

Suraj Iyer

www.surajiyer28.com | 812-822-7872 | surajiyer@iu.edu | linkedin.com/in/surajiyer28 | github.com/surajiyer28

Summary

Software Engineer with 2+ years building and deploying cloud-native microservices on AWS using JavaScript, Node/Express.js, and TypeScript/React. Experienced in CI/CD through automated testing, service-level monitoring, and distributed system reliability. Quick to close skill gaps and ship production-ready software across fast-moving environments.

Education

Master of Science, Computer Science

Aug 2024 – May 2026

Indiana University, Bloomington (GPA: 3.7/4)

Bloomington, IN

- **Achievements:** Winner at Innoquest 2025 (Third Eye), Winner at Plutus x ServiceNow Agentic AI Case Competition (Agentic Leakage Detection), Winner at Luddy Hackathon 2026
- **Coursework:** Software Engineering, Applied Algorithms, Applied Machine Learning, Computer Vision, Computer Networks

Bachelor of Engineering, Information Technology

Aug 2018 – May 2022

University of Mumbai (GPA: 9.22/10)

Mumbai, India

- **Achievements:** Certificate of Merit (2019), Award for Creative Excellence (2022)
- **Coursework:** DSA, Object-Oriented Programming, Operating Systems, Cloud Computing, Mobile App Development, DBMS

Skills

Languages: Java, Python, JavaScript/TypeScript, HTML/CSS, SQL, C

Frameworks & Tools: Spring Boot, REST APIs, Gradle, Node.js, Express.js, FastAPI, React/Next.js, Angular, Postman, VS Code

Databases & Cloud: PostgreSQL, MongoDB, ChromaDB, AWS (EC2, S3, EKS, CloudWatch), GCP(Compute Engine, VertexAI), Azure DevOps(AKS)

DevOps & Methodologies: Git, Docker, Kubernetes, CI/CD (GitHub Actions), JMeter, Agile/Scrum (Jira), Microservices

LLMs/AI tools: LangGraph, LangChain, RAG, Embedding Models, Claude API, Gemini API, OpenAI API, Cursor, Claude Code

Experience

Software Engineer

Jun 2022 – Jul 2024

Merkle

Mumbai, India

- Designed and developed microservices and REST APIs using Node/Express.js and PostgreSQL, that aggregated survey data from Qualtrics and Alchemer, handling 100,000+ daily requests per service to enable faster access to survey insights for project teams
- Containerized and deployed distributed microservices on AWS (EC2, EKS) using Docker, integrating service-level monitoring via CloudWatch to support reliable, repeatable deployments across 150+ engineers and multiple independently deployable services
- Powered analytics across 200+ monthly campaigns by building a real-time reporting dashboard with multi-tier components and responsive design using TypeScript, React, HTML and CSS, enabling teams to track survey responses and quota progress
- Owned CI/CD pipeline infrastructure using GitHub Actions, authoring automated unit and integration test suites that reduced deployment lead time by 20% and maintained release reliability across containerized services
- Collaborated with cross-functional Agile/Scrum teams to define service-level objectives, implementing production monitoring that maintained 99.9% uptime across containerized deployments while conducting code reviews to ensure consistent release quality
- **Achievement:** Recognized with Climb High Award for owning end-to-end delivery of 5+ concurrent projects in a single quarter, demonstrating quick learning with minimal oversight, contributing to a 7x revenue return on quarterly compensation

Full Stack Developer Intern

May 2021 – Dec 2021

Careeristic

Mumbai, India

- Co-developed a full-stack web application using React, Node.js, Express, MongoDB and AWS (EC2, S3) to assist 20,000+ high school students access career insights through video podcasts and blogs.
- Led the development of a Content Management System with secure REST APIs, optimized database schemas, and role-based content workflows, eliminating 90% of content update errors and reducing update turnaround time by 75%
- Led migration from React to Next.js, improving page load performance by 30%, while building reusable component libraries and driving technical decisions across cross-functional sprints with designers and content stakeholders

Projects

Third Eye - LLM-Powered Monitoring System | *FastAPI, LangChain, ChromaDB*

Apr 2026

- Built an end-to-end LLM-powered multi-agent monitoring system processing 4 simultaneous camera feeds via async FastAPI, WebSocket Streaming and a React frontend, using Gemini API for structured scene analysis and Claude API for multi-step reasoning
- Implemented semantic event detection using RAG with ChromaDB and sentence-transformer embeddings, enabling natural language monitoring task definition and scene analysis matching, delivering sub-5 second alerts when user-defined conditions were met

Agentic Insurance Claim Leakage Detection | *FastAPI, ChromaDB, LangGraph*

Mar 2026

- Architected a 4-agent LangGraph system (Similarity, Reserve Audit, Invoice Audit, Portfolio Analytics) with tool calling for autonomous DB read/write operations, integrating with ServiceNow's existing data model to surface leakage flags across insurance claims
- Implemented RAG pipeline using ChromaDB and sentence-transformer embeddings, enabling the Similarity Agent to benchmark incoming claims against historical distributions across geography, coverage, and claim type across 4+ LOBs

Microservices-based Room Reservation System | *Java, Spring Boot, Angular, Docker, Kubernetes*

Dec 2025

- Designed and deployed a 7-service Spring Boot architecture (Edge, Reservation, Room, User, Notification, Config, Event Broker) with an Angular frontend, independent databases per service, and end-to-end system orchestration on Azure AKS with Docker and Kubernetes
- Executed load testing using JMeter simulating 100,000 concurrent users, validating Horizontal Pod Autoscaling under peak load and fault-tolerant Kubernetes deployments on Azure AKS, with CI/CD automation via GitHub Actions and JUnit test coverage