

JOAQUÍN CUBERO

Senior AI Engineer — Agentic & Generative AI Systems

jcubero@gmail.com • +506 8689-5526 • linkedin.com/in/joaquin-cubero • Costa Rica (Remote)

12

AI Products in Production

50+

Agentic Workflows Orchestrated

30–50%

Cost Reductions via Prompt/Inference Optim.

99.9%

Uptime on AI Platforms

ROLE FIT — SENIOR AI ENGINEER

- ▶ Agentic AI architectures (multi-agent, planning, tool-use, memory) with LLMs as first-class system actors
- ▶ Context engineering beyond prompts: long-term memory, embeddings, retrieval, structured context pipelines
- ▶ Observability & evals: golden sets, guardrails, prompt/versioning telemetry, continuous improvement loops
- ▶ Self-adapting systems: feedback-driven refinements, budget-aware policies, auto-scaling + cost governance
- ▶ AI-driven code generation/orchestration for accelerated delivery, testing and refactoring
- ▶ Latency/throughput & token-cost optimization across inference and agent execution graphs
- ▶ Progressive delegation: evolve from human-in-the-loop to agent autonomy with safety gates and oversight

EXPERIENCE

Founder & Chief Technology Officer (CTO)

Jan 2026 – Present

NomadaLLM — Edge AI SDK (Remote · San José, Costa Rica)

- ▶ Architecting the infrastructure for the **Internet of Consciousness (IoC)**: high-performance on-device agentic AI
- ▶ Built **NomadaLLM.reason()** engine delivering < 50ms local inference latency without internet dependency
- ▶ Engineered **air-gap certified** framework enabling total data sovereignty (Healthcare, Finance, Defense)
- ▶ Shifted AI from variable token expenses to **fixed infrastructure assets** for enterprise ROI
- ▶ Standardized developer experience via **GitHub Copilot SDK** and **Model Context Protocol (MCP)**

Stack: Python/C++, ONNX Runtime, TensorRT, Raspberry Pi 5, Edge GPUs, Local LLMs, MCP, Copilot SDK, Docker

Edge AI Data Sovereignty SDK

AI Developer

Oct 2025 – Jan 2026

Paradigm Store — AI Multi-Tenant Commerce Platform (Remote · San José, Costa Rica)

- ▶ Architected and deployed production multi-tenant e-commerce AI platform with **12 specialized agents**
- ▶ Built **LangChainOrchestrator** using AgentExecutor, tool adapters and prompt chains
- ▶ Designed **Marketing Performance Agent** with MemGPT memory and scientific-method optimization loops
- ▶ Implemented dual memory: **MemGPT** (working/archival) and **GilgameshAI** (decay, contextual categorization)
- ▶ Developed multimodal computer vision: **Gemini 2.5 Pro** (face recognition, product image search, layout extraction), self-hosted **LLaVA**, **YOLO** fallback
- ▶ Delivered WhatsApp customer service agent with NL intent detection; built payment agent (Stripe, BAC Credomatic, PayPal, SINPE, 13% IVA)
- ▶ Shipped catalog/search (2,500+ products), admin dashboard, marketing automation (BigQuery, Mailchimp 50K+ contacts), and full lifecycle agents
- ▶ Lifecycle agents: **order management, shipping, delivery, SEO, CRM, tax, and image processing**
- ▶ Deployed on DigitalOcean/Contabo VPS with **99.9%+** uptime, Docker, Nginx, Prometheus
- ▶ Integrated **REST** and **GraphQL** APIs for payment, CRM, inventory, and order orchestration

Stack: Python, FastAPI, LangChain, MemGPT, Vector DBs, Gemini 2.5 Pro, LLaVA, YOLO, WhatsApp API, Stripe/PayPal/SINPE, Docker, Nginx, Prometheus

Senior AI Developer & Tech Lead

Dec 2024 – Oct 2025

Axos Bank — Enterprise FinTech (Remote)

- ▶ Designed and implemented Azure AI solutions using **Azure OpenAI (GPT-4/4o)** for fraud detection and customer service
- ▶ Developed and deployed ML models with **Azure Machine Learning** for fraud detection
- ▶ Built automated customer service with **RAG** architecture on Azure
- ▶ Implemented **Azure AI Search** for intelligent retrieval and knowledge base management
- ▶ Used **Azure Document Intelligence** for automated document processing and extraction
- ▶ Led Azure cloud-native architecture; deployed via **Functions, App Service**, and **AKS**
- ▶ Integrated **Cosmos DB**, **Azure Storage** and **Queues** for distributed systems
- ▶ Implemented NLP evaluation metrics (**BLEU**, **ROUGE**) and human evaluation frameworks
- ▶ Integrated **LangChain.NET Memory Manager** for persistent context in .NET apps
- ▶ Managed AI development with **Cursor** and **Windsurf** in Azure DevOps; reduced bugs and improved performance
- ▶ Developed **AI Fraud Agent** for real-time fraud detection

Stack: Azure OpenAI (GPT-4/4o), Azure ML, AI Search, Document Intelligence, Functions, App Service, AKS, Cosmos DB, Storage/Queues, LangChain.NET, RAG, Azure DevOps

Compliance RAG Throughput/Latency

AI Engineer & Technical Lead Aug 2021 – Jul 2024

Magaya Corporation — Supply Chain (Remote)

- ▷ Implemented custom features and RESTful integrations with .NET backend services for logistics workflows
- ▷ Coordinated with Autonoma to deploy **AI-driven E2E testing** and automated test generation
- ▷ Collaborated with UI/UX and product to refine accessible, consistent mobile experiences
- ▷ Mentored junior developers and improved development practices; integrated cloud services and APIs for sync

Stack: .NET, REST APIs, Azure DevOps

Mobile .NET MAUI Quality & E2E

Senior Full-Stack Engineer (Freelance)

Feb 2015 – Aug 2021

InteliSoft Solutions — Global Clients (Remote)

- ▷ Built **Tributación Directa** and **Factura Electrónica** mobile apps
- ▷ Created **UberWash** (Xamarin.Forms) for on-demand car washes, Uber-like UX
- ▷ Developed **Star Wars** fan information app with cross-platform optimized UI
- ▷ Shipped native iOS app for **INS Costa Rica** using Swift for emergency dispatch workflows
- ▷ Implemented chat iOS apps for **Zermat** and **Teledolar** using Telegram OSS for real-time support

Stack: Xamarin.Forms, Swift/Xcode, TypeScript/React, PostgreSQL/MongoDB, Telegram, REST

Mobile Engineer

May 2020 – Dec 2020

YPrime — Clinical Trials (Remote)

- ▷ Developed **YPBYOD** (Xamarin Native iOS/Android) for subject/caregiver clinical study data collection
- ▷ Implemented secure storage and transmission meeting clinical compliance requirements
- ▷ Drove UI/UX improvements, rigorous testing and platform reliability across iOS and Android

Stack: Xamarin.iOS, Xamarin.Android, Secure Storage, REST

Senior Xamarin.Forms Developer

Jan 2020 – Apr 2020

Proto-brz — Remote

- ▷ Built **Frebo** with Xamarin.Forms/iOS/Android; optimized cross-platform performance
- ▷ Delivered tracking and community features; continuous UX iteration with stakeholder feedback

Stack: Xamarin.Forms, Xamarin.iOS, Xamarin.Android

Senior Xamarin.Forms

Jun 2018 – Jan 2020

BIS S.A. — Remote

- ▷ Developed **Mercaderistas** (iOS/Android) for supermarket reviewers; cross-platform sync and UI optimization
- ▷ Built **Factur-e** for electronic billing to **Hacienda** with comprehensive reporting

Stack: Xamarin.Forms, Xamarin.iOS/Android

Sr Mobile Developer

Jan 2014 – Feb 2015

PwC — Mexico (Remote)

- ▷ iOS apps for consulting Mexican tax laws; large-dataset performance, secure sync, and up-to-date content

Stack: iOS (Swift/Xcode), REST, Data Sync

Mobile Developer

Feb 2014 – Jul 2014

Jiffy Software Solutions — Remote

- ▷ Maintained and enhanced **FilmTastic** (iOS); integrated Facebook AudioTool; improved media sharing UX

Stack: iOS (Xcode), Facebook AudioTool, Performance Tuning

Freelancer — iOS (Elance/oDesk)

Jun 2012 – Feb 2014

Elance / oDesk — Remote

- ▷ **Cooking Plan it** v2.0 (iPhone/iPad); **iBirds Journal** fixes; **Pushwood** maps; **RPE Mobile** POS integration

Stack: iOS (Objective-C), Core Data, REST/JSON, Google Maps API

EARLIER EXPERIENCE (SELECTED)

- ▷ **Number8** — Senior Developer (Contract) at Practice Law Institute of New York (Feb 2014 – Feb 2020): .NET apps, performance tuning, ticket resolution
- ▷ **nVision, LLC** — Senior Developer (May 2008 – Jun 2010): Inventory systems, internal frameworks, coding standards
- ▷ **Govansoft** — Developer (Nov 2008 – May 2009): VS2008/C#/SQL/Reporting Services
- ▷ **Fiserv / Fiserv CBS Costa Rica** — VS.NET Developer/Senior Developer (Feb 2005 – May 2008): E-Banking and teller apps for Bank of America
- ▷ **Proyectica** — Developer (Feb 2004 – May 2008): ASP.NET, VB.NET, XML, UML
- ▷ **Zona Franca Metropolitana** — Developer (Jan 2003 – Feb 2004): ASP.NET/VB.NET, Solaris/Apache systems

PROJECTS & DEMOS (AGENTIC/GENERATIVE)



Nomada Lucid — Dual-Agent System

Planner/Coach agents with memory, RAG, and speech I/O for anxiety support and guided flows.

lucid.nomadahealth.com



NomadaLLM — Embedded LLM SDK

Privacy-first SDK with prompt tooling, eval harness, cost controls, and pluggable vector backends.

nomadallm.nomadahealth.com



HelloStore AI — Agentic Commerce

Shopping assistant with recommendation engine, order orchestration, and retrieval-enhanced chat.

TECHNICAL SKILLS

[LLMs](#) [GPT-4/4o](#) [Claude](#) [Mistral](#) [Frameworks](#)
[LangChain](#) [Llamaindex](#) [Custom Agents](#) [Vector DBs](#)
[Pinecone](#) [Weaviate](#) [FAISS](#) [Graphs](#) [Knowledge Graphs](#)
[Infra](#) [Ray](#) [ONNX Runtime](#) [NVIDIA Triton](#)
[Cloud](#) [Azure](#) [AWS](#) [GCP](#) [APIs](#) [REST](#) [Event-Driven](#)
[Microservices](#) [Languages](#) [Python](#) [TypeScript](#) [Java](#)
[Go](#)

- ▷ Context engineering: hierarchical memory, embeddings, retrieval policies, session/state stores
- ▷ Agent patterns: planner/worker/reviewer, toolformer-style adapters, scratchpads, task graphs
- ▷ Observability: eval datasets, drift checks, prompt/version telemetry, red-teaming & guardrails
- ▷ Efficiency: token pruning, caching, model routing, streaming, small-model fallbacks

EDUCATION & CERTIFICATIONS

MBA in AI for Business (In Progress) —
Panamerican Business School (2026)

B.Sc. in Systems Engineering — Universidad
Internacional de las Américas (UIA)

Available immediately • 100% Remote • Feb 2026