

Nikhil Shankar

nikhilshankar221@gmail.com | (510) 591-9587 | github.com/heyits-nick | linkedin.com/in/nikhilshankar08

EDUCATION

University of Michigan-Dearborn

Master of Science, *Computer and Information Science*

Aug 2024 - Apr 2026

GPA: 3.8/4.0

Visvesvaraya Technological University

Bachelor of Engineering, *Computer Science Engineering*

Aug 2019 - Aug 2023

WORK EXPERIENCE

Healmenow Inc. | AI & Infrastructure Engineer | San Francisco, CA

Sep 2026 - Present

- Built a digital-health platform spanning 19 web pages, 23 API endpoints, and a native iOS app using React/TypeScript, Python/FastAPI, SwiftUI, and a self-hosted FHIR repository backed by PostgreSQL and Redis.
- Enabled record-grounded healthcare AI across 4 LLM providers with cited answers, clinician-ready summaries, and configurable privacy controls; required human approval of all AI/OCR clinical proposals before updating patient records.
- Deployed the platform on AWS with Docker Compose and automated TLS across 4 subdomains; isolated databases on an internal network and scoped service credentials to least-privilege FHIR access policies to protect clinical data.
- Prototyped a voice-guided medication dispenser using Raspberry Pi, ESP32, and speech AI, combining scheduled dose delivery, patient check-ins, adherence tracking, and caregiver visibility for users with memory or dexterity challenges.

Complama Technologies, Inc. | DevOps Engineer | New York, NY

Sep 2025 - May 2026

- Reduced p95 latency 30%, maintained 99.9% uptime through peak traffic, and saved ~\$2,000/month using AWS Infrastructure as Code (IaC) with Terraform and CloudFormation and tuning EKS workloads with Helm, autoscaling, and resource limits.
- Cut manual deployment effort ~40% by building test/dev/prod CI/CD pipelines with GitHub Actions, Jenkins, Ansible Tower, and Docker; gated every release with image scanning, integration tests, and approval workflows.
- Cut mean time to recovery (MTTR) 30% through SQL-backed health probes, Prometheus/CloudWatch observability, and alerts; established incident-response runbooks, an on-call rotation, and post-mortems for production AWS services.
- Migrated complama.com DNS from WordPress to AWS Route 53 with zero downtime for 1,000+ monthly visitors, preserving web and email services through record validation, health checks, and planned rollback procedures.

University of Michigan Transportation Research Institute (UMTRI) | Research Associate – AI/ML & Multi-Cloud

May 2025 - May 2026

Systems | Ann Arbor, MI

- Supported vehicle AI research on the Mcity Data Engine with repeatable AWS, Azure, and GCP environments for multi-TB camera archives and GPU workloads processing ~1.3M high-resolution samples/day on 8 H100 GPUs.
- Cut Data Engine environment setup ~70% with a one-command Python/boto3 and Terraform deployer that provisioned compute, IAM, networking, and storage, configured GPU drivers, validated video inputs, and exposed the FiftyOne visualization UI.
- Decreased post-deployment errors 40% with Python test automation that checked REST endpoints, JSON schemas, and SQL assertions in BigQuery, PostgreSQL, and DynamoDB, validating analytics jobs and data consistency across cloud environments.
- Secured research stakeholder sign-off on the multi-cloud architecture by translating functional requirements into design documents and diagrams covering data flows, IAM boundaries, and ingestion pipelines.

Getronics | Software Engineer – Full-Stack Public Cloud | Bengaluru, IN

Mar 2023 - Aug 2024

- Operated 100+ Windows/Linux servers across Azure and AWS, handling patching, diagnostics, and network troubleshooting; identified anomalies and security threats 40% faster with Azure Monitor, Network Watcher, and log-query diagnostics.
- Cut manual operations 50% by integrating OpsRamp AIOps into internal tooling to automate VM scaling, alerting, and runbook execution; used PowerShell and Bash scripts to streamline routine server maintenance and operational checks.
- Deployed and managed Docker and Kubernetes clusters for ML workloads, sustaining 99.9% uptime through proactive L1/L2 incident triage, documented recovery procedures, and clear handoffs to engineering teams.

LykiQ Labs | Cloud Data Engineer | Bengaluru, IN

Mar 2022 - Dec 2022

- Enabled near-real-time EV-charging telemetry with an MQTT/OCPP data pipeline on GCP; reduced data-collection latency 20% using Firestore, serverless Cloud Functions, and App Engine to process device data for monitoring and analytics.
- Supported a 40% increase in user transaction volume without security regressions by securing authentication and payment APIs with OAuth2, tokenization, encryption, and audit logging in collaboration with product and data science teams.

PROJECTS

DodgeIT – Instruction-Driven Feed Filter | Agent Harness Hackathon Winner

- Built a JavaScript Chrome extension that filters X and Reddit feeds from plain-language instructions using OpenAI structured outputs and TypeSafe's Jev classifier; kept scrolling responsive with caching and concurrent queues, verified by 89 automated tests.

SKILLS

Languages & Development: Python, Java, TypeScript, JavaScript/Node.js, SQL, Bash, PowerShell, React, FastAPI, REST APIs, Git

Cloud & Infrastructure: AWS, Azure, GCP, EC2, EKS, Lambda, S3, IAM, Docker, Kubernetes, Helm, Terraform, CloudFormation

AI/ML: Generative AI, LLM APIs, RAG, AI agents, computer vision, PyTorch, TensorFlow, scikit-learn, OpenCV

Data & Streaming: PostgreSQL, MySQL, MongoDB, Redis, DynamoDB, BigQuery, Firestore, Kafka, Spark/PySpark, pandas, NumPy

DevOps & SRE: CI/CD, GitHub Actions, Jenkins, Grafana, unit/integration testing, Claude Code, GitHub Copilot

Security & Systems: OAuth2, IAM, encryption, distributed systems, event-driven architecture, Agile, Scrum, SDLC

CERTIFICATIONS & ACHIEVEMENTS

- AWS Certified Cloud Practitioner - AWS
- Microsoft Certified: Azure Administrator Associate (AZ-104) - Microsoft