Satya Sai Srinivas Yendru

Software Engineer building cloud-native, AI-powered systems

satyasaisrinivas955@gmail.com Boston, MA github.com/Satya-Sai-Srinivas LinkedIn

Authorized to work in the U.S. (F-1 OPT, active since June 2026)

EDUCATION

Master of Science — Informatics (Cloud Computing) Sep 2024 – Apr 2026

Northeastern University · Boston, MA · GPA 4.00 / 4.00

Bachelor of Technology — Electronics & Communication Engineering Aug 2019 – Jul 2023

St. Martin's Engineering College (JNTUH) · Hyderabad, India

EXPERIENCE

Software Engineer Aug 2023 – Aug 2024

Suntek Corp Solutions Pvt. Ltd. · Hyderabad, Telangana, India

- Developed key backend features for a campus placement LMS, managing data pipelines and routing architecture.
- Cut monthly AWS costs by 30% by moving from always-on EC2 instances to an event-driven architecture using AWS Lambda and auto-scaled RDS.
- Implemented automated failover and RDS snapshot scheduling, achieving 99.9% availability and an RTO under 15 minutes during peak load.

SELECTED PROJECTS

CineCrack

AI-powered movie discovery platform with a natural-language recommendation agent.

- Built a production AI recommendation agent bridging semantic understanding and structured search, delivering sub-second natural-language responses.
- Implemented LLM tool-calling with SSE streaming, autonomously converting user queries into complex database queries across 5 filter dimensions.

LangChain pgvector GCP Cloud Run Server-Sent Events Python

Agent-Dome

Autonomous threat-response agent for simulated cybersecurity incidents.

- Engineered an autonomous AI agent to interpret, analyze, and respond to simulated cybersecurity threats.
- Structured LLM reasoning loops so the agent autonomously routes tasks and executes multi-step logic chains.

LangChain FastAPI Python

AI-Driven CPL Assistant

AI system that automates Credit-for-Prior-Learning academic credit evaluation.

- Owned an AI-driven credit-evaluation system end to end, replacing 4+ hours of manual review per cohort with dynamic prompt pipelines in Azure OpenAI.
- Built a RAG pipeline (Python, PyPDF2) to parse and inject structured evidence into the LLM context at inference time.

Azure OpenAI Python RAG PyPDF2

SKILLS

Languages	Python · TypeScript · JavaScript · Java · C · C#
AI / LLM	RAG (Retrieval-Augmented Generation) · Agentic Workflows · LangChain · Azure OpenAI · Prompt Engineering · LLM Evaluation · Conversational AI Design
Backend & Frameworks	FastAPI · Pydantic · SQLAlchemy · Node.js · React
Cloud	AWS (Lambda, EC2, S3, VPC) · Azure · GCP
Data	PostgreSQL · pgvector
DevOps & Tools	Docker · Git · CI/CD · Visual Studio

Patent: "Deep Learning Based Inspection Device for Sheet Having Tissue and Image Processing" — Indian Patent Application No. 202341026700, Journal 18/2023 (Mar 2023)