

## VICTOR RAMIREZ

Greenbrae, CA • (415) 299-4692 • [VRamirez75@gmail.com](mailto:VRamirez75@gmail.com)

LinkedIn: [linkedin.com/in/victor-hugo-ramirez-mids](https://linkedin.com/in/victor-hugo-ramirez-mids)

Portfolio: [vramirez-mids.github.io](https://vramirez-mids.github.io) • GitHub: [github.com/vhr1975](https://github.com/vhr1975) • Medium: [medium.com/@vhr1975](https://medium.com/@vhr1975)

## SUMMARY

Developer Experience Engineer & AI Architect with 15+ years building developer tools, cloud platforms, AI prototypes, onboarding workflows, and technical education programs. UC Berkeley MIDS graduate with hands-on expertise in RAG, embeddings, LLM APIs, multimodal systems, function calling, and AI agents. Creator of multiple open-source AI demos, instructor for community AI workshops, conference speaker, and monthly AI podcast host.

## DEVELOPER EXPERIENCE PORTFOLIO

### OpenAI Tutorials Showcase

🔗 <https://github.com/vhr1975/OpenAI-Tutorials>

Full-stack, production-ready demos with onboarding guides, architecture diagrams, and extensibility patterns.

### Function Calling Demo — FastAPI + React

🔗 <https://github.com/vhr1975/function-calling-demo>

- Multi-tool orchestration with retries and structured outputs
- Includes telemetry hooks and onboarding documentation

### RAG PDF Demo — FastAPI + React + ChromaDB

🔗 <https://github.com/vhr1975/rag-pdf-demo>

- Complete RAG pipeline (ingestion → embeddings → retrieval → synthesis)
- Demonstrates RAG patterns, chunking, and best practices

### Multimodal Vision Demo — GPT-4o Vision

🔗 <https://github.com/vhr1975/vision-qa-demo>

- Vision Q&A over screenshots, documents, and diagrams
- Showcases multimodal prompting + vector search integration

## AI Architecture Showcase

🔗 <https://github.com/vhr1975/ai-architecture-showcase>

- Maps traditional software architecture patterns (Layered, Event-Driven, Microservices, CQRS) to LLM + RAG systems

## AI EDUCATION, COMMUNITY WORK & CONTENT CREATION

### AI Builders: LatinX Edition — Creator & Host

🔗 <https://open.spotify.com/show/ai-builders-latinx>

- Monthly podcast explaining AI concepts and interviewing LatinX technologists

### Techqueria — AI Instructor & Curriculum Author

- Teach RAG, embeddings, prompting, function calling, LLM fundamentals
- Authored open-source AI education materials:
  - ↳ <https://github.com/techqueria/mult-docs-rag-llama>
  - ↳ <https://github.com/techqueria/GenAI-Tutorial>
  - ↳ <https://github.com/techqueria/ai-para-todos>

### Medium (Technical Writing)

🔗 <https://medium.com/@vhr1975>

- Monthly articles on LLM workflows, AI architectures, and developer tooling

## PROFESSIONAL EXPERIENCE

### Moody's Analytics — Director, Developer & Platform Experience

#### 2024–Present | San Francisco, CA

- Lead developer experience strategy across cloud + AI platforms
- Build onboarding guides, CI/CD templates, architecture diagrams, troubleshooting flows

- Enable engineering teams adopting LLM-based workflows

### **Moody's Analytics — AI Architect**

#### **2023–2024 | San Francisco, CA**

- Taught teams RAG, embeddings, prompting, LLM evaluation, and AI architectural patterns
- Authored conceptual diagrams and documentation for AI adoption
- Supported LLM prototype development

### **Moody's Analytics — Associate Director, DevOps Architect**

#### **2015–2023 | San Francisco & New York**

- Built enterprise CI/CD pipelines, cloud standards, developer tooling, platform documentation
- Improved developer velocity through automation & cloud modernization

### **Moody's Analytics — Senior DevOps Engineer**

#### **2009–2015 | San Francisco, CA**

- Built automation workflows, monitoring systems, and early cloud infrastructure

## **EDUCATION**

**UC Berkeley** — Master of Information and Data Science (MIDS)

**CSU Monterey Bay** — B.S. Computer Science

## **TECHNICAL SKILLS**

**AI/LLMs:** GPT-4o, GPT-4, embeddings, RAG, function calling, multimodal, vector DBs (Chroma, FAISS)

**Languages:** Python, Java, C#, C++, SQL, TypeScript

**Frameworks:** FastAPI, React, Next.js, Node.js

**Cloud/DevOps:** AWS, Azure, Docker, Kubernetes, Terraform, GitHub Actions

**ML Tools:** Pandas, NumPy, Scikit-learn

**Developer Tools:** Jupyter, Postman, Mermaid, GitHub, Confluence

## **Conference Speaking**

- **Techqueria Tech Summit (Oct 2025):** *Building AI Agents for Real-World Workflows*

- **Oakland Tech Week (Nov 2025):** *AI Agents from Scratch: A Developer's Guide*
- **Latino AI Summit (Jan 2026 – Scheduled):** *RAG Chat + AI Agents*