

HARSH GUPTA

Software Development Engineer II with nearly 5 years of experience building distributed systems at scale, leading cross-functional technical delivery, and designing cloud-native architectures. M.S. in Computer Science with specialization in generative AI.

CONTACT

LOCATION

Seattle, WA

PHONE

(925) 204-9495

EMAIL

imharsh26@gmail.com

LINKEDIN

linkedin.com/in/thatcoderguy

GITHUB

github.com/Harsh-Git-Hub

EDUCATION

M.S. Computer Science

CSU Sacramento | 2018-2021

GPA: 3.69/4.0

Certification in Intelligent Systems

B.S. Information Technology

Mumbai University | 2013-2016

GPA: 3.8/4.0

SKILLS

Languages: Java | Kotlin | TypeScript | Python | Node.js | Golang | Smithy | GraphQL

Cloud and Infrastructure: AWS CDK | CloudFormation | API Gateway | Lambda | ECS Fargate | IAM/STS | DynamoDB | SQS/SNS | CloudFront | Route 53 | CloudWatch | WAF | PrivateLink | AppConfig | GCP BigQuery

Systems and Architecture: Distributed Systems | Microservices | API Design | Event-Driven Architecture | Multi-Region and Sovereign Cloud | Authorization and Security Design

Engineering Practices: CI/CD Pipelines | Infrastructure as Code | Automated Testing | Monitoring and Observability | Cross-Team Technical Leadership

AI/ML: Large Language Models | Agentic AI | MCP Servers | GANs (WGAN-GP) | NLP | TensorFlow | Airflow | Spark

WORK EXPERIENCE

Software Development Engineer II

Apr 2025 - Present

Amazon Web Services, Bellevue, WA

- Led cross-team delivery of 8 services for a new cloud region as technical lead — defined the architecture and rollout plan, managed a team of 6 engineers, and coordinated dependencies across 5+ external platform teams. Delivered all services on schedule for the region's public launch.
- Designed and implemented a custom authorization layer resolving a critical security vulnerability where users could bypass access controls through an alternate API path. Evaluated 3 architectural approaches, performed impact analysis across 1,800+ affected accounts, and deployed incrementally to production with zero customer issues.
- Architected agentic AI workflows with MCP server integrations for CI/CD automation, enabling touchless deployment and auto-remediation across the team's services.

Software Development Engineer

Aug 2022 - Apr 2025

Amazon, Seattle, WA

- Designed and led migration from monolithic architecture to distributed microservices by implementing a CDN-based traffic routing layer as the service entry point — achieved zero-downtime cutover and improved multi-region availability.
- Built a version-agnostic GraphQL API client with token-based authentication, automatic retry with backoff, and error-translation middleware — reducing integration failures and developer toil across merchant-facing services.
- Designed and maintained scalable event-driven microservices on container and serverless compute, handling high-throughput event processing with end-to-end observability, automated deployments, and Infrastructure as Code.

Software Engineer - Data Platforms

Jan 2022 - Aug 2022

Bolt, San Francisco, CA

- Built and optimized data pipelines using BigQuery, Airflow, Spark, and Python for analytics and machine learning workflows. Contributed to an open-source protobuf-to-JSON-schema compiler plugin.

PROJECTS

Batoonee - GAN-Based Conversational AI

M.S. Project, 2021

- Developed a conversational question-answering system using Wasserstein GAN with Gradient Penalty and attention-based Seq2Seq architecture. Proposed and evaluated 3 distinct GAN architectures on Stanford CoQA dataset. Best model achieved BLEU score of 59.7.