

Andy Wu

(925) 683-1390 | andyw98@berkeley.edu | [LinkedIn](#) | [GitHub](#)

Summary

Seasoned software engineer specializing in cloud-native platforms and infrastructure automation. Led operation of a multi-team OKE platform across three clusters, 100 nodes and 50 applications serving 1,000 users, and built a schema-driven Terraform generation framework covering 967 OCI resource types. Developed and deployed the orchestration layer for an internal sales RAG assistant handling ~5,000 daily queries, earning Oracle's Platinum Award. Aiming to apply expertise in Kubernetes, Terraform, and AI-driven services to deliver scalable, secure cloud solutions.

EDUCATION

University of California, Berkeley

Bachelor of Science, Computer Science

• GPA: 3.71

May 2024

Berkeley, CA

SKILLS

Cloud & Infrastructure: Kubernetes, Terraform, Helm, Docker, AWS, OCI, Argo CD, Jenkins, Linux

Networking & Observability: CNI, NetworkPolicy, Ingress, IAM, Prometheus, Grafana, OpenTelemetry

Backend & Distributed Systems: Go, Java, Spring Boot, PostgreSQL, Kafka, Redis, OpenSearch

AI: RAG, MCP, LangGraph

EXPERIENCE

Oracle

Sep 2024 - Present

Software Engineer, OCI

Redwood Shores, CA

- Operated a multi-team OKE platform spanning 3 clusters, 1000 nodes, and 50 applications serving 1,000 users, covering cluster upgrades, scaling, RBAC, service discovery, ingress, NetworkPolicy, CNI networking, and Prometheus/Grafana observability.
- Implemented OCI workload identity by upgrading OKE clusters and replacing broad node-level tenancy permissions with workload-scoped IAM, reducing credential blast radius and enforcing least privilege.
- Built a schema-driven Terraform-generation framework covering 967 OCI resource types, grounding LLM-generated modules in provider schemas and validating dependencies, naming conventions, and tags, which accelerated infrastructure provisioning and reduced configuration errors
- Built and deployed the orchestration and serving layer for an internal sales RAG assistant on OKE, including a PDF ingestion pipeline and vector retrieval to serve ~5,000 daily queries, earning Oracle's Platinum Award.
- Migrated legacy WebLogic applications to Spring Boot microservices on Kubernetes; migrated Elasticsearch workloads to OpenSearch and built ETL pipelines to index and query 200K+ sales records.
- Built a multithreaded Spring Boot search service with bounded concurrency and asynchronous OpenSearch/Elasticsearch queries across 5.6M leads, 5M company profiles, and 20K products, delivering sub-200ms parameterized search.

PROJECTS

KubeVista github.com/andyw98w/Kubernetes-Dashboard

Aug 2026 - Present

- Architected a production-grade Go/React Kubernetes operations platform on AWS EKS with a three-AZ Terraform network, private managed nodes, IAM Identity Center, Pod Identity, encrypted storage, and Cognito-protected HTTPS delivery through Route 53, ACM, and ALB.
- Operated 12 application and platform services through Argo CD and Helm, integrating Prometheus, Grafana, Loki, Tempo, OpenTelemetry, External Secrets, ExternalDNS, and AWS controllers with autoscaling, disruption budgets, default-deny networking, and least-privilege RBAC.
- Built a keyless GitHub OIDC-to-ECR delivery pipeline with signed immutable images, SBOM/SLSA attestations, and vulnerability scanning; demonstrated 100% request success during controlled failures.

Illuma

Jun 2026 - Present

- Founded and engineered an AI admissions platform, building a multi-agent orchestration system that coordinates 39 typed tools with shared state and deterministic validation across 79K applicant outcomes.
- Trained and calibrated an XGBoost admissions-probability model across 130 universities using applicant, admissions-plan, and school-level features; achieved a 0.108 Brier score on held-out validation.
- Grew the product to 5,000+ users and \$10K+ revenue through a TikTok and Instagram content strategy, generating 4M+ views.