Enrique Rodriguez Garrido

San Francisco, Bay Area

enrique.rodriguez.garrido@gmail.com
https://www.linkedin.com/in/enriquerodriguezgarrido
https://github.com/ensamblador

Professional goal:

For my next step, I seek to lead enterprise-scale Generative AI solutions that transform businesses through innovative, responsible implementation.

Summary

Senior Specialist Solutions Architect at AWS with deep expertise in enterprise Generative AI and cloud architectures. Demonstrated success in architecting and deploying production-grade GenAI solutions for cross-industry organizations, leveraging Amazon Bedrock, Foundation Models, RAG patterns, and multi-agent systems. Combined with a Master's in Data Science and hands-on experience implementing Foundations Models at scale. Passionate technical educator with proven track record in mentoring teams and enabling customer success.

Key Achievements:

- 10+ years pre-sales/architect enterprise cloud solutions as trusted technical advisor to C-suite and engineering teams.
- Led 60+ enterprises through complex cloud adoption journeys, from discovery to production deployment
- Expert in LLM integration patterns, evaluation frameworks, and scalable cloud architectures
- Passionate technical educator: regular speaker at major events (400-1500+ attendees), extensive content creation
- Deep expertise with LLM frameworks (LangChain, agents), Python development, and MLOps best practices

Core Technical Focus: Development & Implementation:

- Python Expert: Production-grade development for GenAI applications, API integrations, and ML workflows
- Proficient in LLM frameworks: LangChain, LangGraph, Strands, Anthropic SDK, Amazon Bedrock SDK, MCP
- Full-stack development: Python backend, JavaScript/TypeScript frontend
- Experience with: API Gateway, Streamlit, async programming, testing frameworks

Experience

Generative AI Experience

Enterprise Solutions & Architecture

- Led architecture and deployment of production Generative AI solutions top Latam Companies, integrating foundation models with RAG patterns, vector databases, and intelligent agent workflows.
- Recently with focus on applying AI to Customer Experience in the Contact Center space.
- Designed and implemented AI-powered legal assistant enabling instant analysis of court cases, reducing report generation time from days to 2 minutes. Implemented evaluation frameworks for LLM performance measurement, including accuracy metrics, safety benchmarks, and business alignment.
- Architected multi-agent collaboration systems with human escalation capabilities for customer service applications, including Chat and Voice integrations. This proves to be a responsible GenAI implementation by including human in the loop.
- Developed reference architectures and reusable solutions for enterprise-scale GenAI implementations using infraestructure as code, showcasing RAG patterns, LLM optimization, and agent orchestration

Technical Leadership & Innovation

- Created comprehensive solutions demonstrating GenAI implementation patterns:
  - WhatsApp Integration with AI Agents for automated customer service
  - Contact Center solution featuring AI agents with human escalation workflows
  - Full-stack Generative AI Chat Application showcasing RAG, LLM optimization, and agent capabilities
- Guide 60+ customers in their AI/ML transformation journey, focusing on responsible and effective GenAI implementation.
- Regular speaker at major AWS and AI events (400-1500+ attendees), delivering insights on enterprise GenAI adoption and best practices

Research & Academic Contributions

- Completed Master's Defree program on GPT-2 implementation for political tweet generation.
- Active contributor to AI/ML community through technical content creation, including blog series on Generative AI assistants
- Continuous professional development through advanced courses in LLM applications, prompt engineering, and AI agents
- Expertise in latest GenAI technologies including RAG implementations, multi-modal AI, and agentic systems, RLHF, LLM Reasoning.

Notable Projects

- Led architecture of AI-powered legal Assistant Case Study for Legaltech Startup, enabling instant analysis of 6M+ court cases and reducing report generation time from days to 2 minutes for 25K+ active users. **Technologies:** Python, LangChain, AWS Bedrock, Claude, OpenSearch, Lambda, DynamoDB
- WhatsApp Integration: Developed Bedrock Agents integration for automated customer service
- Multi-Agent System: Architected contact center solution with AI agents and human escalation
- Generative AI Chat Application: Built full-stack implementation showcasing RAG, LLM optimization and Agents.

Responsible AI Leadership

- Advocate for safe AI deployment with focus on human-in-the-loop architectures, explainability and Fairness. Aligned with business values.
- Impleneted guardrail systems for preventing hallucinations and ensuring factual accuracy in legal and financial applications
- SME speaker on AI ethics, bias mitigation, and responsible GenAI adoption

Specialist Solutions Architect, Applied AI

June 2025 - Present

Serving as a technical domain expert within AWS's Worldwide Specialist Organization, with focus in Amazon Connect Partners across the region to architect scalable, AI-powered customer experience solutions that transform legacy contact centers.

Senior Solutions Architect ISV @ AWS

Jan 2020 - June 2025

Lead technical strategist for 25 strategic Independent Software Vendor (ISV) customers, with primary focus on enterprise-scale Generative AI implementations and cloud-native solutions. Architect and deliver production-grade solutions leveraging AWS Services

Key Achievements:

- Lead the architecture and implementation of an AI-powered Legal Assistant that processes 6M+ court cases, reducing analysis time from days to minutes while serving 25K+ active users
- Designed and implemented embedded analytics solutions using Amazon QuickSight across multiple enterprises, resulting in: 45% reduction in operational costs, 50-66% decrease in development time, creation of 700+ client dashboards, Real-time ESG and safety analytics for high-risk industries, Integrated financial insights for 300+ businesses.
- **AWS Pathfinder Award (2021)**. Recognized for innovative problem-solving and pioneering solutions that enhanced efficiency and drove business growth.
- Promoted to **Senior Solution Architect** in 2022

Public Speaking & Thought Leadership:

- IADevs 2024 @ Universidad del Desarrollo
- Generative AI Roadshow Argentina 2024 (400+ attendees)
- Cloud Experience 2023 Chile (1300+ attendees)
- Cloud Experience 2023 Argentina (1500+ attendees)
- Chile Partner Summit 2022 (1100+ attendees)
- Regular presenter on enterprise Generative AI implementation strategies and best practices

AI / ML Tecnical Field Community (TFC) @ AWS

Mar 2021 - Present

Lead technical advisor specializing in enterprise AI/ML implementations across LATAM, guiding over 60 organizations from ideation to production. Focus on delivering practical ML solutions while advancing

Generative AI adoption through community building and knowledge sharing.

Key Achievements:

- Helped over more than 30 organizations in implementing end-to-end ML solutions, from problem definition to production deployment
- Early adoption of Large Language Models and Generative AI applications in the region, resulting in successful enterprise implementation

Technical Leadership & Content Creation:

- Authored comprehensive blog series on Generative AI assistants
- Created and delivered hands-on workshops on AI/ML technologies including NLP, LLMs, and MLOps
- Featured speaker on AWS Spanish Podcast discussing critical AI topics: ***Bias and Explainability in Machine Learning Models***, ***Artificial Intelligence for Contact Centers***
- Developed technical content and reference architectures for ML implementation patterns

Generated Content:

- [Blog Serie: Asistente IA Generativa](#)
- [Entregue experiencias naturales y eficaces de servicio al cliente de Telefonía Móvil con Amazon Connect y Amazon Lex](#)
- [Podcast AI/ML en AWS: Sesgo y explicabilidad en modelos de machine learning](#)
- [AI/ML en AWS: Inteligencia Artificial para Contact Centers](#) -[Webinar: Despliegue de un Contact Center en la nube con Amazon Connect en 30 minutos](#)

Manager Presales @ Telefónica

Mar 2016 - Dec 2019 (3 years 10 months)

- Managed 14-person Presales Engineering team, optimizing solution design and cost estimation
- Developed self-service quoting tools to scale sales operations
- Mentored team members resulting in 3 IC promoted.

Senior Presales Consultant @ Telefónica

Sep 2014 - Mar 2016 (1 year 7 months)

- Provided technical advisory and solution architecture for enterprise clients
- Created standardized proposal templates and built internal knowledge base

Presales Engineer @ Telefónica

Apr 2011 - Sep 2014 (3 years 6 months)

- Designed technical solutions aligned with customer requirements and budgets
- Automated evaluation processes through VBA, improving team efficiency

Service Manager @ Telefónica

Oct 2008 - Apr 2011 (2 years 7 months)

- Managed continuous service improvement for key enterprise accounts

Education

Universidad del Desarrollo

- Master Degree in Data Science (2019 - 2021)
- Degree Project: ***This Politician Does Not Exist: Using GPT-2 for Political Tweet Generation*** (Public Repo)

Universidad de Santiago de Chile

- Bachelor: Electrical Engineer, Electronic Systems and Telecommunications (1998 - 2006)
- President of the Student Branch of the Institute of Electrical and Electronic Engineers (IEEE)

Certifications and Courses

AWS Certifications:

- AWS Certified Generative AI Developer - Professional
- AWS Certified AI Practitioner
- AWS Certified Machine Learning – Specialty
- AWS Generative AI Foundational
- AWS Certified Solutions Architect - Professional
- AWS Certified Solutions Architect – Associate
- AWS Certified DevOps Engineer - Professional
- AWS Certified SysOps Administrator – Associate
- AWS Certified Developer - Associate
- AWS Certified Alexa Skill Builder – Specialty

Courses:

- Advanced Prompt Engineering for LLMs - Machine Learning University
- Building LLM Applications with Bedrock Agents - Machine Learning University
- Generative AI with Large Language Models
- Build LLM Apps with LangChain.js - DeepLearning.AI
- LangChain Chat with Your Data - DeepLearning.AI
- LangChain for LLM Application Development - DeepLearning.AI
- ChatGPT Prompt Engineering for Developers - DeepLearning.AI
- AI Agents in LangGraph - DeepLearning.AI
- Knowledge Graphs for RAG - DeepLearning.AI

Skills

Technical Expertise

Generative AI & LLMs:

- Foundation Models: Anthropic Claude (Opus, Sonnet, Haiku), Amazon Nova, GPT family
- LLM Frameworks: LangChain, LangGraph, LlamaIndex, Haystack, Strands, MCP
- Implementation Patterns: RAG, multi-agent systems, fine-tuning, prompt engineering (CoT, ReAct, ToT)
- Reasoning: Experience with reasoning models and advanced prompting techniques

Development & Engineering:

- Languages: Python, JavaScript/TypeScript, SQL
- Architecture: Microservices, event-driven, serverless, API-first design
- APIs: REST, GraphQL, WebSocket implementations

ML/AI Operations:

- Evaluation: Experience in frameworks for accuracy, safety, business alignment metrics
- Safety: Guardrails, content filtering, bias detection, explainability
- MLOps/FMOps: Model versioning, A/B testing, monitoring, continuous evaluation
- Methodologies: CRISP-DM, Well-Architected Framework

Cloud & Infrastructure:

- AWS Services: Bedrock, SageMaker, Lambda, Step Functions, EKS, OpenSearch, DynamoDB
- Vector Databases: pgvector, FAISS, ChromaDB, OpenSearch vector search, S3 Vector
- Observability: CloudWatch, model performance monitoring, cost optimization

Customer Success & Leadership:

- Pre-Sales: Technical discovery, PoC design, stakeholder management, complex sales cycles
- Communication: Executive briefings, technical deep-dives, public speaking (400-1500+ audiences)
- Content: Technical blogs, reference architectures, workshops, training materials
- Mentoring: Team leadership, knowledge transfer, community building