

KAMAL SHARMA

Solution Architect | Cloud Platform Engineering & AI

Location: RTP, North Carolina, USA
Tel: +1 (646) 283-5718
Email: contacts@kamalsharma.me
Web: kamalsharma.me · [LinkedIn](#) · [GitHub](#)

SUMMARY

Senior Software Engineer with 13+ years building enterprise platforms, distributed systems, and cloud infrastructure on AWS and Google Cloud. Leads platform modernization across AWS and Amazon EKS using Terraform, Kubernetes, observability, cloud security, and CI/CD automation, while maintaining deep hands-on Java and microservices expertise from over a decade of production systems. Extends that platform background into Generative and Agentic AI: built an enterprise RAG platform and an AI-driven infrastructure drift-detection agent that turn LLMs, vector search, and automation into practical engineering tools. Works fluently across application, SRE, platform, and security teams to turn ambiguous requirements into reliable, scalable systems.

CORE SKILLS

Cloud	AWS, Google Cloud Platform	AWS Services	EKS, EC2, IAM/IRSA, VPC, Route 53, ALB, ACM, CloudWatch, RDS, OpenSearch, Secrets Manager
Containers & IaC	Kubernetes, Docker, Helm, Terraform	CI/CD	Jenkins, GitHub Actions, Git
Observability	Grafana, Prometheus, Alertmanager, CloudWatch, OpenSearch/Kibana	Data & Messaging	Kafka, Pub/Sub, Spanner, BigQuery, Dataflow, Apache Beam, SQL/NoSQL
Generative / Agentic AI	RAG, LLMs, Vector Search, AI Agents, Prompt Engineering	Languages & Frameworks	Java, Kotlin, Python, Spring Boot, Micronaut, REST/Microservices
Platform Engineering	Reliability (SRE), Capacity Planning, Incident Management	Leadership	Team Leadership (5 engineers), Cross-team Collaboration, Mentoring

EXPERIENCE

Nov 2024 - Present
Research Triangle Park, NC

Solution Architect – Cloud Infrastructure & Platform Engineering Cisco Systems · TPSCRM Platform

- Lead platform modernization for Cisco's TPSCRM platform across AWS and Amazon EKS, improving reliability, observability, security, automation, and deployment consistency.
- Built reusable platform engineering capabilities using Terraform to standardize provisioning of EKS, IAM, IRSA, ALB, Route 53, ACM, Secrets Manager, and CloudWatch across enterprise environments.
- Architected and managed production AWS EKS clusters supporting enterprise microservices with RBAC, network policies, Prometheus/Grafana observability, CI/CD automation, and high availability.
- Build enterprise Grafana dashboards integrating Prometheus, CloudWatch, and OpenSearch for centralized platform visibility and faster incident detection.
- Develop automated Kubernetes sizing recommendations from historical Prometheus metrics, improving cluster utilization without sacrificing performance.
- Build AI skill-based operational runbooks that combine Kubernetes, Prometheus, CloudWatch, OpenSearch, and AWS context to investigate production issues and guide root-cause analysis.
- Strengthen security and CI/CD through Secrets Manager adoption, least-privilege IAM/IRSA, Jenkins pipeline improvements, and GitHub Actions modernization.

HACKATHON

AI-Powered Infrastructure Drift Validation Agent

Built an agent integrating Terraform, AWS, and Generative AI to detect infrastructure drift, classify changes, assess operational risk, and generate remediation recommendations with human-in-the-loop review.

Dec 2018 - Sep 2024
Philadelphia, PA

Senior Software Engineer Urban Outfitters (URBN)

- Designed Docker-based microservices and backend APIs for internal business operations using Java, Micronaut, Google Cloud, Kafka, and Spanner.

Nuuly Thrift – resale marketplace, 18,000+ listings

Jul 2020–Nov 2024

- Built order, invoice, cart, and backend microservices handling 100,000+ requests per day.
- Built ETL pipelines processing and visualizing Stripe financial data to surface seasonal sales trends using Kafka, Spanner, and Spring.

Nuuly Rental – subscription platform

Dec 2018–Jun 2020

- Designed a warehouse database and reporting ETLs with Dataflow, SQL, and BigQuery to process terabytes of streaming and batch data.
- Designed reporting dashboards that gave business and operations teams visibility into inventory, order processing, sales performance, and subscription growth.
- Developed microservices and backend APIs supporting core rental workflows, including order management and internal business operations.
- Built reliable streaming and batch data workflows that transformed operational data into analytics-ready BigQuery datasets.
- Led five offshore developers on delivery of an internal platform, coordinating design, implementation, and technical execution.

■ EXPERIENCE (CONTINUED)

Feb 2013 - Nov 2018
Gurgaon, India

Technical Lead

MediaAgility India (now part of Persistent Systems)

- Contributed to the architecture and development of Dista, an AI-enabled location-intelligence platform for delivery, field service, field sales, logistics, and fleet operations.
- Designed core location-intelligence capabilities serving millions of transactions through Google Maps APIs and cloud-native backend services.
- Developed microservices and APIs using Google Cloud Endpoints, Java, Struts, and SQL/NoSQL databases for enterprise customer workflows.
- Built streaming and batch data pipelines with Dataflow and Cloud Composer to process operational and location data at scale.
- Implemented Angular user-tracking features integrated with Google Maps and collaborated across frontend, backend, and data engineering workstreams.
- Led internal technical training and mentoring sessions that helped new developers learn the platform and become productive faster.

■ SELECTED AI & SOFTWARE PROJECTS

Enterprise RAG Platform

github.com/ksharma-git/enterprise-rag-platform

FastAPI · Streamlit · PostgreSQL (pgvector) · Ollama (Llama 3.2) · Docker Compose

Designed and developed a production-ready Retrieval-Augmented Generation (RAG) platform for enterprise knowledge management. Implements document ingestion, chunking, embeddings, hybrid vector/keyword retrieval, citation-backed responses, streaming chat, and persistent conversations using locally hosted LLMs.

Little Wins – Kids Routine Planner

littlewins.kamalsharma.me

Next.js · React · TypeScript · Tailwind CSS · Cloudflare Workers & D1 · Gemini 2.5 Flash

Designed and developed a full-stack AI-powered family productivity platform enabling parents to manage routines, rewards, shared calendars, and streak tracking. Integrated Gemini AI to generate personalized weekly plans based on real family context while enforcing secure parent-only administrative controls.

■ EDUCATION & CERTIFICATIONS

B.Tech, Computer Science

Rajasthan Technical University, Jaipur, India · 2007–2011

- Google Cloud Professional Cloud Architect
- Google Cloud Professional Cloud Developer
- Google Cloud Associate Cloud Engineer

■ COMMUNITY & PUBLICATIONS

- Shares software engineering and cloud learnings through a [YouTube channel](#) reaching approximately 57,000 viewers.
- Published a free [System Design Fundamentals](#) course on Udemy, used by more than 7,500 students.