

# Rahul Shah

San Francisco, CA | +1(516)979-5019 | [rshah107@asu.edu](mailto:rshah107@asu.edu) | [Portfolio](#) | [Github](#) | [LinkedIn](#)

## EXPERIENCE

### Software Engineer

May 2024 – Present

*Solutions Unified LLC.*

*Califon, NJ*

- Designed a scalable **file upload pipeline** using pre-signed **S3** URLs with chunked uploads, S3 multipart support and **Lambda** triggers, reducing core service load and improving file processing latency by **35%**.
- Automated multi-service (**Elastic Beanstalk**, **S3**, **DynamoDB**) deployment using **AWS CDK**, reducing customer environment setup time by **75%** through seamless integration with **CloudFormation**.
- Developed an admin dashboard using **React**, **Tailwind**, and **Go** backend services to monitor real-time usage metrics and system logs, enhancing system observability and reducing debugging time by **60%**.
- Built a **notification service** to process usage data from client environments via **SNS** and **SQS**, triggering Lambda-based metering and **Stripe** integration for billing, reducing errors by **25%**.

### AI Full-Stack Developer

Jan. 2024 – Jan. 2025

*Enterprise Technology(Mastercard Foundation), Arizona State University.*

*Tempe, AZ*

- Architected three full-stack platforms serving over **10,000+** users to build, compare, and interact with various **LLMs** using **AWS Bedrock**, **OpenSearch**, and **Llamaindex** with audio and image support.
- Engineered a **voice-enabled assistant** by integrating **Whisper** for transcription and **Amazon Polly** for speech synthesis, allowing seamless bi-directional audio interaction with LLMs for accessibility-focused use case
- Architected a modular backend system supporting plug-and-play integration of third-party **LLM** providers, enabling flexibility in model experimentation without downtime.
- Created a **multi-agent** orchestration system using **LangGraph** and **FastAPI**, enabling chained LLM workflows for planning, memory, and tool use, reducing user interaction steps by **50%**.

### Software Engineer II (Full-Stack)

June 2020 – July 2023

*Allegion (Schlage).*

*Bengaluru, India*

- Drove the migration of an internal API from **REST** to **GraphQL**, optimizing query complexity and efficient data fetching, reducing the need for multiple API calls & improving developer efficiency by **35%**.
- Led the integration of an automated **CI/CD pipeline** using **Jenkins**, **Docker**, and **Kubernetes**, reducing deployment cycles by **40%** and implementing blue-green deployment for zero-downtime production updates.
- Developed an onboarding workflow using **React** and **Redux**, integrating REST APIs and custom hooks, improving task completion rates and reducing onboarding time by **60%**.
- Built a middleware-based request tracing system in **Node.js** using OpenTelemetry, enabling distributed tracing across microservices and reducing debugging time by **40%**.
- Designed and implemented an **end-to-end testing** framework using **Selenium WebDriver** and **Mocha**, enabling automated browser testing reducing manual QA efforts by **70%**.

## PROJECTS

**Conversense** ([link](#)) | *Go, React, PostgreSQL, LiveKit, Google Gemini, AWS S3, Inngest* Oct 2025 – Nov 2025

- Built an AI-powered meeting platform with real-time video conferencing, automated transcription, and intelligent summarization using LiveKit and Google Gemini, featuring AI agents for in-meeting assistance and post-meeting chat interaction.

**Vistruct** ([link](#)) ([Demo](#)) | *React, TypeScript, FastAPI, PostgreSQL, Docker, GenAI* Jul. 2025 – Aug 2025

- Developed a full-stack educational video platform that is currently used by 20+ professors from my undergrad university, leveraging AI for automated content generation, featuring real-time chat-based script development, AI voice-over synthesis, and Manim animation integration.

**InfoFetchBot** ([link](#)) | *Python, Discord API, Open AI, AutoGen, NotionAPI* Sep. 2024 – Oct 2024

- Created a Discord bot that processes incoming messages, triggers web searches, and scrapes website content using AI agents, delivering concise summaries directly into Notion for data organization and retrieval.

## TECHNICAL SKILLS

**Languages & Frameworks:** Go, Java, TypeScript/JavaScript, Python, FastAPI, React, Node.js, Spring Boot

**Cloud & Tools:** AWS (EC2, API Gateway, S3, ECS, Bedrock), Kubernetes, Docker, Git, Jenkins, Selenium

**Data & Messaging:** DynamoDB, PostgreSQL, MySQL, MongoDB, Redis, Kafka, RabbitMQ, SQS, SNS

## EDUCATION

**Arizona State University**

Tempe, AZ

*Master of Science in Computer Science GPA: 4.0/4.0*

*Aug. 2023 – May 2025*