

Srujay Reddy Vangoor

Location: Sacramento, CA
Phone: +1 (916) 940-5986

Email: srujayreddyv@icloud.com
LinkedIn: linkedin.com/in/srujayreddyv
GitHub: github.com/srujayreddyv

SUMMARY

- Full Stack Software Engineer with 6+ years of experience building production full stack applications, scalable backend APIs, and cloud native systems. Strong in Python with FastAPI, React, and AWS, with deep expertise in system design, interoperability architectures, and high throughput distributed services. Known for modernizing legacy systems, designing reliable data pipelines, and integrating GenAI into production systems, including LLMOps pipelines, RAG architectures, and multi agent workflows that automate complex operational and compliance processes.

SKILLS

- **Backend & APIs:** Python (FastAPI, Django, Flask), Node.js (Express), Go (Golang), C# (ASP.NET Core), PHP (Laravel), REST, GraphQL, gRPC, WebSockets, Swagger, OpenAPI, Postman
- **Frontend & Mobile:** React, Next.js, Nuxt.js, TypeScript, JavaScript, React Native, Angular, Bootstrap
- **Cloud & DevOps:** AWS, Azure, GCP, Docker, Kubernetes, Terraform, CI/CD (GitHub Actions, Azure DevOps), CloudWatch, Datadog, Grafana, Prometheus
- **Data & Analytics:** PostgreSQL, MS SQL Server, DynamoDB, Redis, Snowflake, Amazon Redshift, Airflow, Kafka, Spark, Elasticsearch, Power BI, Tableau, Amazon QuickSight
- **AI & ML:** AWS Bedrock, LangChain, CrewAI, Hugging Face Transformers, RAG Pipelines, Vector Databases (FAISS, Pinecone, OpenSearch), PyTorch, TensorFlow, MLflow, Scikit learn

WORK EXPERIENCE

- | | |
|--|----------------------------|
| Software Engineer | Sacramento, CA (Hybrid) |
| ■ <i>California Department of Developmental Services (DDS)</i> | <i>Feb 2025 – Present</i> |
| ○ Increased statewide data processing throughput by 30% by architecting and deploying 31 HL7 FHIR compliant REST APIs, enabling cross program interoperability and real time data exchange across State Operated Facilities. | |
| ○ Scaled FastAPI services handling 20K plus daily requests on AWS ECS and Lambda, implementing Cognito based authentication and EventBridge driven workflows to achieve 99% uptime for mission critical healthcare integrations. | |
| ○ Modernized legacy COBOL, JCL, and Db2 batch systems into Python based services, mapping historical datasets to FHIR aligned models, cutting nightly processing time by 60% and eliminating more than 200 manual operational steps. | |
| ○ Reduced analyst review effort by 85% by designing an LLMOps QA platform using AWS Bedrock, LangChain, FAISS, and Titan embeddings to automate claims policy review, utilization summaries, and regulatory document search. | |
| ○ Delivered a production multi agent GenAI system using FastAPI, React, CrewAI, and Claude, providing source grounded answers across healthcare and compliance data for internal users. | |
| ○ Built and deployed a React and TypeScript portal with Express and PostgreSQL on Terraform provisioned AWS infrastructure, reducing AI Assessment entry time from 30 to 10 min through automation and UX optimization. | |
| Software Engineer | Sacramento, CA (Remote) |
| ■ <i>California Department of Conservation (DOC)</i> | <i>Jan 2024 – Feb 2025</i> |
| ○ Built a production seismic hazard modeling platform using Python, Perl, PHP, D3.js, and OpenQuake, delivering ground motion predictions for California fault zones with response times under 3 seconds. | |
| ○ Designed and automated large scale seismic data pipelines processing 10+ GB per day, improving ingestion efficiency and reducing data retrieval time by 50% to support high throughput detection workflows. | |
| ○ Strengthened system security for 60+ remote monitoring stations by developing an authentication gateway with AES encryption, AWS IAM roles, and AWS Secrets Manager, passing all compliance audits with zero findings. | |
| ○ Improved reliability of geospatial services by resolving ArcGIS server performance issues and enhancing map rendering with GeoServer and Leaflet.js, reducing downtime and improving system stability. | |
| Software Application Developer Intern | Sacramento, CA |
| ■ <i>Population Research Center (PRC), Sacramento State</i> | <i>Jan 2023 – Jun 2024</i> |
| ○ Built a staff management system using C#, Entity Framework, and SQL Server, optimizing stored procedures, indexing, and query performance to reduce administrative workload by 40%. | |
| ○ Designed and automated health data ETL pipelines using .NET, SQL Server, and PowerShell, migrating legacy Access and Oracle datasets into centralized SQL analytics stores for Power BI reporting under HIPAA compliant workflows. | |

- Developed interactive data entry tools and analytics dashboards for statewide public health studies, increasing data collection efficiency by 60% through automated validation and quality checks.
- Supported high volume CATI survey operations with system upgrades, PowerShell diagnostic utilities, and HIPAA compliant workflows, reducing survey delays by 30%.

■ Software Engineer

Hyderabad, IN

Human Sciences Research Group (HSRG), IIIT-H

Jan 2021 – Aug 2022

- Designed a cloud native archival and data processing platform for social media and news data, enabling realtime analysis of more than 1 million data points during the Indian Farmers Protests.
- Fine-tuned multilingual BERT and IndicBERT models using Hugging Face Transformers for sentiment analysis and topic modeling to achieve an 8% accuracy improvement over baseline models.
- Developed secure REST APIs with C# (ASP.NET Core) to support high volume data ingestion and metadata management, with interactive React dashboards for exploratory data analysis.
- Containerized services with Docker and implemented automated build, test, and deployment pipelines using GitHub Actions, ensuring reproducible environments and reducing release cycles 40%.

■ Software Engineer

Hyderabad, IN

Change and Continuity in Spiti Valley, ICSSR Sponsored Project

Jul 2018 – Dec 2020

- Engineered a cross platform GIS system using Python (Django, PostgreSQL, PostGIS) to map 280+ Buddhist heritage sites across 25 villages, enabling spatial search, analytics, and interactive visualization.
- Built responsive React web and offline first React Native mobile clients with Leaflet.js, achieving 60% faster load times in low bandwidth field environments.
- Designed and deployed RESTful APIs with Django REST Framework to support spatial and attribute based queries with pagination and caching, reducing query latency and improving scalability.
- Developed Python based ETL and spatial data processing workflows that integrated survey, census, and archival datasets into a single geospatial repository, improving metadata accuracy by 40% and enabling near realtime updates.

EDUCATION

■ California State University, Sacramento

Sacramento, California

Master of Science (MS) in Computer Science

Aug 2022 – Jan 2025

■ International Institute of Information Technology (IIIT-H)

Hyderabad, India

Bachelor of Technology (BTech) in Computer Science

Aug 2015 – July 2019

RELEVANT PROJECTS

■ FastChat - Real-Time Chat Application

- Developed a Dockerized full stack application with Python (FastAPI), React (TypeScript + Material UI), and PostgreSQL, supporting realtime messaging, user presence, and typing indicators via WebSockets.
- Automated CI/CD pipelines with GitHub Actions, enforcing rate limiting, input validation, structured logging, PyTest based integration tests, and performance monitoring for production reliability.

■ Local Retrieval Augmented Generation (RAG) Implementation

- Built a local RAG pipeline using Hugging Face Transformers for embeddings and Pinecone for vector search enabling context aware semantic retrieval across a 1,200 page nutrition textbook with 93% accuracy.
- Designed an evaluation workflow in MLflow to benchmark retrieval performance and refine prompt response relevance, improving contextual precision and ensuring readiness for cloud deployment.

■ TaskHub - Task Management Platform

- Architected a JWT secured, modular React (NextJS) SPA backed by Express microservices and PostgreSQL, enabling realtime collaboration across 100+ workspaces.
- Implemented caching, load balancing, and blue green deployments through Azure DevOps pipelines, achieving 99% uptime and 70% faster API responses under concurrent load.

■ Sacverse - Augmented Reality (AR) Campus Tour App

- Prototyped an AR campus navigation app in Figma, featuring geospatial search, route planning, and customizable avatars optimized for low bandwidth environments.
- Conducted UX research with 50+ pilot users, improving AR overlay accuracy and interaction design, leading to a 40% rise in usability scores and higher engagement.