

TIRTH GAJJAR | 9+ YOY

contact@tirth.dev • (+91) 85 114 55 160 • linkedin.com/in/tirth-gajjar • tirth.dev

SKILLS

AI / AGENT SYSTEMS

- Agentic pipelines, ReAct & CodeAct loops, planner + critic engines, multi-agent orchestration, task-graph execution, reflection loops, async long-running tasks, memory-driven architectures, safety & fallback heuristics
- MCP: custom typed MCP servers & tool meshes, OpenMCP-style tool discovery, vault-backed browser-use MCP, AgentKit-driven runtimes, custom Python tool servers
- LangChain, LangGraph, LlamalIndex, OpenAI Agents SDK, Anthropic SDK, Google Vertex AI, HuggingFace, Docling, LitServe, Ilmlite
- RAG: pgvector, Pinecone, Weaviate, Qdrant, hybrid semantic + BM25 retrieval, embeddings, rerankers, context compression, Agentic Search, PageIndex
- Fine-tuning: LoRA, instruction tuning, embedding refinement, preference optimization (Unsloth); constrained schema-first decoding & output validators
- Evals & observability: Langfuse, eval harnesses, experiment tracking, prompt & run analytics, confidence scoring, hallucination control

BACKEND

- Python (Advanced); FastAPI, Django, Flask, Celery, asyncio), Node.js, Express.js
- REST API Design, GraphQL, Redis Queue
- AWS (Lambda, ECS, S3, SQS, SNS, DynamoDB, RDS, CloudFront, CloudWatch, EC2, IAM, VPC, KMS), Terraform
- Serverless Framework, Docker, Kubernetes, AKS

FRONTEND

- React.js, React Native, Next.js, Remix.js, Electron.js, TypeScript, React Router, Redux, Redux-Sagas, React Hook Form, Jest, Cypress, Radix UI, ESNEXT, ES6
- TailwindCSS, Styled Components, SCSS

DATABASES

- PostgreSQL, MySQL, MongoDB, DynamoDB, Redis

TOOLING

- Docker, Docker Compose, Webpack, Babel, Turbopack, Lerna, Yarn Workspaces, GitHub Actions, Husky, Prettier, ESLint

PROFESSIONAL EXPERIENCE

[Indexa Exchange Group](#)

Fractional CTO

Remote

March 2026 - Present

- Define and execute the technical strategy for Indexa, an institutional-grade system for transparent benchmark construction and index methodology
- Architect multi-tenant infrastructure for capital markets participants, with full traceability, permissioned access, and audit logging at every layer
- Design deterministic scoring and calculation frameworks producing reproducible, auditable results across every pillar and constituent
- Lead the design of eligibility engines, index composition systems, and calculation pipelines with the precision and rigor expected of benchmark infrastructure

- Establish data governance and methodological consistency standards required for regulatory credibility and institutional adoption

Nurix.ai (Anyteam) (Backed by Accel)

Remote

AI Engineering Lead (Contract)

March 2025 - Feb 2026

- Designed AI call companion using LangChain + LlamaIndex with reflection loops, context compression, and knowledge-tool execution
- Implemented pgvector-based RAG for product knowledge, competitor intel, and customized sales playbooks with dynamic reranking
- Built evaluation & telemetry layer using Langfuse to score accuracy, objection-coverage quality, and sales-context alignment
- Developed dual-pipeline speech system: local STT + cloud transcription fallback for accuracy and resilience in live sales environments
- Built cross-platform desktop client using Electron.js running local Whisper inference for low-latency meeting transcription
- Streamed and stored transcripts to Firebase, enabling real-time guidance, objection handling prompts, meeting notes, and follow-ups
- Integrated CRM, email, and calendar workflows to automate post-call summaries, next-steps drafting, and account research tasks
- Orchestrated multi-agent tool execution for live objection handling, competitor lookups, and dynamic playbook retrieval during calls

UBIK (BigCircle) (Backed by Antler Malaysia)

Remote

AI Systems Architect

Jan 2025 - March 2025

- Built a planner + ReAct / CodeAct + critic engine that turns high-level goals into executable task graphs across email, calendar, docs, sheets, CRM, and ERP tools
- Designed a unified MCP mesh with OpenMCP-style tool discovery, including a vault-backed browser-use MCP for authenticated (SSO/MFA) web apps with no public APIs
- Added local RAG + planner memory so agents ground in the user's own data and self-improve planning strategies, lifting no-intervention plan success from ~55% to ~80%
- Implemented a CodeAct IPython MCP for analytics (Thought to Code to Run to Observation to Refine), cutting multi-hour spreadsheet work to sub-15-minute reproducible reports
- Integrated eval + telemetry layer using Langfuse to track response quality, hallucination control, and task reliability
- Deployed (PoC) inference layer with LitServe + lllite for low-latency streaming responses and scalable agent execution
- Built AgentKit-driven voice runtimes (VAD to streaming STT to LLM + typed MCP tools to TTS) for multi-tenant, low-latency conversational automation with strict guardrails

Brandelligence (BigCircle)

Remote

AI + Tech Lead

Jan 2022 - July 2024

- Built a legal intelligence engine for trademark similarity and confusion analysis, combining instruction-tuned LLMs, RAG, and rule-based heuristics into attorney-style opinions
- Used constrained, schema-first decoding plus validators to enforce a fixed 5-section opinion structure, jurisdiction tags, and citation formats, lifting template adherence from ~40% to ~98%
- Added a RAG layer over whitelisted case law, statutes, and registry data, cutting fabricated/invalid citations by ~90% in a precision-first regime
- Leveraged Unsloth-optimized LoRA fine-tuning on attorney-written opinions to improve semantic + phonetic similarity accuracy and formal legal tone

- Built a 0-10 legal-integrity eval harness with production logging to compare base vs fine-tuned models and detect regressions across structure, reasoning, and citation accuracy
- Delivered multi-agent pipelines for NICE class mapping, case law lookup, and conflict scoring

Gartner Inc. (Contract)

Remote

Engineering Contractor

Oct 2021 - Oct 2024

- Led the technical transition of two main applications to GraphQL, planned and executed over nine months for a scalable, future-proofed ecosystem, with backend components powered by both Node.js and Python
- Implemented server actions to resolve cross-departmental client-side challenges during Next.js migration, enhancing performance and security across SEO, Security, DevOps, and Engineering departments
- Initiated a security overhaul by decommissioning a vulnerable REST endpoint (re-implemented in Python Flask services), thwarting bot attacks and eliminating the need for costly anti-bot solutions
- Designed and deployed a unified data fetching strategy within a monorepo, integrating React.js apps with Python backend services across 13 BX Engineering Apps
- Built major features around category (product listing) pages; developed a framework for running A/B tests on incoming traffic
- Created migration strategies for moving to Next.js from the in-house SSR framework, adopted by several teams; researched Contentful CMS integration and caching strategies
- Migrated to latest Serverless framework after breaking changes; contributed to the companywide Design System component library; worked with DevOps on a DR environment for all apps

Blueland (Backed by Shark Tank USA)

Remote

Tech Lead

Jan 2022 - July 2024

- Built custom experiences for e-commerce platform using Next.js, deployed on Vercel for seamless performance and improved Core Web Vitals
- Managed and optimized the Shopify backend, ensuring robust, scalable operations; migrated checkout to Shopify's new extensibility framework
- Developed custom discount extensions on Shopify, enabling personalized promotions and offers to enhance engagement and boost sales

Mission Sustainability

Remote

Fractional CTO

Jan 2022 - July 2024

- Engineered an advanced sustainability platform using AWS Lambda (Python) for serverless computing and S3 for scalable storage, creating a centralized hub for eco-friendly practices and companies
- Designed Python-based data processing pipelines for real-time environmental metrics visualization
- Architected an Electron-based dashboard aggregating environmental data, streamlining reporting and enabling rapid decision-making
- Implemented robust, scalable infrastructure across AWS (CloudFront, Lambda, S3) ensuring agility for sustainability-focused projects
- Cultivated and guided a team of developers in Python and AWS best practices to deliver impactful environmental solutions

Rax Labs Inc.

Remote

Product Architect / Tech Lead (Core team)

Jan 2020 - Oct 2021

- Created, led, and mentored a team of 8 engineers from 0, filtering 500+ applications across 100+ interview rounds
- Redesigned the architecture to achieve 4 nines of scalability and cost reductions up to 34%; removed 8 single points of failure

- Migrated 25+ services to event-driven, serverless architecture (AWS Lambda with Python, DynamoDB, S3, SQS, SNS, ECS, RDS)
- Oversaw gradual deprecation and migration of legacy services into Python-based and serverless architectures
- Introduced Terraform and Ansible, migrating 13 services to IaC; built CI/CD pipelines for 50+ backend services with GitHub Actions
- Leveraged advanced Python concurrency (asyncio, multiprocessing) to optimize backend throughput and reduce latency

SalesHandy

Ahmedabad, India

Senior Software Engineer cum Team-Lead

Oct 2018 - Aug 2019

- Spearheaded a team of 4 front-end developers; trained the team in React and Express, test-driven development, and introduced Python-based services
- Coded major features like read-receipt for emails in Gmail in the product's browser extension, improving product adoption by 30%
- Revamped 5 core features including time-zone-aware email campaign scheduling (Node.js backend + Python background processing), cutting support queries 27%
- Introduced and moderated 10+ engineering processes; ran weekly cross-team code reviews and bi-weekly UI audits; stabilized legacy Node.js APIs

MAQ Software

Mumbai, India

Technology Research Engineer

Jan 2017 - Oct 2018

- Automated the data validation process, validating 5000+ data points daily generated by 20+ teams across the organization
- Managed a technical research team of 3 exploring Microsoft Azure and Power BI; coached 10+ service teams on technology adoption
- Built a custom Power BI reporting tool for internal use, now a product offering by MAQ Software
- Maintained and extended Python-based reporting/data-processing tools with MySQL; onboarded 28+ interns via boot-camp sessions

EDUCATION

DA-IICT (Dhirubhai Ambani Institute of Information & Communication Technology)

Gandhinagar

Bachelor of Technology, Information and Communication Technology (ICT)

Jun 2013 - Apr 2017

- Computer-science focus; favourite courses included Software Engineering, Computer Networks, DBMS, and Human-Computer Interaction