

DHRUV ARORA

Platform Engineer | Kubernetes · GitOps · GCP · Internal Developer Platforms

Ahmedabad, India | +91 9898333360 | dhruvmarora6670@gmail.com | [LinkedIn](#) | [GitHub](#) | dhruv-arora.com

SUMMARY

Platform engineer who co-built Deploy Genie, a self-service internal developer platform (IDP) on GCP and Kubernetes that has run **1,000+ deployments** across **15+ environments** for **150 engineers**. Hands-on with GKE, Argo CD (GitOps), Helm, KEDA, Kyverno, Pub/Sub, and Python/FastAPI; focused on golden paths, guardrails, and deployment automation.

EXPERIENCE

Techolution | *AI Engineer – Platform Engineering* | Oct 2024 – Present

Deploy Genie – Internal Developer Platform

- Co-architected a multi-tenant, self-service IDP that deploys frontend, backend, and AI-agent services to GKE and Cloud Run, cutting average deployment time from **2-3 days to 15 minutes** across 10+ teams.
- Owned the FastAPI orchestration engine (9 stages on GKE, 5 on Cloud Run): CI builds, pod sizing from measured usage +25% headroom, health checks, Argo CD handoff, and retries that clean up partial resources.
- Co-designed the split of the platform's FastAPI monolith into three services (deployment orchestration, builds, deployment management/environment admin) so each uses the scaling method that fits its load.
- Bypassed Cloud Build's **600-trigger limit** with a GitHub App - Pub/Sub - expander pipeline replacing trigger-based CI: ordered per-deployment queues, a 45-retry dead-letter queue, and PR-summary gates.
- Implemented Argo CD GitOps delivery with pod-readiness gating and auto-generated Helm charts versioned in GitHub; Kyverno admission policies block out-of-band changes to prevent configuration drift.
- Shipped self-service autoscaling (CPU HPA or KEDA request-based) and per-environment ingress (NGINX, GCE Ingress, or Kubernetes Gateway API); the pipeline adapts routing, with **zero direct Helm access** needed.
- Designed MongoDB-backed ACL RBAC with **up to 3-level approval chains** gating prod and client deployments, plus Cloud Tasks-scheduled auto-reverts that roll trial releases back to the last known-good build.
- Built route-level audit logging across all lifecycle events (deploys, retries, ACL, config, CI pushes, promotions) with AI-generated change summaries and one-action revert to any prior stable version.

AI Infrastructure & Operations Tooling

- Built GPU ML infrastructure: model deployment on GPU-backed GKE node pools and Cloud Run, plus a training-job scheduler that auto-provisions VMs, scales vertically via persistent disk, and stops idle compute.
- Cut a production RAG chatbot's first-response time from **~40s to 8s** by streaming tokens to post-processing in chunks.
- Automated promotion of 90%+ of each AI agent's configuration across environments, cutting setup from **45 to 5 minutes**.
- Architected an event-driven incident-triage service (FastAPI, PostgreSQL, Pub/Sub) linking Slack/Teams to Jira/ServiceNow with severity-based escalation, load-tested at **5,000+ events/hour**.
- Root-caused a Python Kubernetes client bug in bulk deletions; added safeguards and swapped affected calls to kubectl.

Smollan | *Python Developer Intern* | Apr 2024 – Aug 2024

- Built an Amazon Textract + OCR extraction pipeline with **~95% accuracy** on structured product listings, and 15+ Flask API endpoints powering live kiosk-usage analytics dashboards.

Ernst & Young (EY) | *Data Analyst Intern* | Dec 2023 – Mar 2024

- Automated suspicious-behaviour detection on assessment clickstream data (SQL, Python), cutting manual querying by 20%; built a KD-Tree face-similarity engine that ran **60 billion** pairwise comparisons.

TECHNICAL SKILLS

Platform: Kubernetes (K8s), Helm, Argo CD, GitOps, CI/CD, KEDA, HPA, Kyverno, NGINX Ingress, Gateway API, Docker

Cloud (GCP): GKE, Cloud Run, Compute Engine, Cloud Build, Artifact Registry, Pub/Sub, Cloud Tasks, Vertex AI

Languages & Backend: Python, Bash, SQL, FastAPI, Flask, asyncio, Pydantic, LangChain

Data & Tools: MongoDB, PostgreSQL, Redis, BigQuery, Firestore, AWS (S3, Textract), GitHub, Jira

PROJECTS

Voice-Controlled Smart Home Agent | *Python, FastAPI, LangChain, Gemini, MQTT, Ubuntu*

- Reverse-engineered the vendor app's device-control protocol: intercepted its traffic with a MITM proxy and packet captures, identified MQTT as the control channel, and decoded payload formats to drive devices from Python.
- Built a LangChain + Gemini agent with Pydantic tool schemas that turns free-form voice commands into multi-device, multi-step actions, served via FastAPI to iPhone Shortcuts and traced with LangSmith.
- Self-hosted it 24/7 on a headless Ubuntu server with a DHCP-reserved IP, SSH workflow, and boot-time startup scripts.

EDUCATION

Manipal University Jaipur | B.Tech, Information Technology | CGPA: 8.17/10

July 2025