Romeu Palos de Gouvea

Chief AI Officer · Agentic AI & RAG · Commerce & Payments Platforms

romeu@gouvea.us

SUMMARY

Technology executive with 15 years of experience across the full stack, from FPGA and PCB design to planet-scale e-commerce platforms and applied AI. Founding team member and Chief AI Officer at Firmly, where I architected the agentic-commerce platform launched with partners including Meta, PayPal and Perplexity. Previously one of the principal architects of Samsung.com US e-commerce during its growth from $50M to $2.4B in gross revenue. Hands-on leader who still designs the systems, builds the agents and writes production code.

PROFESSIONAL EXPERIENCE

Firmly Inc.

Mar 2022 – Present

Chief AI Officer

Mar 2025 – Present

- Own company-wide AI strategy, implementation and governance; set the AI research and development roadmap and led the company through enterprise AI-governance reviews with global payment networks.
- Architected a retrieval-augmented generation (RAG) pipeline that crawls, chunks and indexes merchant sites and catalogs into vector and hybrid search, grounding agent answers in merchant content with source citations.
- Built Ask Agent, a conversational shopping assistant embedded on merchant sites, with tool calling for catalog search, cart building and in-chat checkout, plus voice and multilingual conversations.
- Designed a multi-agent system that autonomously investigates a merchant’s e-commerce platform, generates the integration adapter, tests it and monitors it over time, cutting merchant onboarding to self-service.
- Built MCP servers and tool interfaces that expose commerce capabilities (discovery, cart, checkout) to third-party AI agents, supporting agentic-commerce protocols such as ACP and UCP.
- Lead model selection and multi-provider LLM orchestration (Gemini, Claude), prompt engineering, agent guardrails (scope limits, loop guards) and model change management.
- R&D: fine-tuning of conversational speech models (full fine-tuning and PEFT) on multi-node H100 GPU clusters.

VP of Engineering

Mar 2022 – Mar 2025

- Founding team member; architected the core platform and integrations behind Firmly’s launches with Meta, PayPal and Perplexity.
- Established and scaled an offshore engineering subsidiary, from hiring and legal setup to delivery culture.
- Defined foundational processes across engineering, product, HR and finance, built for lean, low-overhead operations.

Samsung Electronics America

Oct 2019 – Sep 2022

Principal Systems Architect

- Defined and evolved the architecture of the Samsung.com US e-commerce platform during gross revenue growth from $50M to $2.4B.
- Led design across order management, fulfillment, operations, DevOps, security and data warehouse pipelines.
- Solved major scaling challenges; the platform was adopted globally across Samsung as a best-practice reference.

Samsung R&D Center Brazil

Jul 2013 – Sep 2019

Senior Software Engineer → Technical Architect

- Owned fulfillment and carrier integration for Samsung.com; architected integrations with the four major US carriers with a focus on data privacy, security and scalability.
- Built parts of the Samsung ecosystem, from Android base services with hardware control and I/O (TV decoder) to customer-service Android/iOS apps and the global Samsung.com web and mobile storefronts.

EARLY CAREER: HARDWARE & EMBEDDED ENGINEERING

CPqD, Telecommunications R&D Center (Brazil)

Jan 2013 – Jul 2013

Electrical Engineer

- Developed an LTE Layer-1 transceiver on an Altera Cyclone FPGA; implemented Turbo decoder, FFT and PRACH in VHDL and managed radio-unit control registers.

Wimobilis Digital Technologies (Brazil)

Jul 2011 – Dec 2012

Electrical Engineer / Technical Lead

- Led an ultra-low-latency, low-bandwidth broadcast radio link between TV stations and mobile studio units using Speex and Fraunhofer codecs.
- Designed a PCIe H.264 (SBTVD) video-encoder board on a Xilinx Virtex with MicroBlaze: 10 Gbps signaling, impedance and length matching, and termination.

PEET, University Junior Enterprise (Brazil)

Jul 2009 – Jul 2011

Financial Officer

- Ran accounting, finance and legal; grew revenue and headcount by winning automation projects in the Poços de Caldas region.

EDUCATION

M.S. Electrical Engineering (coursework), UNICAMP – DECOM/FEEC

2011 – 2013

- Research: partitioning a video encoder across heterogeneous co-processors (CPU + FPGA).

B.S. Electrical Engineering, Telecommunications & Automation, PUC Minas (Brazil)

AI COMPETENCIES

- **Agentic systems:** multi-agent orchestration, tool/function calling, autonomous code-generating agents, sandboxed execution, guardrails and loop control
- **Retrieval & grounding:** RAG pipelines, web crawling and ingestion, chunking, embeddings, vector and hybrid search, citation-grounded answers
- **Agent interoperability:** Model Context Protocol (MCP) servers and apps, agentic-commerce protocols (ACP, UCP), agent payments
- **LLM engineering:** prompt engineering, multi-provider orchestration (Gemini/Vertex AI, Claude), structured output, streaming and voice agents
- **Model training:** PyTorch, Hugging Face Transformers & Diffusers, full fine-tuning and PEFT/LoRA, distributed multi-GPU training (H100), TensorBoard
- **AI governance:** enterprise AI risk questionnaires, data scoping and zero-retention considerations, model change management, production monitoring

ENGINEERING SKILLS

- **Expert:** Python (ML training, agents, data pipelines) · JavaScript/TypeScript & Node.js (agent services, MCP servers, commerce APIs) · Java (Samsung.com commerce services) · C/C++ (Android system services, hardware control, performance tuning)
- **Hardware description:** VHDL, Verilog: LTE Layer-1 and H.264 encoder cores on Xilinx/Altera FPGAs
- **Specialties:** distributed commerce and payments platforms, checkout and cart architecture, high-scale e-commerce, embedded and real-time systems
- **Cloud & platform:** Cloudflare Workers, Durable Objects, D1, R2, Workflows and Zero Trust; Kubernetes; CI/CD and DevOps; React; SvelteKit
- **Data:** PostgreSQL, CockroachDB, Cassandra, RethinkDB, Elastic Stack (Elasticsearch, Logstash, Filebeat), vector indexes
- **Systems & tooling:** Linux, Android, ROS, FreeRTOS; Valgrind, KCachegrind, QEMU, strace; Playwright, Vitest
- **Hardware design:** high-speed PCB design (10 Gbps, impedance matching): Altium Designer, KiCad, Eagle
- **Languages (spoken):** English, Portuguese (native)