

Ganesh Gunasegaran

AI Systems Architect | Staff Software Engineer

Pondicherry, India | India/Global Remote

me@itsgg.com | www.linkedin.com/in/itsgg | github.com/itsgg

Full career record and project links: itsgg.com/resume

PROFESSIONAL SUMMARY

I build AI and distributed systems, and the controls that keep them reliable. At Fluxon, I lead engineering for Project Alexandria and contribute to architecture and delivery for the Google Sustainability Performance Hub. My work combines hands-on implementation with system design, delivery and operations across AI, fintech, payments and real-time platforms.

CORE EXPERTISE

AI systems: AI system reliability, LLM systems, AI evaluation, Provider routing and fallback, Prompt-injection defenses, Cost controls, Observability

Architecture: System design, Distributed systems, API architecture, Event-driven systems, Platform modernization, Reliability engineering

Implementation: Python, TypeScript, Kotlin, Java/Spring Boot, Ruby on Rails, PostgreSQL, Docker, Kubernetes

PROFESSIONAL EXPERIENCE

Fluxon, Remote

Staff Software Engineer | Jun 2025 to Present

Project Alexandria

Lead engineering for Project Alexandria, a privately backed initiative aiming to translate 1,000 public-domain classics into major languages and publish them free as PDFs, EPUBs and audiobooks.

- Designed and implemented LLM cost controls with per-user attribution by provider and model, configurable daily caps, drill-validated kill switches, and provider fallback that completed requests during real provider failures without user action.
- Built prompt-injection defenses validated through automated and manual adversarial testing, and hardened the translation pipeline to reject incomplete responses, recover unreadable source PDFs, route books by detected language, and process books that exceeded prior limits.
- Added proactive alerts that detected failures before user reports, shared rate limiting across application instances, signed webhooks and access controls, and load tests that validated expected production traffic.

Google Sustainability Performance Hub

- Contribute to architecture and delivery with engineering, product and design while working directly on Angular/TypeScript interfaces, Kotlin services, API integrations, automated tests and production incident resolution.
- Contributed to service-boundary design and shared integration patterns that reduced coupling, enabled faster independent delivery, and improved consistency, reliability and maintainability; helped establish architecture reviews, coding standards, quality gates and testing practices.

Turing, Remote

Full Stack Architect | Feb 2022 to May 2025

Q.ai, AI investing

- Led backend architecture and engineering with ML teams, productionizing model outputs from evaluation and data contracts through serving infrastructure and real-time metrics; cut analytics processing time by about 40% through processing, caching, parallel execution and service-design changes.
- Designed and drove adoption of a federated API architecture across Rails, Kotlin, Python and Spring Boot, giving clients unified access while allowing services to evolve independently and improving delivery speed and integration reliability.

SELECTED PUBLIC WORK

- **vazhi** (github.com/itsgg/vazhi): Built and published a TypeScript LLM router for the pi coding agent with live provider probes, automatic fallback, shared cooldowns, constrained endpoints, nightly pool updates, CI and tests.
- **Dream-RSI Local** (github.com/itsgg/dream-rsi-local): Reimplemented an unreleased agent-search method from its paper and ran a controlled three-seed test on a local model. It found no score improvement, traced 11 crashes across 27 generated policies and repaired 7 from tracebacks alone.
- **Nodi** (github.com/itsgg/omarchy-nodi): Built and shipped an Omarchy command bar for application, window, file and clipboard workflows, with questions routed to Claude, Codex, Gemini or user-owned agents over the Agent Client Protocol.

SELECTED ARCHITECTURE EXPERIENCE

BukuKas (Lummo), Remote

Chief Architect | Dec 2020 to Feb 2022

- Architected campaigns, payments and notifications across product, growth and operations; designed gamification flows that more than doubled participation and led P2P and merchant transaction architecture.
- Built a unified SMS, WhatsApp, push and WebSocket notification service with provider failover, and implemented fraud controls, OTP verification and transaction observability.

BigBinary | Senior Architect | Aug 2019 to Nov 2020

MoneySmart | Senior Software Architect | Jan 2018 to Jul 2019

Crossover (Engine Yard) | Chief Software Architect | Jan 2017 to Dec 2017

EARLIER SYSTEMS HIGHLIGHTS

- Founded and architected Akshi, a real-time learning platform with adaptive video, chat and whiteboarding that supported more than 4,000 concurrent students.
- Led the IBM/HCL team that ported WebSphere MQ to Tandem NSK and designed a core SSL module in C/POSIX for enterprise messaging systems.

EDUCATION

Bachelor of Technology (Distinction with Honours)

Pondicherry University, Pondicherry