Aleksandr (Alex) Shamilyan

Senior Platform & Site Reliability Engineer

Yerevan, Armenia (UTC+4) · open to remote | alexshamilyan@gmail.com | linkedin.com/in/alexandr-shamilyan | aleksandrshamilyan.com

SUMMARY

Platform and site reliability engineer with 5+ years in DevOps (since October 2021). I take over production infrastructure, make it safe and observable, and turn it into code teams can ship on: OpenTofu and Terraform with Terragrunt and Atlantis, Kubernetes on EKS and GKE, private AWS networking, Databricks and SageMaker. On the AI platform of a global mobility company I own private DNS, the multi-region model-deploy path, SLO alerting and the recovery design, and I take production on-call. I led the DevOps team on a multi-tenant retail platform.

SKILLS

Cloud and networking: AWS (VPC, PrivateLink, Route 53, Cloud WAN, IAM, KMS, EKS, ECR, Lambda, API Gateway, Step Functions, S3, CloudFront), GCP (GKE, Cloud Run), Azure (AKS, App Service, Azure SQL)

Infrastructure as code: Terraform, OpenTofu, Terragrunt, Atlantis, Ansible

Containers and CI/CD: Kubernetes (EKS, GKE, AKS), Helm, Docker, Harbor, GitHub Actions, GitLab CI, Jenkins shared libraries, Azure DevOps, release-please

Reliability: SLO design and alerting, on-call, Prometheus, Alertmanager, Grafana, CloudWatch, Datadog, PagerDuty, disaster-recovery design, PostgreSQL HA (Patroni, HAProxy, Keepalived)

Security: Cloud security audits, least-privilege IAM, HashiCorp Vault, OIDC and workload identity, Keycloak, Trivy, distroless images, CVE remediation

Data and ML platforms: Databricks (Unity Catalog, Asset Bundles, cluster policies), MLflow, SageMaker, Kafka, Delta Lake, BigQuery, Liquibase

Languages and tools: Python, Bash, Go, TypeScript, Groovy; Claude Code skills and sub-agents for PR and Terraform-plan review, with a person approving every write

EXPERIENCE

SRE / DevOps / MLOps Engineer, DataArt

May 2022 – Present

Yerevan, Armenia. Client engagements below; client names withheld.

AI platform, global mobility company

May 2026 – Present

Platform / MLOps engineer in a small platform team. Author the team ADRs on security, recovery and platform limits; review Terraform changes as a merge gate; on the production on-call rota.

- Built the private DNS behind PrivateLink-only Databricks access: Route 53 private hosted zones with cross-account VPC associations in every region and environment. Traced a missing AWS Cloud WAN return route and raised the fix with the networking team.
- Found that SageMaker async inference cannot pull from a private registry and designed a dual-registry path: CI copies images from Harbor to ECR under OIDC; async endpoints pull from ECR, real-time endpoints from Harbor. Rejected an ECR pull-through cache, which cannot use a self-hosted Harbor as its upstream.
- Hardened model deploys for eight target regions: moved a racing version-file write into one post-deploy job, bounded endpoint-creation time, and added autoscaling with zero-scaled standby replicas.
- Built the SLO alerting chain for the data platform, from a SQL-warehouse detector through a Lambda webhook and an SNS FIFO topic (ordered breach and resolve events) to PagerDuty; exported Databricks and SageMaker metrics to Grafana through CloudWatch, with per-team cost attribution.
- Wrote the approved multi-region disaster-recovery design for ML serving: one-hour RTO target, RPO equal to the latest release, zero-scaled standby replicas, release checks and a runbook.
- Audited the account IP allowlist against 3.19M audit-log events (37 entries to 28, each with a named owner); remediated a critical Linux kernel CVE across GCP VM fleets through Terraform; replaced a personal token with GitHub App authentication for Atlantis, key held in Vault; closed 80+ Dependabot alerts across two ML monorepos.
- Migrated a legacy Kubeflow estate (51 pipelines, about 22 unique images and packages) from GKE to Databricks on AWS: scored every asset, mapped Kubeflow Pipelines to Databricks Jobs and Asset Bundles, moved 1,500+ MLflow models to Unity Catalog aliases, retired about 25% instead of moving it.
- Ran the move in scored waves, starting with pilots I built as Asset Bundles, and collapsed 75 per-city Kubeflow jobs into one parameterised job.

Multi-tenant e-commerce platform

2026

- Led the platform’s DevOps team: planned the work as tasks and ran the team meetings.

- Authored onboarding changes for three retailers across 24 service repositories (SSO, retailer and banner settings, fulfilment, carriers and charges, S3-delivered configuration, BigQuery and Terraform, themes, CI paths); all three reached production.
- Spotted late repository-access failures as a repeat launch blocker and built the module that finds them first (~3,400 lines: Python, React/TypeScript, GitLab, Slack, scheduled CI): audits 79 repositories, routes gaps to 13 owners, defaults to dry run with post caps and 24-hour dedupe; owners approve every grant.
- Built an idempotent workflow step that generates tenant CI deploy-wave jobs, validates YAML and opens merge requests; it edits only the target blocks, and pytest proves a rerun changes nothing.
- Planned and ran tenant migrations between platform environments: one cut that tenant's running cost by about 80%; another carried 64 feature flags across 4 banners, each read back for parity.

Stock exchange trading platform 2024 – 2026

- Managed Amazon EKS clusters as code with Terragrunt and Terraform, and delivered workloads onto them through Jenkins shared-library pipelines and Helm; wrote Liquibase database delivery jobs; S3, CloudFront and ACM hosting; Lambda, API Gateway and Step Functions workflows.
- CloudWatch alarms through SNS to PagerDuty, least-privilege IAM and security groups; Level 3 production support and on-call for the trading platform.

Reinsurance platform on Azure 2022 – 2023

- Designed the Azure platform in Terraform (AKS, Data Lake, Azure Database for PostgreSQL, Azure Cache for Redis, Storage) and moved hundreds of Java service pipelines from Jenkins to reusable Azure DevOps templates under regulated change management; built cleanup that removed terabytes of obsolete packages from JFrog Artifactory.

Internal engineering platform, DataArt 2022 – 2023

- Migrated and maintained DataArt's internal platform, where each team describes its services in one descriptor: built its Kubernetes cluster with Prometheus and Grafana monitoring, reachable only over private VDI, upgraded the Python engine behind it, and held on-call.
- With a front-end engineer, built a per-team dashboard of resources and logs, plus an editor and publisher for descriptors; moved hosting from VMs to App Service slots and SQL to Azure SQL.

Crypto exchange and database reliability 2022 – 2023

- Built a highly available PostgreSQL cluster alone, on a short deadline, for a client that could not use a managed database: Patroni, HAProxy and Keepalived, Ansible provisioning, backups, access controls, and Prometheus/Alertmanager alerts on replication lag and failover.
- Automated templated deployments with Ansible and Jinja2, ran identity and access through Keycloak, designed VPN DNS on PowerDNS.

Independent security review Oct – Nov 2026

Solo, read-only pre-launch review of a healthcare application: findings with severity, file-level evidence and fixes, written for the team's coding agent; re-audit two weeks later.

DevOps Engineer, freelance Oct 2021 – Feb 2022

Content-creator platform: automated WordPress deployments to GKE with Datadog uptime checks; folded duplicated GitLab CI pipelines into templates; added role-based authorization to Jenkins; moved data workloads to Amazon RDS and VM services to Cloud Run.

React Developer (part-time), Stone Valley LLC Apr – Sep 2021

PROJECTS

Conatus (conatus.day) 2026

Built and shipped an offline-first journaling app to the App Store: Next.js static export inside Capacitor, every entry stored on the device, no backend.

EDUCATION AND TRAINING

Bachelor's degree, Information Technology, National Polytechnic University of Armenia 2017 – 2023

DevOps bootcamp (CI/CD, containers, infrastructure as code, observability), Nana DevOps Bootcamp Jul – Sep 2021

LANGUAGES

Armenian (native), English (B2), Russian