

NIKHIL SAXENA

+91-8920212782 | myemailnikhilsaxena@gmail.com | in/nikhil-saxena-codes | github/nikhilsaxena04 | Portfolio

Education

Dr. A.P.J. Abdul Kalam Technical University

Sep. 2023 – Expected May 2027

Bachelor of Technology, Computer Engineering

Open Source Contributions

Graphify (120K+ Stars, 7M+ PyPI downloads, YC S26) | Python, C++, Kotlin, Tree-sitter, ASTs 🔗 Sep. 2026

- **PR #3843** – Fixed **Python monorepo import resolution for nested scan roots** by inferring namespaces from `__init__.py` ancestors and projecting absolute imports onto the scan root; **shipped in v0.9.70** with 6 tests.
- **PR #3848** – Fixed **Kotlin AST extraction for annotations and primary-constructor properties**, restoring field/type/attribute relationships; **shipped in v0.9.69** with 9 regression tests.
- **PR #3461** – Added **first-class PHP closure extraction** with stable per-scope IDs, nested-closure traversal, and route-aware naming for HTTP handlers; **shipped in v0.9.67**.

Checkmate (11K+ Stars) | React, gRPC, REST, MongoDB, Zod

🔗 Sep. 2026

- **PR #3975** – Built click-through **90-day incident history** for public status pages: **lazy REST API** with **Zod validation** and public-data sanitization, **React/MUI** modal with **i18n/timezone** support, **MongoDB compound indexing**, and 403/404 regression tests.
- **PR #3974** – Fixed **gRPC monitors** stuck in an initializing state due to **health.proto** resolution failures; corrected runtime path handling and updated the CI/CD build pipeline to package the protobuf into **dist/**.

Projects

OmniRoute – Multi-Provider LLM Gateway | Go, Python, pgvector, Prometheus, Langfuse, Docker

🔗 Sep. 2026

- Architected a multi-provider LLM gateway in Go, exposing a **unified OpenAI-compatible API** for drop-in integration, dynamically routing requests using **YAML priority tiers** and real-time token cost/rate tracking.
- Implemented a **Semantic Cache** using **OpenAI Embeddings** and **pgvector**, executing **cosine similarity lookups** to serve responses for semantically identical prompts, slashing **P95 latency ~80%** and API spend.
- Engineered an automated **LLM evaluation and benchmarking engine** in Python/Go, executing concurrent fan-out runs to measure token latency (TTFT), cost per 1K tokens, and response quality across upstream providers.
- Built zero-dependency resilience primitives from scratch: a **sliding-window circuit breaker** for instantaneous model fallback during upstream rate-limits/outages, and an atomic **token-bucket rate limiter**.

High-Throughput Notification Broker | Go, gRPC, Redis, AWS, Prometheus, Grafana, Jaeger

🔗 May 2026

- Architected a decoupled **3-tier microservice** notification engine in Go, sustaining **9,200+ req/s** with **13.8ms average latency (<30ms P95)** under intensive local stress testing.
- Engineered a zero-loss worker pool with **200 concurrent goroutines** utilizing Redis **BLMOVE** for atomic queue draining, integrating **exponential backoff, jitter, and a Dead Letter Queue (DLQ)**.
- Serialized **OpenTelemetry traceparent headers** into Redis message payloads, preserving distributed tracing context across async boundaries in **Jaeger**.
- Deployed to **AWS (EC2, ECR, S3)** using IAM instance profiles; engineered a background DLQ archiver draining **100 jobs/cycle** into S3 JSON archives via **multi-stage Docker builds (<20MB)**.

Technical Skills

Languages: Go, Python, C++, SQL (PostgreSQL), TypeScript, Bash

AI & LLM Engineering: LLM Gateways, LLM Evals, pgvector, RAG, Semantic Caching, MCP, Multi-Agent Workflows

Backend & Cloud: AWS (EC2, S3, ECR, IAM), Linux, Docker, GitHub Actions, Redis, gRPC, Protobuf, REST APIs, WebSockets

LLMOps & Observability: OpenTelemetry, Prometheus, Grafana, Jaeger, Langfuse, Distributed Tracing

Core Fundamentals: Data Structures & Algorithms, Distributed Systems, Concurrency, Microservices, DBMS, OS

Experience

Creative and Tourism Club

Aug. 2024 – May 2026

General Secretary

ABES Engineering College

- Led a **45+ member organizing committee** to execute college-wide events (“ABES Got Talent”, “SAARANG’26”) for **800+ attendees**, successfully pitching and securing corporate sponsorship from **Royal Enfield**.