

NISHANT MODI

Principal Architect | Tech Lead • GenAI & Distributed Ecosystems

Pune, India
modinishant90@gmail.com
linkedin.com/in/nishant-modi
nishant-modi.vercel.app

A visionary orchestrator of digital complexity with over a decade of experience bridging the gap between bleeding-edge GenAI research and enterprise-grade reality. Specializing in autonomous agentic intelligence, high-concurrency Python architectures, and cloud-native ecosystems for Fortune 500 clients.

CORE EXPERTISE

GenAI & Agentic AI

LangChain RAG LLM Tool-Calling
Vertex AI Gemini OpenAI

Backend & Distributed Systems

Python FastAPI Django Flask
Concurrency Microservices

Infrastructure & Cloud

GCP AWS Kubernetes Docker
Terraform CI/CD

Architecture & Leadership

Technical Vision Team Mentorship
POC-to-Production Systems Thinking

PROFESSIONAL EXPERIENCE

Globant Solutions

2021 — Present

Technical Lead (GenAI Focus)

- Architected GenAI infrastructure powering Deloitte's global analytical intelligence.
- Managing an elite squad of engineers to harden POCs into production-grade assets for Fortune 500 clients.
- Integrated LLMs into core Python development cycles achieving a 30% increase in deployment velocity.

Xoriant Solutions

2014 — 2021

Senior Software Engineer

- Led automation initiatives for Hitachi Deployment Manager, deploying nodes globally using VMware APIs.
- Architected Python-Oracle middleware for Cisco Systems, optimizing complex version control workflows.
- Engineered investment automation for high-volume users at Lending Club to maximize portfolio yield.

KEY PROJECTS & TECHNICAL IMPACT

WINGMAN (NL2SQL Intelligence)

Vertex AI LangChain PgVector RAG

A revolutionary natural language interface granting analysts power to converse with massive datasets. Adopted as the primary intelligence layer at Deloitte.

- Implemented advanced RAG patterns for accurate schema mapping and query generation.
- Reduced report generation time from hours to seconds through automated SQL synthesis.

VOICE AI (Conversational Agent)

AWS Nova Sonic 2 OpenAI Twilio Redis

A conversational bot with emotional intelligence, replacing enterprise call-center logic with automated reasoning.

- Orchestrated low-latency audio streams for near-instant human-like response times.
- Enhanced customer engagement by bridging human empathy with automated decision-making.

VALUED (High-Speed Valuation Engine)

FastAPI Pandas Celery Concurrency

A valuation engine utilizing Pythonic concurrency to automate complex financial modeling globally.

- Engineered high-throughput processing pipelines for millisecond-precision financial calculations.
- Streamlined global valuation services, significantly reducing manual overhead.

HITACHI OPS (Infrastructure Automation)

Docker Python VMware API Bash

Automated configuration and deployment of mission-critical Datastore nodes globally.

- Achieved zero-touch deployment for global data centers via custom OVA scripts.
- Implemented rigorous security hardening for CentOS/Linux environments.

CERTIFICATIONS

Generative AI Leader Certification

Google Cloud • 2025

AI Agents Fundamentals

Hugging Face • 2025