

Gonzalo Martin Perez

AI Engineer | Software Engineer | Product Engineer

LLM Agents · Agentic AI · RAG · GraphRAG · LLM Evals | LangChain · LangGraph · OpenAI API · Claude API · MCP · pgvector · Python · TypeScript · FastAPI · Next.js · AWS

gonzalomartinperez2002@gmail.com | +54 9 2984 686913 | Bahía Blanca, Argentina · 100% Remote (UTC-3) · English B2
linkedin.com/in/gonzalo-martin-perez | github.com/gonzalomartinperez | gonzalomartinperez.com

PROFESSIONAL SUMMARY

AI Engineer with 2.5+ years shipping production software, most recently LLM agents, RAG systems and eval suites. Lead engineering at an early-stage fintech and blockchain startup whose generative interface opens digital assets to everyday people, and own its AI assistant end to end: the LangGraph agents in English, Spanish and Portuguese, the eval suite that gates every release, and the backend, web and mobile app around them. Previously built a GraphRAG assistant for Payway's merchant platform.

PROFESSIONAL EXPERIENCE

AI Engineer | Rampy

Sep 2026 – Present | United States · Remote

Leading engineering at an early-stage fintech and blockchain startup whose non-custodial app opens digital assets to everyday people through a generative interface. Working closely with the three founders and with specialists in blockchain and integrations.

- **Built the eval suite that gates every release of a multilingual AI assistant: 1,000+ scenarios replicated in English, Spanish and Portuguese (3,000+ cases), with the agents passing 90%+.** An LLM-as-a-judge in LangSmith, validated against human review, grades every run.
- **Migrated the assistant from Agno to a multi-agent LangGraph system** with cycles, financial tools, background workers, retries and failure recovery; Claude is the primary model, with DeepInfra and Vertex AI fallbacks, GraphRAG (Neo4j, pgvector) and Mem0 memory.
- **Made agent responses 2× faster and cut tokens by 40%** by removing redundant tool calls and tightening memory and context, verified with load tests and evals.
- **Designed the structured-output contract that turns agent answers into generative UI in the React Native app:** native buttons, cards and deep links into swap, vault and perps flows on live market data. Users confirm every transaction in a non-custodial Privy wallet.
- **Re-architected the inherited backend and frontend around Clean Architecture, cutting build-and-deploy from about 1 hour to 10 minutes.** Built the trading integrations and fee model on top: Hyperliquid perps (200+ markets), LI.FI swaps (100+ tokens) and Morpho, Aave and Compound vaults, tracked via webhooks.

Software Engineer | Teamcubation

Dec 2025 – Sep 2026 | Argentina · Remote

Consultant on Payway's merchant promotions platform (350,000+ merchants, 13 countries).

- **Built a GraphRAG assistant that passed 95%+ of a 500+ case eval suite**, graded in LangSmith by a human-validated LLM-as-a-judge. Used LangGraph, Neo4j and PostgreSQL/pgvector, and added guardrails against off-topic questions and data leaks.
- **Delivered the event-driven pipeline that ingested 20M+ historical promotions into production** (S3 and SQS into a Python/FastAPI Lambda, then Java/Spring Boot and NestJS services on Docker and Kubernetes). Wrote tooling to replay loads locally and reconcile failed records.

Software Engineer | Cooperativa Obrera

Dec 2024 – Dec 2025 | Bahía Blanca, Argentina · On-site

Primary engineer on the permissions back office of Argentina's largest consumer cooperative (2.7M+ members).

- **Integrated 10+ internal systems into one admin experience** through a Python/FastAPI BFF and a Next.js front end; a 10-minute manual permission lookup now takes under 10 seconds.
- **Cut permission tickets and manual assignments by 80%+ and recurring assignment bugs by 95%+** by standardizing how new systems are onboarded.

AI Engineer & Full-Stack Developer | Self-employed

Jan 2024 – Present | Argentina · Remote · Project-based

AI and full-stack projects for clients selling on Shopify and Hotmart; several are live in production.

- **Cut manual processes from hours or days to minutes** with WhatsApp assistants, n8n, ManyChat and Meta Ads automations, and MCP content agents on the Claude or OpenAI APIs.
- **Delivered websites, custom software and e-commerce projects in days or weeks that generated USD 10,000+ in client profit**, including Pequeverso's store and its GraphRAG shopping assistant on LangGraph.

TECHNICAL SKILLS

AI Engineering: AI Agents, Multi-Agent Systems, RAG, GraphRAG, LLM Evaluation, LLM-as-a-Judge, Tool Calling, Generative UI, Prompt Engineering, AI Guardrails

AI Frameworks & Platforms: LangGraph, LangChain, LangSmith (LLM Observability), OpenTelemetry, Claude API, OpenAI API, Vertex AI, MCP, Neo4j, pgvector

Backend & Data: FastAPI, NestJS, Spring Boot, Laravel, REST APIs, Microservices, Clean Architecture, Event-Driven Architecture, PostgreSQL, Redis

Languages & Frontend: Python, TypeScript, Java, SQL, Node.js, React, Next.js, React Native, Kotlin, Swift

Cloud & Delivery: AWS Lambda, Amazon S3, Amazon SQS, Docker, Kubernetes, GitHub Actions, SSE

Fintech & DeFi Integrations: Hyperliquid, LI.FI, Morpho, Aave, Compound, Privy, Non-Custodial Wallets

SELECTED PROJECTS

Portfolio Assistant · Bilingual RAG Assistant with Evals

Sep 2026 – Present

- **Open-source bilingual RAG assistant that answers recruiters' questions about my work:** FastAPI and LangGraph with hybrid lexical and vector retrieval plus GraphRAG (pgvector, Neo4j), cited answers streamed over SSE and per-request cost limits. Implemented and evaluated; not yet deployed.
- **Built a 1,000-question bilingual eval bank plus adversarial cases**, a 244-turn live-model evaluation on OpenAI and 800+ automated tests in GitHub Actions CI, with OpenTelemetry tracing.

github.com/gonzalomartinperez/portfolio-assistant-api

Filomena · Production Exam Platform for Health Sciences

Jan 2025 – Dec 2025

Capstone project now used in production by **five national institutions in Argentina** for health sciences exams, serving 1,000+ concurrent users. Principal contributor on a three-person team (Laravel, Next.js, Redis); graded **10/10**.

gonzalomartinperez.com/work/filomena

EDUCATION

Information Systems Engineer (Software Engineering equivalent) | Universidad Nacional del Sur

Jan 2020 – Oct 2025 | Bahía Blanca, Argentina | Average grade: 8.67 / 10