

MARCO ANTONIO GONZALEZ JUNIOR

m@rco.sh · +55 (11) 91577-7715 · São Paulo, Brazil

<https://linkedin.com/in/marcoantoniogonzalezjunior>

Solutions Architect | Cloud Engineer | Artificial Intelligence Engineer

PROFESSIONAL SUMMARY

Solutions Architect with 15+ years building scalable enterprise systems and 10+ years in AI/ML. Led technical teams delivering solutions that generated 30% revenue increases and 200% improvement in customer retention across fintech, retail, and digital marketing. Currently specializing in generative AI, fine-tuning foundation models, RAG systems, agentic architectures, and AI-assisted development workflows with MCP integrations.

SKILLS

Enterprise Architecture: Domain-Driven Design, Event-Driven Architecture, Microservices, Enterprise Integration Patterns

AI & Machine Learning: Machine Learning, Generative AI, LLMs, RAG Systems, Fine-tuning, Multi-Agent Systems, MLOps, Knowledge Graphs

Cloud & DevOps: AWS/Azure/GCP, Kubernetes, Docker, Terraform, CI/CD Pipelines, Infrastructure as Code

Leadership: Technical Team Management, Cross-functional Collaboration, Technology Strategy, Digital Transformation

EXPERIENCE

Capco

São Paulo

Solutions Architect - Artificial Intelligence

September 2024 – Present

Leading AI-driven legacy modernization initiatives using knowledge graphs and LLMs to transform enterprise COBOL systems.

- Architected knowledge graph-based modernization platform using Neo4j, LangGraph, and AWS Bedrock, processing legacy COBOL applications across multiple industries, reducing migration timelines by 60%
- Built production RAG systems with vector databases (Pinecone, Weaviate) and knowledge graphs (Neo4j) achieving 95% accuracy in business logic preservation across language translations from COBOL to Java/.NET
- Led team of 8 engineers implementing LLM orchestration frameworks with Python and TypeScript, enabling automated documentation of 20M+ lines of legacy code
- Delivered \$400k+ in cost savings using AWS cloud infrastructure and Kubernetes deployments, eliminating technical debt and reducing operational risk

IBM

São Paulo

Senior Solutions Architect

July 2022 – September 2023

Specialized in implementing IBM WatsonX AI solutions, fine-tuning foundation models, and architecting advanced RAG systems with continuous learning pipelines for enterprise financial services clients.

- Architected 10+ foundation model fine-tuning pipelines and RAG knowledge bases using WatsonX, PyTorch, and TensorFlow, delivering conversational AI solutions that achieved 85% automated resolution rate and improved customer satisfaction by 25%
- Designed multi-cloud MLOps infrastructure with automated continuous learning pipelines on AWS and Azure, supporting real-time model evaluation and retraining workflows that maintained 99.9% availability while processing 5M+ monthly interactions through fine-tuned conversational AI systems
- Implemented comprehensive model evaluation frameworks using BLEU/ROUGE scores, semantic similarity metrics, and hallucination detection systems, achieving 15% higher accuracy than generic models while reducing customer query handling time by 40% through intelligent context-aware response generation
- Implemented Docker-containerized model deployment strategies and IaC solutions using Terraform, enabling client teams to achieve 90% faster AI model deployment cycles and maintain 90%+ task completion rates

banQi

São Paulo

Senior Solutions Architect

June 2019 – July 2022

Led cross-functional technical architecture initiatives, ensuring system scalability and enterprise strategic alignment.

- Architected microservices platform using Docker, Kubernetes, and Go supporting 10x traffic growth, reducing system response times by 65%
- Led adoption of containerization across 15+ applications using Helm, ArgoCD, and GitOps workflows, achieving 99.95% uptime and reducing deployment times by 80%
- Managed technical team of 12 engineers implementing agile development practices with GitHub Actions CI/CD, increasing delivery velocity by 45%

Petlove
Tech Lead

São Paulo
May 2015 – June 2019

Led technical architecture for Brazil’s largest pet supplies e-commerce platform, driving digital transformation initiatives.

- Architected prime membership platform using Ruby on Rails, Elixir, and PostgreSQL, achieving 30% of monthly revenue (\$2M+) and 500% increase in customer repurchase rate
- Led ERP system modernization using Python and distributed warehouse management, enabling multi-state expansion and supporting 200% business growth
- Built Business Intelligence team of 6 engineers, implementing big data solutions with ETL pipelines that improved inventory forecasting accuracy by 85%
- Designed automated supply chain optimization using machine learning algorithms, reducing procurement costs by 15% and eliminating 90% of stockout incidents

Jüssi
Solutions Architect

São Paulo
November 2012 – May 2015

Architected digital marketing automation solutions as Facebook Marketing Partner and Google Partner for enterprise clients.

- Led Facebook Marketing Partner certification, collaborating with Facebook Engineering using GraphQL APIs to deliver marketing automation for biggest e-commerces in Brazil
- Architected Google Ads automation platforms using Java and RESTful APIs, managing \$5M+ monthly ad spend and achieving 40% improvement in ROAS
- Built real-time bidding optimization algorithms using statistical modeling, reducing customer acquisition costs by 25% across 5k+ enterprise campaigns
- Designed scalable attribution platform using message queues and distributed processing, processing 10M+ daily events with sub-second response times

Editora Abril
Senior Software Engineer

São Paulo
September 2008 – November 2012

Developed large-scale content management systems for Brazil’s leading media and entertainment publisher.

- Built Veja SP platform using HTML5, CSS3, JavaScript, and MySQL, serving 500K+ monthly users with 99.9% uptime
- Architected enterprise content management system using Java and SOA, supporting 20+ publications and 200+ concurrent editors
- Optimized database performance using advanced SQL queries and caching strategies, reducing page load times by 60% and supporting 10x traffic growth
- Led migration from legacy systems to modern web architecture using agile methodologies, improving development velocity by 75%

PAST EXPERIENCES

Serasa Experian (Java Enterprise Architect), Western Union (Senior Software Engineer), Toyota do Brasil (Software Engineer), Wdotnet Internet (Owner), Flyever Electronics (Software Engineer)

EDUCATION

Faculdade de Informática e Administração Paulista
Faculdade Impacta de Tecnologia
Universidade Federal de São Carlos

MBA Data Science (2015-2016)
BS Information Systems (2011-2015)
BS Materials Engineering (1996-2000)

CERTIFICATIONS

AWS: Solutions Architect Professional (2025), Data Engineer Associate (2025), ML Engineer (2024), ML Specialty (2024), AI Practitioner (2024), Solutions Architect Associate (2022), Developer (2022)
Specializations: Stanford Machine Learning and Deep Learning (2024), IBM AI Engineering Professional and AI Developer (2024), Data Science (2015), Recommender Systems (2014)

LANGUAGES

Portuguese: Native

English: Fluent

Spanish: Conversational