

Utkarsh Garg

631-215-6568 | utgarg1993@gmail.com | linkedin.com/in/utk1801 | https://utk1801.github.io/

Senior Backend Engineer with 9+ years driving enterprise identity scale and agentic AI infrastructure, bringing deep expertise in high-availability IAM platforms, with hands-on architecture of zero-to-one AI agent security frameworks (cross-domain delegation, least-privilege access). Targeting Staff Backend, Platform, and AI Infrastructure roles.

Experience

Oracle America

Feb 2021 – Present

Senior Member of Technical Staff

Santa Clara, California

- Engineered resilient backend infrastructure for Identity and Access Management (IAM), managing over 10 million daily authentications and ensuring exceptional system availability with 99.999% uptime.
- Led the architecture and design of next-gen AI-agent authentication/authorization capabilities by establishing trust boundaries, building CIMD-based metadata validation frameworks, and enforcing least-privilege controls—safely scaling dynamic agent client provisioning by 3x, while reducing identity onboarding bottlenecks..
- Built cross-domain agent delegation on ID-JAG identity assertions and RFC 8693 token exchange, so agents act on behalf of users across trust domains, with DPoP sender-constrained tokens to prevent token replay.
- Built network-aware OAuth controls and geofencing into login workflows, blocking 99% of cross-border attacks and cutting credential fraud by 45%
- Engineered event-driven, server-side session termination for high-risk identity events, cutting incident response times by 80%.

Sony Interactive Entertainment

May 2020 – Aug 2020

Software Engineering Intern

San Diego, California

- Developed an automated vendor-onboarding system on AWS to integrate new payment service providers, improving operational throughput by 50%.
- Proposed a chargeback orchestration flow, reducing processing errors and enhancing dispute resolution efficiency.

Visa Inc.

July 2015 – July 2019

Associate Systems Analyst → Software Engineer

Bangalore, India

- Automated CI/CD pipelines for a high-availability Tier-0 gateway serving 20+ microservices, slashing provisioning cycles by 70%.
- Migrated manual gateway configurations to Terraform, reducing drift errors by 95% and cutting resource provisioning times from days to minutes.
- Fine tuned alert-monitoring and health check scripts, centralized dashboards and log aggregation pipelines using the ELK Stack/Prometheus

Projects

Recovering Visual Encodings from Chart Images | Python, Flask, AWS, OpenCV, D3.js

- Engineered an OCR and CNN pipeline with 92% classification accuracy that uses AWS Rekognition and Text-to-Speech to make visual charts accessible for visually impaired users.

FinLLM: Financial Q&A and Document Intelligence Platform | Airflow, ChromaDB, OpenAI SDK, Streamlit

- Developed an LLM-powered financial Q&A app (Streamlit, OpenAI SDK) utilizing an AirFlow ingestion pipeline and hybrid retrieval (ChromaDB, BM25) to deliver rapid insights from Fortune 500 earnings calls, with system telemetry monitored via Grafana.

Git Deployer | AWS S3/ECS, Docker, Node.js, Express, Redis, Socket.io

- Bootstrapped a zero-touch pipeline that provisions dynamic Docker containers to build GitHub repositories, upload assets to S3, and stream live build logs via Redis.

Skills

- Security & Identity:** OAuth 2.0, OIDC, SAML, SCIM, RBAC
- Languages:** Java, Python, TypeScript, Node.js, Go
- AI / LLM Engineering:** LangChain, LangGraph, RAG, Vector Search (QdrantDB, ChromaDB), MCP, LLM Evaluation
- Cloud & DevOps:** AWS, GCP, OCI, Docker, Kubernetes, Terraform, CI/CD
- Data & Observability:** PostgreSQL, Redis, FastAPI, Grafana, OpenTelemetry

Education

Stony Brook University, State University of New York

Master of Science in Computer Science (GPA: 3.77 / 4.00)

Aug 2019 - Dec 2020

Stony Brook, New York

SRM University, India

Bachelor of Technology in Computer Science (GPA: 9.52 / 10.00)

Aug 2011 - May 2015

Chennai, India