

CHETAN VIRANI

AI / GenAI Engineer • LLM & Agentic Systems • Full-Stack & Data Engineering

Ahmedabad, Gujarat, India • +91 98255 31581 • viranichetan666@gmail.com •
<https://www.linkedin.com/in/chetan-virani-178a79114> • <https://github.com/viranichetan666>

PROFESSIONAL SUMMARY

AI / GenAI Engineer with 10+ years of software engineering experience, now focused on building production LLM and agentic AI systems on top of modern data platforms. Design and ship multi-agent workflows with LangChain and LangGraph, Retrieval-Augmented Generation (RAG) pipelines, and Model Context Protocol (MCP) servers that give domain-specific AI assistants tool access and persistent context across systems. Complementary strength in Python data engineering — dbt, Dagster, Trino/Starburst, Snowflake — so AI features stay grounded in reliable, well-modelled data. Also a proven full-stack engineer across the MERN and MEAN stacks (React, Next.js, Angular, Vue, Node.js, NestJS, GraphQL), with a record of architecting greenfield products and leading cross-functional teams.

TECHNICAL SKILLS

Generative AI & LLM Engineering: LangChain, LangGraph, Retrieval-Augmented Generation (RAG), embeddings & vector search, prompt engineering, structured output & function calling, LLM evaluation and observability, token/cost optimisation, streaming, guardrails; AWS Bedrock (Agent Core, Knowledge Bases), Anthropic Claude API, OpenAI API (GPT), Google Gemini

Agentic AI & Orchestration: Multi-agent systems, agentic workflows, tool use & function calling, Model Context Protocol (MCP) servers and clients, human-in-the-loop review, stateful graph-based agents, AI-assisted SDLC (GitHub Copilot, Claude Code, Cursor), n8n workflow automation

Data & Analytics Engineering: Python, dbt (data build tool), Dagster, Trino / Starburst, Snowflake, SQL, ELT / ETL pipelines, dimensional & semantic modelling, data quality testing, orchestration & lineage, analytics engineering

Backend Engineering: Node.js, Express, NestJS, Python (FastAPI), Go, REST APIs, GraphQL, Apollo, WebSockets, microservices, serverless, event-driven architecture

Frontend Engineering: React.js, Next.js, Angular, Vue.js, TypeScript, JavaScript (ES6+), HTML5, CSS3, Tailwind CSS, Bootstrap, responsive & accessible UI (WCAG)

Databases & Storage: PostgreSQL, MySQL, MongoDB, Redis, Snowflake, vector stores / pgvector

Cloud & DevOps: AWS (Bedrock, Lambda, S3, API Gateway), Docker, CI/CD, Git & GitHub, Linux, monitoring & observability (Datadog)

PROFESSIONAL EXPERIENCE

Lead Software Engineer — AI & Data Engineering | HRIMSYS

Sep 2020 – Present | Ahmedabad, India

- **Lead engineer for AI and data initiatives** — shipped production LLM and agentic systems (LangChain, LangGraph, RAG, AWS Bedrock, MCP servers) on governed data platforms built in dbt, Dagster, Trino/Starburst, and Snowflake.
- **Applied advanced prompt engineering and evaluation** — structured JSON outputs, few-shot and chain-of-thought patterns, and regression test suites over prompt changes — improving answer accuracy and reducing token spend.
- **Embedded AI into the software development lifecycle** with GitHub Copilot, Claude, and custom internal agents for code review, test generation, and documentation, measurably shortening delivery cycles for the team.
- **Architected and re-engineered full-stack products** end to end across React, Next.js, Angular, Node.js, NestJS, GraphQL, and PostgreSQL — **achieving a 75% improvement in application performance** through query optimisation, caching, and payload reduction.
- **Led cross-functional teams** of engineers, designers, and product managers; ran code reviews, mentored engineers, and set architecture and quality standards.

Full Stack Engineer | Volansys Technology

Sep 2019 – Sep 2020 | Ahmedabad, India

- **Designed and architected scalable software solutions** for enterprise clients using Angular, Node.js, React, and Go, applying microservices and API-first patterns.
- **Resolved complex performance bottlenecks** across backend services and databases, improving system responsiveness under production load.
- **Drove engineering quality** through systematic code review and proactive technical proposals that influenced architecture decisions, translating ambiguous business requirements into delivery-ready specifications and estimates.

Full Stack Engineer | OpenXcell Technolabs

Jan 2018 – Aug 2019 | Ahmedabad, India

- **Delivered production web applications** with React, Angular, Vue.js, and Node.js, translating UI/UX designs into performant, component-driven interfaces.
- **Optimised frontend performance** by reducing network requests, code-splitting bundles, and eliminating render bottlenecks; resolved cross-browser compatibility issues and implemented WCAG accessibility standards.

Software Engineer | Webclues Infotech

Jan 2016 – Dec 2017 | Ahmedabad, India

- **Developed and maintained backend systems** — coding, testing, debugging, and troubleshooting services — and integrated them with frontend applications and third-party APIs, following established protocols and security practices.

KEY PROJECTS

AI Engineer — Conversational AI Knowledge Assistant & On-Call Automation Agent

Stack: LangGraph, AWS Bedrock Agent Core, Bedrock Knowledge Bases (RAG), Claude / Bedrock-hosted foundation models, Python, AWS Lambda, API Gateway, S3

- **Designed and deployed a Retrieval-Augmented Generation (RAG) knowledge assistant** on AWS Bedrock Knowledge Bases, letting business users query domain-specific operational knowledge in natural language instead of navigating static documentation.
- **Built a multi-step on-call agent with LangGraph and AWS Bedrock Agent Core**, orchestrating tool-calling workflows to triage and route operational issues autonomously.
- **Implemented agent state management and tool-routing logic in LangGraph** supporting conditional multi-turn reasoning across knowledge retrieval, escalation, and notification workflows.
- **Integrated document ingestion pipelines** to keep the knowledge base synchronised with source-of-truth documentation, improving answer relevance and reducing stale-content errors.

Data Engineer — E-Commerce Advertising Analytics Platform

Stack: Snowflake, dbt, Dagster, Python, SQL, PostgreSQL, AWS S3 (S3 Tables / Apache Iceberg), Power BI, REST API integration

- **Architected and maintained ELT pipelines in dbt and Snowflake** that consolidated advertising performance data from multiple raw sources (EAV-structured event data) into governed dimensional data marts, replacing a legacy BI semantic layer of 500+ calculated measures.
- **Built and orchestrated automated decisioning pipelines in Dagster**, integrating dbt-computed business logic with third-party ad-platform APIs to execute bid and keyword-targeting changes at scale.
- **Designed a tiered data model** with target_id-grain fact tables supporting a multi-stage optimisation workflow, enabling incremental promotion of ad targets through automation tiers.
- **Diagnosed data-quality discrepancies across pipeline stages** — transformation logic versus live API state — by tracing lineage through multi-hop fact and dimension joins, reducing the mismatch rate in downstream reporting.
- **Modelled sales-attribution logic** joining budget and spend data to performance facts, defining grain mapping across previously siloed data domains.
- **Resolved transient cloud storage failures** in S3 Tables credential vending through retry-based fault handling, improving reliability of daily data builds.

Dripshop — Live Shopping & Collectibles Marketplace

Built real-time live-streaming commerce features — socket-driven auction and chat flows, checkout, and a GraphQL API layer serving concurrent viewers. (React.js, Next.js, Node.js, GraphQL, Apollo, TypeScript, PostgreSQL, WebSockets)

Jack Henry Financial Performance Suite (FPS)

Developed cloud-native budgeting, profitability, and board-reporting modules for banks and credit unions, including a daily metrics dashboard. (Go, Angular, Vue.js, GraphQL, microservices, TypeScript, PostgreSQL)

Additional Projects: **Dovetail** — paperless dental practice management: patients, appointments, insurance, treatment plans, role-based access control (React.js, Node.js, Python, GraphQL) · **HeyCollab** — team collaboration and kanban platform with Slack / Drive integrations · **Dealertouch** — dealership inventory, CRM, and service scheduling · **Sukshi Academy** — NEET / JEE exam-prep learning platform · **SNHU** — student-facing web experiences for online degree programmes

EDUCATION

B.E. — Information Technology, Vishwakarma Government Engineering College (GTU) | 2014 – 2018