

Abdus Samad

Software Engineer · Gen AI & Agentic Systems · Full-Stack Node.js / NestJS · React / Next.js · AWS & DevOps
Delhi, India (IST, UTC+5:30) · open to remote and relocation | samad.abdus3535@gmail.com · +91 96544 05340 · [linkedin.com/in/asamad35](https://www.linkedin.com/in/asamad35) · github.com/asamad35
PORTFOLIO asamad.vercel.app · case studies with architecture, metrics and screenshots, plus a live Boondi chat

SUMMARY

5+ years shipping production LLM agents and large backends: led **Boondi**, a WhatsApp AI concierge on the Claude API and MCP (**250+ customer chats a day at \$0.02 each**), and built hoichoi's NestJS/GraphQL services (**10M+ requests an hour, 40M records migrated**, zero critical downtime).

SKILLS

Languages & UI	TypeScript, JavaScript, Python, SQL, React, Next.js, Redux, React Query, Tailwind CSS, Playwright
Backend & data	Node.js, NestJS, GraphQL, REST APIs, PostgreSQL, pgvector, MongoDB, TypeORM, Redis, Kafka, Socket.IO
Applied AI	Anthropic Claude API, MCP (Model Context Protocol) servers, agent runtimes (Gantry, LangGraph), tool calling, prompt caching, RAG, evals, guardrails
Cloud & DevOps	AWS (ECS Fargate, RDS, S3, CloudWatch, X-Ray), Terraform, Docker, Kubernetes, GitHub Actions CI/CD

EXPERIENCE

KnackLabs (formerly CAW Studios) · SDE 2, promoted from SDE in Oct 2024 Remote · Hyderabad, India · Oct 2023 – Present

Boondi · WhatsApp AI concierge for Bombay Sweet Shop Lead engineer · May 2026 – Present · [chat live on WhatsApp](#)

- PROBLEM Festival chat volume outran the support team and buried corporate gifting leads inside ordinary threads.
- Architected four services around one Postgres (Gantry agent host, read-only Shopify and CRM MCP servers, Next.js console) serving two personas from one webhook; the model only calls typed, least-privilege tools and never sees a key.
 - Made identity unforgeable: HMAC-signed caller identity on every MCP call, tools with no phone or email arguments, a boot check refusing write-capable tools, and a cached catalogue snapshot so every price and stock claim is grounded.
 - Built the eval and safety loop: 43 live scenarios gate every prompt change, simulation numbers can never reach a customer, DB-lease ordering keeps parallel workers in sequence, and a 35-message cap forces handoff in code.
 - Load-tested on real ECS and RDS with faked Anthropic and Shopify upstreams, finding a global advisory lock and then the pool ceiling (40 to 100); shipped via Terraform and one ordered GitHub Actions pipeline with OpenTelemetry tracing.

IMPACT Built solo, first commit to production in **12 weeks** · 250+ customer chats a day · **1,200+ sales leads** in two months · **\$0.02** per conversation · **94.6%** prompt-cache hit rate, spend –83.8% · median turn **3.5 s**, p95 ~6 s · **285** concurrent turns

Gantry · open-source, self-hosted agent runtime (KnackLabs, MIT) github.com/knacklabs/gantry · Apr 2026 – Present

- First production adopter; maintain the production fork (360+ commits: per-channel personas, memory digests, message caps, handoff hooks, tool ledger, X-Ray tracing) vendored as a package, and fed the load-test findings (global advisory lock, pool sizing) and settings-contract limits upstream to the core team.

hoichoi · Bengali OTT platform (SVF), 100+ countries Full-stack engineer · Oct 2023 – 2025 · hoichoi.tv

- PROBLEM A licensed SaaS OTT stack had become the ceiling; rebuild it in-house without the paying audience noticing.
- Built NestJS services behind GraphQL and REST on MongoDB for catalogue, subscribers and entitlements, schema designed from the read paths back (dataloader batching, Redis), serving millions of daily users at **10M+ requests an hour**.
 - Owned the **40M-user, 10K-title migration**: dual-write, checkpointed backfill, a verifier run to zero diff, per-domain cut-over behind flags with tested rollback; **99.9% uptime**, zero critical downtime.
 - Led the content CMS frontend (publishing time –50%), moved SMS and email into queue-fed microservices (core load –30%), and cut CloudWatch spend **70%** (\$10K to \$3K a month) with alerting and X-Ray tracing intact.

Xicom Technologies · React Developer On-site · Delhi, India · Jan 2022 – Oct 2023 · gogmgo.com

- Built a QR food-ordering platform and ERP (100+ restaurants, 50K+ daily users) in React/Redux; migrated the PHP backend to Node.js REST APIs (30% faster), added Redis caching (200K+/day) and Socket.IO chat (+20% retention).

Career break · medical leave and recovery after surgery Jan 2021 – Jan 2022

Gurucool · Founding Front-End Engineer Delhi, India · Mar 2020 – Jan 2021 · gurucool.xyz

- First front-end hire at an edtech network: built the Next.js app and component library, wired React Query with mock adapters against a moving backend, and shipped the investor site ahead of deadline; the pre-seed round followed.

PROJECTS

Portfolio (asamad.vercel.app): Next.js 16, React 19, React Three Fiber, a 22,000-point custom-shader hero, Lighthouse 100/100/100/100 on desktop · **Doodle-It** ([live](#) · [GitHub](#)): canvas drawing engine from scratch in TypeScript with undo/redo

EDUCATION

Bachelor of Computer Applications · Ambedkar Institute of Technology (now DSEU), Delhi 2020 – Jul 2023
CGPA 8.7/10 · All-India Rank 24, GGSIPU CET (BCA) 2020 · completed while working full-time from the second semester