

Divyam Ahuja

Systems & Software Engineer

☎ +91 7015600547 ✉ ahujadivyam@gmail.com 🌐 divyam.dev 📄 in/ahujadivyam

🔗 git.divyam.dev 📁 gh/ahujadivyam

Experience

Microsoft

Software Engineer

Noida

Mar 2025 - Present

Tech Stack: Python, TypeScript, React, Azure OpenAI, Claude Code, Copilot, Playwright

- Owned the end-to-end Responsible AI (RAI) evaluation pipeline, executing critical harms assessments to gate model releases. Engineered custom retry mechanisms and architectural optimizations, reducing average pipeline execution time by **50%**.
- Engineered an automated synthetic data generation framework using LLMs, generating **70,000+** diverse queries and multi-step interaction trajectories across **20,000+** real-world Excel workbooks to power large-scale model evaluation pipelines.
- Served as on-call Shield Engineer, driving infrastructure reliability across evaluation pipelines—resolving session instabilities, engineering deterministic failure classification, and automating bug workflows to maintain **>95%** uptime.
- Spearheaded evaluation observability by building KQL-driven health dashboards, modularizing queries to monitor infrastructure failures, clean run rates, and overall platform reliability.
- Designed and shipped RAI refusal handling for the UI-less agent evaluation path, ensuring policy refusals propagate correctly with full diagnostics and achieving evaluation parity between UI and headless execution modes.

Advanced Micro Devices (AMD)

System Software Designer

Bangalore

Aug 2022 - Feb 2025

Tech Stack: C/C++, Java (OpenJDK), Python, Django, React, SQLite

- Engineered a visualization dashboard and RESTful API for OpenJDK telemetry, increasing internal data accessibility by **40%**.
- Automated nightly Java benchmarking and tuning pipelines using machine learning models, achieving a **30%** reduction in manual effort and earning the **Spotlight Award**.
- Optimized JVM flags, GC policies, and runtime parameters for OpenJDK on High Performance Computing (HPC) servers, achieving a **2%** throughput improvement on SPECjbb benchmarks.

Chromium

Google Summer of Code Contributor (Open Source Developer)

Remote

May 2022 - Aug 2022

Tech Stack: C++, JavaScript/TypeScript, WebAssembly, WebAudio, WebWorkers

- Engineered a lock-free C++ ring buffer as a reference library for real-time audio streaming between WebAudio, WebAssembly, and WebWorkers, benchmarking multiple concurrency strategies.

Skills

Programming Languages: C/C++, Python, Go (Golang), Java, TypeScript, Bash

Frameworks & Tools: PyTorch, Azure OpenAI, React/Next.js, Django/Flask, Node.js, gRPC, Playwright

Infrastructure: Kubernetes, Docker, Linux/POSIX, WebAssembly, WireGuard, PostgreSQL, Redis

Projects

fastbin: a code sharing webapp | HTMX, Go (Golang), gRPC, PostgreSQL, Docker, Kubernetes

- Code sharing service with a gRPC microservices architecture, containerized and live-deployed on a self-hosted Kubernetes cluster.
- git.divyam.dev/divyam/fastbin, fastbin.lab.divyam.dev

Homelab | Raspberry Pi 5, WireGuard, Kubernetes, Rancher, MetalLB

- Personal infrastructure hosting containerized services, with WireGuard VPN tunneling to a public VPS for secure external access.

Education

Deenbandhu Chhotu Ram University of Science and Technology

Bachelor of Technology (B.Tech.) Computer Science

Haryana, India

Aug 2018 - Aug 2022

Certifications

- [MITx 6.86x — Machine Learning with Python: Linear Models to Deep Learning](#)

- CS50AI: CS50's Introduction to Artificial Intelligence with Python