/2021

Analista de Ciencia de Datos · Tiempo completo

- Gestioné Big Data a través de SQL en BigQuery de Google Cloud Platform (GCP)
- Realicé análisis de datos en Python (Pandas) para crear KPIs
- Desarrollé REST APIs con Django

Python, Pandas, SQL, BigQuery, GCP, Django

HABILIDADES TÉCNICAS

AI & LLM: OpenAI API, Claude API, Gemini / VertexAI, AWS Bedrock, LangChain, LangGraph, Langfuse, RAG Systems, Prompt Engineering

Backend: Python, FastAPI, Django, Node.js, TypeScript, ASP.NET / C#

Frontend: React, TypeScript, HTML & CSS, Tailwind CSS

Ingeniería de Datos: Airflow, Spark / PySpark, Kafka, Databricks, BigQuery, ETL Pipelines

Cloud & DevOps: AWS, GCP, Azure, Docker, Kubernetes, GitHub Actions CI/CD

Bases de Datos: PostgreSQL, MongoDB, Redis, Opensearch

Testing & Monitoreo: Pytest / UnitTest, Playwright, Datadog, Grafana, CloudWatch

PROYECTOS & CLASES

Trampoline Gymnastics AI App - RAG-based
2025 · Buenos Aires, AR

App full-stack de gimnasia en trampolín con calculadora de dificultad y puntuación, más un asistente de AI que responde sobre el Código de Puntuación FIG usando RAG sobre los reglamentos y tool calls al motor de dificultad

AI, React, TypeScript, FastAPI, Gemini, pgvector, Langfuse

https://tra-app.clever-club.com/bot

Data Analysis Class
2022 · Buenos Aires, AR

Curso de análisis de datos dictado por mí en HUMAI

Python, Pandas

https://www.youtube.com/watch?v=V2t4brIm6xM

FastAPI Workshop
2025 · Buenos Aires, AR

Taller de FastAPI: ideal para estudiantes, profesionales e investigadores que quieran aprender con un ejemplo real

Python, FastAPI

https://youtu.be/EjY_KGjovIs

Tuclases.com Web Scraper

2022 · Buenos Aires, AR

Herramienta de web scraping automatizado para extracción de datos de plataforma educativa

Python, Automation

https://github.com/engcarlosperezmolero/tus_clases_automatic_response_scraper

Time of Fly Measurement ESP32 Device

2022 · Buenos Aires, AR

Dispositivo IoT para medición de tiempo de vuelo usando microcontrolador ESP32 con interfaz web

C++, JavaScript, HTML, CSS

https://www.linkedin.com/feed/update/urn:li:activity:6982465610228088832/

Ser Dev en Tiempos de IA

2025 · Buenos Aires, AR

Stream en vivo sobre las ventajas reales de ser dev en tiempos de IA, con temas que están marcando el presente del desarrollo

AI, Development

https://www.youtube.com/live/t8o2N41fo1I

EDUCACIÓN

Ingeniero Químico

02/2013 - 06/2019

Universidad del Zulia

Promedio: 18.66 / 20 (GPA equivalente: 9.33 / 10) | Graduado con honores

Pasantías en Indesca trabajando con un reactor piloto de producción de PVC. Tesis desarrollada usando software de simulación de procesos ASPEN y HYSYS.

IDIOMAS

Español: Nativo

Inglés: Avanzado | Profesional