

NIRAV GONDALIYA

Senior Full-Stack Engineer · AI Infrastructure

Rajkot, Gujarat, India (Remote) • nirav.gondaliya@icloud.com • +91 6354010189
linkedin.com/in/niravzxv • github.com/niravzxv

SUMMARY

Senior full-stack engineer with 7+ years shipping production systems across fintech, healthtech, and enterprise SaaS. I build the infrastructure that makes AI products actually work — MCP servers and clients, agentic pipelines, on-device LLMs, and resilient data systems — on a TypeScript / Node.js / React / Next.js / AWS stack. Author of two open-source MCP servers published to npm and a contributor to Google Lighthouse. I've delivered for clients across India, the US, and the UK, mentored 8 engineers, and acted as the technical bridge to enterprise C-suite stakeholders.

CORE SKILLS

Languages: TypeScript, JavaScript (ES6+), Python, SQL

Frontend: React, Next.js (App Router), Redux, Tailwind CSS, Angular, GraphQL

Backend: Node.js, Express, NestJS, REST APIs, WebRTC, Webhooks, OAuth 2.0, microservices, event-driven & serverless architecture

AI / LLM Infrastructure: MCP servers & clients, agent loops, multi-LLM routing, RAG, on-device / local LLMs, OpenAI & Anthropic APIs, LangChain, prompt engineering, vector databases

Data: PostgreSQL, MongoDB, MySQL, Redis, DynamoDB, Aurora, data warehousing, rate-limiting & backpressure

Cloud & DevOps: AWS (Lambda, API Gateway, CloudFormation, EC2, S3, SQS, Route 53, CloudWatch), Docker, CI/CD, GitHub Actions, infrastructure-as-code, multi-tenant provisioning

Practices: System design, performance optimization, TDD, code review, AI-assisted development (Claude Code, Cursor, Aider), Agile/Scrum

EXPERIENCE

Senior Software Engineer, Ratnam Solutions

Sep 2022 – Apr 2026

Remote, India

- Built a real-time B2B analytics platform for a multinational QSR chain operating across India and Indonesia, processing 5M+ data points/day into executive dashboards (SSSG, ADS, ADT) — replaced manual reporting and saved C-suite teams 5–10 hours per week.
- Architected a Node.js, MongoDB, and Redis data-warehousing pipeline, reducing data-transfer latency ~40% across distributed franchise sources.
- Served as primary technical liaison between client stakeholders and engineering; led requirement breakdown and task distribution across a 4-engineer team.
- Embedded AI-assisted development (Cursor, Claude Code) into the team's workflow, cutting feature delivery time on greenfield modules ~30%.
- Owned production deployment, monitoring, and incident response on AWS (EC2, S3, CloudWatch).

Independent Consultant, Shivay Infotech

2023 – Present

Remote · alongside full-time

- Built Consult Doctor — a telehealth platform with WebRTC video consultations, Razorpay payment integration, and digital prescription workflows.
- Built Elios GenAI Suite — an OpenAI-powered app spanning image generation, video processing, TTS, STT, résumé generation, and Google OAuth.
- Mentored 8 junior engineers in full-stack development, system design, and code-review practices.

Software Engineer, Techniverse Software Consulting

Jan 2022 – Aug 2022

Rajkot, Gujarat, India

- Designed and deployed serverless infrastructure (AWS Lambda, API Gateway, CloudFormation) for enterprise SaaS clients — cut infrastructure cost ~35% vs. equivalent EC2-based architecture.

- Built React frontends integrated with serverless backends, owning features end-to-end from API design to UI.
- Implemented infrastructure-as-code with CloudFormation, reducing new-environment provisioning from days to hours.

Software Engineer, Soft 'n' Web

Jun 2021 – Oct 2021

Rajkot, Gujarat, India

- Delivered conversational-AI features using Google Dialogflow and Firebase Functions, scaling to thousands of daily user interactions.

Software Engineer, WebCodeGenie Technology

Jan 2019 – May 2021

Ahmedabad, Gujarat, India

- Promoted from intern to developer within 6 months on shipping velocity and code quality.
- Built production CRM, ERP, Lead Management, and financial loan systems for mid-market clients across India (Node.js, Angular, MongoDB, MySQL).

SELECTED PROJECTS

PhonePe MCP Server *github.com/niravzxv/phonepe-mcp · TypeScript, MCP SDK, Node*

Unofficial MCP server for PhonePe's Payments API, exposing payment, status, refund, and webhook-verification tools to any MCP-capable AI client. Automatic OAuth token caching and refresh; sandbox and production environments; unit-tested against mocked requests.

UPI MCP Server (upi-mcp) *npm: upi-mcp · github.com/niravzxv/upi-mcp · TypeScript, MCP SDK, Node*

Open-source MCP server published to npm that lets any MCP-capable AI client generate and parse UPI payment links and QR codes — fully offline, with no API keys or accounts. Three tools (create link, create QR, parse link); strict TypeScript with discriminated-union result types; tested with Vitest and the MCP Inspector.

upi-agent (AI agent) *github.com/niravzxv/upi-agent · TypeScript, MCP client, Groq*

From-scratch AI agent (no framework) acting as an MCP client: connects to the upi-mcp server, exposes its tools to an LLM, and runs an agent loop where the model autonomously selects, chains, and executes tools. Handles real-world LLM tool-call failures with deterministic settings and retry.

Multi-Store Retail ERP / CRM / PM *Next.js, Node.js, PostgreSQL, RBAC*

Single platform unifying ERP, CRM, and project management for a multi-store retail chain, with role-based access control down to the field level across owners, managers, cashiers, and accountants.

Auto-Provisioning Lead Funnel *Node.js, AWS SDK, Route 53, IaC*

Lead funnel that programmatically spins up an isolated AWS instance of the product on its own subdomain per client — new tenant, DNS, and database provisioned automatically.

Resilient API → SQL Ingestion Pipeline *Node.js, SQL, queueing*

High-volume ingestion pipeline with adaptive rate-limiting, backoff, and queue-based backpressure to protect a fragile upstream API while guaranteeing reliable delivery into SQL.

OPEN SOURCE

- Author of two open-source MCP servers published to npm — phonepe-mcp and upi-mcp — usable by any MCP-capable AI client (Claude Desktop, Cursor, Cline).
- Google Lighthouse — authored a merged pull request to the official repository, the web-performance auditing tool used by millions of developers and built into Chrome DevTools.

EDUCATION

Bachelor of Computer Applications (BCA) — Marwadi University