

Summary

Senior **Agentic AI Engineer** focused on building **production-ready** AI systems for **Enterprise** customers. Experienced in taking ideas from customer discovery through **architecture, implementation, deployment**, and adoption. Strong engineering background across **LLMs, GraphRAG, Multi-agent systems, LLM observability**, and full-stack development, with a collaborative approach to working with **Customers, product**, and engineering teams

Experience

- Senior Engineer - GenAI & Agentic AI, Coforge

03/2024 – Present | Hyderabad

 - Architected and shipped an **AI-Infused Agentic Quality Engineering Platform** end to end; took it from feasibility framing to production with **17 enterprise QA teams** at peak adoption, and deployed it into **customer-managed AWS, Azure, and GCP tenants** with architecture, compute sizing, and identity (**Azure AD, IAM, SSO**) adapted per cloud, all on one codebase with no forks.
 - Slashed production hallucination from **~15% to under 5%** by engineering complex production RAG pipelines from scratch: **GraphRAG on Neo4j** with entity normalization, query transformation, re-ranking, context compression, metadata filtering, and guardrails. Layered **Agentic RAG over Jira and Azure DevOps** so test generation is scoped against linked stories, epics, and acceptance criteria.
 - Cut per-tenant LLM spend through **token and cost engineering** (context compression, prompt-payload reduction, rerouting low-complexity queries to lighter models) and built the **LLM observability and FinOps layer** beneath it: per-tenant cost telemetry plus quality evaluation (**Acceptance Criteria Coverage, Test Design Coverage, RAGAS**) that customer stakeholders use to track spend and output quality live.
 - Eliminated **85-90% of manual QA effort** with agentic test generation, execution, and **CI-integrated automation**; built a custom **Playwright MCP integration** that turns natural language into production-grade Playwright and Cypress scripts and cut test authoring time **~75%**.
 - Secured **multi-tenant isolation** across clouds with **tenant-aware RBAC**: module-level CRUD-X permissions and role templates that let new enterprise customers onboard without a rebuilt permission model.
- Data Scientist - GenAI, ML & Conversational AI, Hewlett Packard Enterprise

11/2022 – 12/2023 | Bangalore

 - Architected the migration of an enterprise conversational system from **rule-based FAQ to LLM-assisted RAG** over ~1,500-1,700 internal documents, cutting repeat user escalations **~40%**.
 - Designed the retrieval pipeline end to end: chunking with overlap calibration, **hybrid search combining semantic and keyword signals, BM25** and metadata filtering tuned for enterprise-jargon queries. Added evaluation loops that measured retrieval precision against a held-out set and improved response relevance **30-40%** in pilot.
 - Built the case for the redesign from **customer usage analysis**, which showed the rule-based system failing on anything outside its predefined set, then owned the rollout.
- Associate Software Engineer - Machine Learning & Data Science, Mphasis

07/2019 – 10/2022 | Bangalore

 - Built **fraud-detection and risk-scoring models** for financial-services clients using tree-based ensembles (**XGBoost, LightGBM, Random Forest**) with engineered feature pipelines, improving precision **~15%** over legacy rule-based systems and cutting manual review volume.

Education

- Product Management in Generative & Agentic AI, BITS School of Management (BITSoM)

07/2025 – 03/2026 | India
- PG Diploma in Machine Learning & Artificial Intelligence, IIIT Bangalore

06/2020 – 03/2022 | Bangalore
- Bachelor of Technology in Computer Science, Bharat Institute of Engineering and Technology

08/2015 – 06/2019 | Hyderabad

Skills

Generative & Agentic AI: LLMs | Prompt Engineering | Agentic Workflows | Multi-Agent Orchestration | LangGraph | LangChain | GraphRAG | Agentic RAG Pipelines | Neo4j | Hybrid Search | Reranking | MCP Servers | LLM Observability | Context Engineering | RAGAS | Guardrails | Responsible AI

Enterprise & Customer-Facing: Forward Deployed Engineering | Solutions Architecture | Customer Discovery | Stakeholder Alignment | Production Deployment | Multi-Tenant Architecture | RBAC Design | Enterprise CRM Integration | Identity & SSO Architecture | OAuth Protocols | Customer-Facing Technical Communication | Product-Informed Engineering Judgment

ML & Data Science: Tree-Based Ensembles (XGBoost, LightGBM, Random Forest) | Feature Engineering | Predictive Analytics | Risk Scoring | Statistical Modeling

Engineering & Delivery: Python | FastAPI | React 19 | MongoDB | PostgreSQL | Docker | CI/CD (Harness, GitHub Actions, Railway) | Cloud Platforms (Azure, GCP, AWS) | System Design | Data Pipelines | REST APIs | Webhook Architecture | Async Architecture | Microservices | Playwright | Cypress

Projects

- Agentic Codebase Intelligence System, Multi-Agent Platform for Enterprise and Architectural Level Codebase Auditing

06/2026 – Present

 - Architected and shipped a **multi-agent codebase-intelligence platform end to end**, cutting manual code-review and audit effort **~50%** by orchestrating specialized LLM agents over a hybrid retrieval layer (vector + knowledge graph) rather than flat file search.
 - Engineered a **codebase knowledge-graph engine** mapping entities, call paths, and cross-file dependencies across polyglot repositories, cutting **false-positive findings ~40%** by grounding every agent decision in structural context instead of raw text matches.
 - Built an **audit orchestrator coordinating 6 specialized agents** (retrieval, reasoning, verification) with MCP tool integration, compressing time-to-findings from **~2 days of manual review to under 20 minutes** and auto-filing prioritized issues into Jira.
 - Designed the **retrieval and guardrail pipeline** (hybrid search, context compression, verification gates), holding **hallucinated or low-confidence findings under 5%** so agent output was trustworthy enough to act on.

Loop Copilot : AI CRM Copilot for Enterprise Sales, Architect · Engineer · Sole Owner | Production, Fortune 500 pilot

- Built and shipped a **production AI CRM copilot** for enterprise sales reps as sole architect and engineer; took it from customer discovery through V2 with an active beta user inside a **Fortune 500 sales org**, cutting CRM logging time from **4-6 min to ~45 sec (~85%)**.
- Shipped a **voice agent pipeline** that listens to sales-client calls (**STT**), turns them into structured **PRD** and **onboarding** documents, and **auto-detects** any newly mentioned accounts. It resolves them against the **account repository** using **unique ID**, then logs them straight into **D365** or the connected **CRM**, so reps never have to manually key in account details after a call.
- Designed the **D365 integration inside the customer's tenant trust model: Power Automate flows** authenticated by internal tenant identity, with direct **Dataverse REST via MSAL** held in reserve for environments that grant broader permissions.
- Architected the full stack (**React 19**, async **FastAPI**, **MongoDB Atlas**, **Groq/Llama 3.1**, **Entra ID** via **MSAL**, **Microsoft Graph**, **Telegram Bot API**, **Railway CI/CD**) and chose **structured event sourcing** over vector RAG for memory, removing hallucinated recall while keeping cross-session context; shipped **10+ feature areas** including calendar integration, bulk upload with **AI summarization**, and admin analytics, with a **multi-CRM expansion shell (Salesforce, HubSpot, Zoho)** architected and ready.

Voice-First Financial AI Copilot, Capstone, BITSoM : Architected, Built & Piloted

- Owned architecture and delivery of a **voice-first financial AI copilot** for India's gig workforce, from customer research through pilot: a **~30-screen MVP with 10-language architecture**, shipped on both web and mobile.
- Designed a **multi-layer compliance architecture** with **SEBI advisory principles and DPDP requirements** as first-class design constraints: RAG over a regulatory-safe knowledge base, system-prompt guardrails, server-side governance, and consent-first data handling.
- Engineered a **central memory layer with RAG** so user context, goals, and prior conversations persist across sessions and across web and mobile from a single source of truth.

Autonomous Wealth Operating System, Multi-Asset Agentic Application | Architect. Engineer. Owner | Ongoing

- Architected and built an **autonomous multi-asset wealth platform** (stocks, mutual funds, ETFs, bonds, NPS) around a **21-agent analyst council**: isolated agent opinions, a **devil's advocate stage**, aggregation weighted by each agent's **historical accuracy**, and a **code-enforced veto gate** so no money moves on LLM output alone.
- Engineered a **five-ring risk system** (position, portfolio, drawdown, budget, and connection-level **broker locks**) plus a **frozen-threshold evolution gauntlet**: any strategy change must beat the incumbent across **fixed statistical gates** before promotion, which rules out **silent strategy drift**.
- Delivered it as a **multi-surface platform**: a shared **framework-free core** behind web and mobile clients, a **Telegram copilot** for conversational instructions and live trade, investment, and report updates, and **user-selectable LLM models per agent role**.

Certifications

Claude Certified Architect
Anthropic : In-Progress

Agile Practices Certificate
SAfe