

# Shitij Mathur

+1 (623) 277-7854 | [shitijmathur@gmail.com](mailto:shitijmathur@gmail.com) | [linkedin.com/in/shitijmathur](https://linkedin.com/in/shitijmathur) | Tempe, AZ

## EDUCATION

### Master's in Computer Science

Arizona State University – GPA: 4.0

Coursework: Statistical Machine Learning, Data Processing at Scale, Cloud Computing, Data Mining

August 2024 – May 2026

Tempe, AZ

### Bachelor of Technology in Electronics & Communication Engineering

Manipal Institute of Technology

July 2015 – July 2019

Karnataka, India

## TECHNICAL SKILLS

**Languages:** Python, JavaScript, SQL, Node.js, Bash, HTML/CSS

**AI/LLM:** RAG Pipelines, Vector Databases (OpenSearch), LangChain, Amazon Bedrock, AI Agents, Prompt Engineering

**AI Dev Tools:** Kiro, Cursor, Claude Code, GitHub Copilot

**Backend & Web:** REST APIs, WebSockets, HTTP/JSON, Serverless Architecture, Microservices, Event-Driven Architecture

**Data:** Apache Airflow, Pandas, PostgreSQL, DynamoDB, Redshift, OpenSearch

**Infrastructure:** AWS (Lambda, API Gateway, S3, SQS, ECS, CDK), Docker, Kubernetes, Git

**Certifications:** AWS Machine Learning – Specialty, AWS Solutions Architect – Associate

## PROFESSIONAL EXPERIENCE

### Platform Engineer Intern (AI & Data Systems)

Quantiphi

May 2025 – January 2026

Boston, MA (Remote)

- Designed and built reusable **AWS CDK** modules in Python to provision platform components for multiple domain teams, enabling config-driven onboarding where new domains self-served by updating config files.
- Orchestrated complex **ELT** workflows using **Apache Airflow** (MWAA) DAGs, dynamically generating tasks from dbt models to transform raw data into cleaned, aggregated tables for downstream analytics.

### Senior Platform Engineer (AI & Web Systems)

Quantiphi

January 2021 – May 2024

Mumbai, India

- Designed and deployed a **Retrieval-Augmented Generation (RAG)** pipeline using **OpenSearch** as the vector database and **Amazon Bedrock** for LLM serving, implementing **embedding, chunking, and indexing** strategies to build a **domain-specific chatbot** that reduced information retrieval time by 95%.
- Built an automated serverless document ingestion pipeline where S3 uploads trigger async **SQS** workers on **ECS** to concurrently parse, chunk, and index 500+ technical PDFs, implementing dead-letter queues and retry logic reducing manual indexing effort by 90%.
- Developed a real-time **WebSocket-based** chat application using **API Gateway, Lambda, and DynamoDB** for session state, integrated with **Amazon Lex** for NLU to automate 80% of support queries without agent assistance.
- Led **cross-functional technical discussions** with client engineering teams to define product scope, guided documentation of engineering processes, and mentored junior engineers on system architecture and API design patterns.

### Platform Engineer (Backend & Cloud Systems)

Quantiphi

July 2019 – January 2021

Mumbai, India

- Built an **event-driven pipeline** in Python using **SQS, Lambda, ECS, and PostgreSQL** to process financial risk reports asynchronously, reducing manual processing effort by 70%.
- Developed **RESTful APIs** in Python with token-based authentication and rate limiting, enabling secure, reliable access to ML inference endpoints for downstream applications.

## PROJECTS

### Zoom Momentum – Active Learning Assistant | React, TypeScript, Node.js, Zoom Apps SDK

Nov 2025 – Present

- Built a full-stack **Zoom Apps SDK** application in React and TypeScript that transforms passive virtual lectures into active learning environments with 4 AI-powered features: real-time topic segmentation, AI quiz generation, check-in polls, and personalized post-class recovery packs.
- Integrated **LLM serving** via OpenAI-compatible API with streaming transcript analysis for real-time AI responses.

### TrustMed AI – Healthcare Knowledge Chatbot | Amazon Bedrock, OpenSearch, Llama 3.1

October 2025

- Built a **RAG-based** healthcare chatbot using **Amazon Bedrock** Knowledge Bases and **OpenSearch**, generating evidence-backed responses with source citations from medical literature to build a **domain-specific chatbot**.

### PassGen – Password Generator Chrome Extension | JavaScript, Chrome APIs

October 2024

- Built a **Chrome extension** that generates unique, strong passwords using multi-layered hashing of website name, username, and secret key. Won 2nd Place in Amazon's Online Safety Track at Hack SoDA 2024.

### Lucky Loo – Multi-Agent AI System | Bedrock, Strands Agents SDK, Flask

November 2025

- Built a multi-agent AI system using **AI APIs** with an Agents-as-Tools pattern, orchestrating a hierarchical jury of LLM sub-agents for real-time facial expression analysis with **multimodal capabilities** and structured decision logic served via a **Flask API**.

## HONORS & ACTIVITIES

### Zoom Fellowship | ASU Next Lab

November 2025 – Present

- Selected for a competitive innovation program building next-generation educational applications on Zoom's developer platform using AI and real-time APIs.

### AWS Road to re:Invent 2025 Hackathon

November 2025

- Selected among 50 developers globally for a live-streamed hackathon across the LA-to-Vegas route.