

ANDREW KAISERAUER

ajkaiserauer@gmail.com | Detroit, MI | [linkedin.com/in/aj-kaiserauer](https://www.linkedin.com/in/aj-kaiserauer) | github.com/kornsour

Engineering and platform leader who builds the systems other engineers build on. Leads a 33-person Developer Experience / Platform Engineering organization, with hands-on depth in Kubernetes, cloud architecture, and AI engineering.

EXPERIENCE

Manager, Platform Engineering - US Infrastructure

Jan 2021 - Present

Deloitte US Member Firm | Detroit, MI

- Built the Platform Engineering organization from the ground up, designing a four-squad structure with dedicated tech leads and BAs. Coach and develop leads across backend, frontend, infrastructure, and platform engineering disciplines, delivering developer tooling and platform products to thousands of developers across the Deloitte US Member Firm
- Designed and built an internal platform management product that evolved from SCIM-based user provisioning into a full product suite: platform health monitoring, adoption analytics, cross-platform access intelligence for security and compliance, and a policy enforcement engine automating governance where vendor platforms lack enterprise controls
- Architected a multi-tenant EKS platform serving 40+ microservices at ~99.95% uptime; standardized CI/CD and container delivery across 15 teams using golden images and GitOps, reducing deployment time from hours to minutes and cutting incidents by ~30%
- Driving GitHub Copilot adoption across the organization, demonstrating AI-native workflows for automated code review on pull requests and agentic issue resolution. Leading enterprise AI platform evaluation for ChatGPT Enterprise, Gemini Enterprise, and Anthropic Enterprise
- Leading org-wide adoption of Backstage as a centralized developer portal for 50+ services, reducing onboarding from weeks to days through a hub-and-spoke model that balances org-level autonomy with firm-wide standardization
- Manage platform strategy across GitHub, Postman, Docker, LucidChart, and more by building self-service tooling, standardized support lanes, and automation that reduces operational burden by 60%

Sr Infrastructure Architect

Jan 2018 - Dec 2020

GE Aviation | Van Buren Township, MI

- Delivered 25+ cloud migrations and new builds resulting in ~\$1.9M in infrastructure cost savings across GE Aviation, Renewables, and Transportation divisions
- Built self-service AWS CloudWatch dashboards for 10+ application teams and designed AWS-native CI/CD reducing deployment time from 1 hour to 2 minutes

SELECTED PROJECTS

- **inference-platform:** Self-directed platform for LLM inference serving: GPU scheduling, high-throughput serving (vLLM, KServe), and inference-aware autoscaling (KEDA) on Kubernetes with cost and SLO discipline (Python, Kubernetes, Envoy AI Gateway)
- **llm-inference-performance:** End-to-end inference performance engineering: device-aware benchmark harness (p50/p95 latency, tokens/sec), measured optimizations across KV-cache, batching, and int8 quantization, tensor-parallel serving, and a custom fused-RMSNorm CUDA kernel (Python, PyTorch, CUDA)
- **ops-triage-agent:** Enterprise agentic system that triages an IT/Ops queue end to end with grounded responses, guarded actions behind human-approval gates, and evaluation-gated releases (Python, FastAPI, React/TypeScript, MCP)
- **deCuisine:** Full meal-planning and grocery SaaS live at decuisine.com: recipe discovery, weekly planner, auto-generated grocery lists, and subscription billing (Next.js, Prisma, Postgres, Stripe)

SKILLS

Management & Product: Engineering team leadership (33 reports), cross-functional stakeholder engagement, product roadmap ownership, internal product development, platform adoption strategy, enterprise tool evaluation, Agile (Scrum, Kanban, SAFe)

AI & LLM: GitHub Copilot (enterprise adoption and enablement), Claude Code, Claude API integration, AI workflow design, agentic systems and MCP, LLM inference serving and performance, enterprise AI platform evaluation

Engineering: Python, FastAPI, TypeScript, React, Next.js, AWS, Azure, GCP, Terraform, Kubernetes (EKS), Docker, CI/CD, GitHub Actions/Apps, Helm, Backstage, Observability (Prometheus, Loki), GitOps (ArgoCD, Kustomize), Linux, Bash

EDUCATION

Michigan State University | BS, Media & Information

May 2016

Minors: Computer Science Engineering, Information Technology, Game Design & Development