

Lucas Pignataro

AI AGENT & FULL-STACK ENGINEER

lucaspignataro@proton.me · +54 9 2241 49-1678 · Buenos Aires, Argentina (UTC-3)
linkedin.com/in/lucas-pignataro-aa7747207 · github.com/LukiPigna · Open to remote, worldwide

I ship products and the LLM systems inside them: a case-file platform a paying client runs in production, a desktop study app with a metered AI tier, and a self-hosted knowledge base served to agents over MCP — along with the Postgres, Docker and networking underneath them. Electronics engineering background, fluent English, remote-first.

SKILLS

LLM & agents	Claude Agent SDK · MCP servers and tools · tool/function calling · RAG and semantic search · vector embeddings (bge-m3) · local inference (Ollama) · prompt caching · multi-provider (Anthropic / OpenAI / Google)
Code	TypeScript · Node.js · Rust · C · Python · SQL
Data & backend	PostgreSQL · pgvector · Supabase (RLS, auth, migrations) · Redis · REST · SSE streaming
Frontend	Next.js · React · PWA / offline-first · OAuth · role-based access control
Infrastructure	Linux · Docker Compose · Caddy (reverse proxy, TLS) · systemd · WireGuard · Matrix/Synapse · LiveKit (WebRTC) · Vercel · GitHub Actions · NVIDIA GPU inference

EXPERIENCE

Co-founder & Software Engineer, Unit Dec 2025 – Present · Remote

- Two-person software studio, co-founded with Manuel Corcos. We design, ship and operate everything below.
- **ArgenINTA — paying client, in production** — digital case-file system replacing a paper-and-email workflow at an agricultural research foundation. Sold as a fixed-price build plus a monthly maintenance retainer; piloted with staff for five months, now live with real users. LLM-based PDF extraction pre-fills forms, role-separated internal/external views, policy validation rules (Next.js + Supabase; cut database latency ~26% via region migration).
 - **Planiar** (Rust + Tauri) — desktop app for university students: subjects, deadlines and timed study sessions, plus an AI assistant that reads the student's own course material to answer questions and prepare exams. 70+ public releases (v0.11) with signed installers and auto-update; subscriptions billed through Mercado Pago; free tier on the student's own API key, paid tiers with metered AI usage across Anthropic, OpenAI and Google models. Cut AI cost per session via fragment retrieval and text caching.
 - **gbrain** — company knowledge base with semantic search over docs, meeting transcripts, notes and code, exposed to LLM clients as an **MCP server** (Postgres + pgvector, bge-m3 embeddings served locally on a consumer GPU: no company data leaves the server). A persistent agent backend on the **Claude Agent SDK** brokers tool access to it over loopback so only the authenticated agent endpoint is exposed.
 - **Self-hosted infrastructure** — operate the full stack the above runs on: Caddy with internal TLS, DNS, WireGuard VPN, federated Matrix/Synapse, LiveKit and coturn video calls, Postgres, automated backups, and hourly health checks across ten services.

PROJECTS

- Health-data pipeline ingesting iPhone exports through a self-hosted endpoint, surfacing recommendations corrected against rolling medians · an ESP32 voice-command device in a parametric 3D-printed enclosure.

EDUCATION

B.S. Electronics Engineering, University of Buenos Aires (UBA) · In progress. Embedded systems, digital design, computer architecture, C programming.

LANGUAGES

Spanish (native) · English (fluent, technical discussion and interviews)