Patrick Ellis

Passionate problem-solver and SRE specializing in building reliable large-scale ML systems at Google.

(+44) 7949 380 594
patrick.ellis.w@gmail.com
linkedin.com/patrickellis
github.com/patrick-ellis

EDUCATION

University of Sheffield   Sheffield, UK
Master of Science (M.S.) in Computer Science   Sep 2017 — Jul 2021

- Research area: Designed and validated a 3D convolutional neural network using Python, TensorFlow, and Keras for the automated segmentation of brain tumors in MRI scans, achieving a Dice similarity coefficient of 0.88 on a novel dataset.

EXPERIENCE

Google / London, UK   Apr 2025 — Present
Site Reliability Engineer — SRE-SWE

- Architected and maintained the reliability of large-scale ML training infrastructure, supporting distributed training for models with over 100 billion parameters across thousands of TPUs.
- Reduced model training pipeline failures by 40% through proactive monitoring of data ingestion, implementation of automated data validation checks, and chaos engineering experiments.
- Developed a custom monitoring solution using Prometheus and Borgmon to track GPU/TPU utilization and ML-specific metrics, leading to a 25% improvement in hardware resource efficiency.
- Led the incident response for a critical P0 outage in a production inference service, restoring functionality 50% faster than projected and authoring a post-mortem that led to three core reliability improvements.

Sky / London, UK   Oct 2022 — Mar 2025
Site Reliability Engineer

- Managed intercontinental-scale CDN infrastructure (Akamai / Fastly) using Infrastructure as Code (Terraform), achieving 99.99% availability for critical video streaming services.
- Led a company-wide initiative to manage DNS-as-code using OctoDNS, reducing DNS-related incidents by 55% YoY and onboarding several hundred critical zones to the automated workflow.
- Mentored two junior team members, providing guidance on SRE principles and support for their successful promotion to mid-level engineers.

Sky / London, UK   Feb 2022 — Oct 2022
DevOps Engineer

- Built fully automated CI/CD pipelines on Jenkins for containerized applications using Docker and Kubernetes, reducing developer idle time due to build failures by 53%.
- Cut developer diagnostic time by over 200% by implementing microservices for UI and regression test logging, enabling teams to rapidly identify and quarantine flaky tests.

Sky / London, UK   Sep 2021 — Feb 2022
Software Engineer

- Implemented a RESTful backend API and a React frontend to display reliability metrics from millions of live user sessions of the Sky Go app.
- Designed and implemented low-latency C++ algorithms for device driver software on the Sky Glass platform, demonstrating strong practical application of data structures.

PROJECTS

MLOps Pipeline for Real-Time Log Anomaly Detection
Personal Project — github.com/patrick-ellis/log-anomaly-detector   2027

- Designed and built a fault-tolerant MLOps pipeline on GCP using Kubernetes and Kubeflow to train and deploy a Transformer-based model for real-time log anomaly detection.

Contributor, OpenTelemetry Project
Open Source Contribution   2026

- Developed and merged a new OpenTelemetry exporter for Python to track model inference latency and data drift, now a part of the official contrib repository.

PUBLICATIONS

- **Patrick Ellis**. Lessons from Scaling Google's Inference Fleet: An SRE Perspective. *Google Cloud Blog*, 2027.
- Speaker, SREcon Americas 2026. Presented a framework for ensuring data integrity in petabyte-scale ML training pipelines.

AWARDS

- **Winner**, Google Internal AI for Social Good Hackathon   2027
- **Recipient**, Google Peer Bonus Award for Incident Response Leadership   2026
- **Program Committee Member**, Usenix LISA Conference   2026

SKILLS

Cloud & Infra:	AWS, GCP, Kubernetes, Docker, Terraform, Ansible, Networking, Linux Kernel
ML & Data:	Kubeflow, TensorFlow Extended (TFX), PyTorch, Pandas, Scikit-learn
Observability:	Prometheus, Grafana, OpenTelemetry, Elasticsearch, Borgmon
Languages:	Python, Golang, C++, Javascript, HCL, SQL
Certifications:	Google Cloud Professional Cloud Architect, Terraform Certified Associate