David Laing

Dublin, Ireland, IE | david@davidlaing.com | github.com/mrdavidlaing | linkedin.com/in/mrdavidlaing

Summary

Platform engineering leader with deep expertise in scaling distributed systems and AI infrastructure. Passionate about building observable, reliable systems that serve millions while maintaining operational excellence. Track record of designing large-scale API platforms, optimizing cloud infrastructure performance, and pioneering AI agent orchestration for complex systems. Thrives in high-trust, mission-driven environments where empirical approaches and long-term thinking create lasting impact. Brings unique combination of Python expertise, cloud platform mastery, and AI systems experience to help scale Anthropic's model serving infrastructure.

Education

University of Cape Town, Bachelor of Business Science (Honours) in Information Systems — Jan 1997 – Dec 2000

- Distributed Systems
- Algorithm Design
- Information Systems
- Statistics

Experience

Staff Software Engineer - AI Infrastructure Lead, Mechanical Orchard – San Francisco, CA — Sept 2023 – present

- Architected AI agent orchestration platform for infrastructure automation - creating self-healing, observable systems
- Built and scaled cloud infrastructure on GCP and AWS, optimizing for performance and cost efficiency
- Designed Python-based service mesh for AI agent communication and coordination
- Implemented observability patterns for distributed AI systems using structured logging and metrics
- Led migration of legacy systems to modern cloud-native architectures

Senior Staff Reliability Engineer, Shopify Ltd — May 2022 – July 2023

- Scaled platform to handle $3.5M/minute peak traffic through performance optimization and capacity planning
- Designed observability systems for 100PB data infrastructure using Python-based monitoring tools
- Implemented SLI/SLO frameworks ensuring 99.99% availability for critical commerce APIs
- Optimized resource utilization across Kubernetes clusters reducing infrastructure costs by 30%
- Built Python automation tools for incident response and system recovery

Staff 2 Software Engineer - Data Platform, VMware Ltd — May 2020 – Apr 2022

- Designed Python data pipelines processing terabytes of telemetry data using dagster.io
- Built distributed systems for cross-portfolio analytics on GCP infrastructure
- Implemented performance optimization for BigQuery workloads reducing query costs by 40%
- Created observable data platform architecture with comprehensive monitoring and alerting

Director of Engineering - Platform SRE, Pivotal Software Ltd — Dec 2015 – Apr 2020

- Built SRE team scaling Cloud Foundry platform to thousands of production deployments
- Designed large-scale API platform architecture serving 1M+ active users on pivotaltracker.com
- Implemented Kubernetes orchestration patterns for multi-cloud deployments on AWS/GCP
- Collaborated with Google SRE on error budget implementation and observability practices
- Led performance optimization initiatives reducing platform latency by 50%
- Mentored engineering teams on Python-based automation and infrastructure-as-code

Co-founder - AI Systems, Make Time Together Ltd – Dublin, Ireland — Aug 2023 – present

- Architected full-stack AI application using Python backend with TypeScript/React frontend
- Implemented cloud-native deployment patterns on GCP with Kubernetes orchestration

- Designed observable system architecture with comprehensive monitoring and alerting

R&D Team Leader - Cloud Platform, City Index Ltd – London, UK Dec 2008 – Mar 2015

- Architected LogSearch platform handling 2TB daily log ingestion for distributed systems
- Led organizational migration to AWS cloud infrastructure
- Implemented Python-based automation for infrastructure provisioning and deployment
- Established engineering best practices including TDD and continuous delivery

Publications

Scaling AI Infrastructure: Lessons from Production Jan 2024

David Laing

Internal Publication - Best practices for scaling AI workloads in production environments, focusing on observability, performance optimization, and resource management

SRE & the Vulnerability Budget Mar 2019

David Laing

SRECon AMER - Applying SRE principles to balance security, reliability, and feature velocity in large-scale systems

Projects

AI Agent Infrastructure Platform 2023

- Production platform for AI agent orchestration handling complex infrastructure automation tasks. Built with Python, focusing on observability, reliability, and scale.
- Designed distributed system architecture for AI agent coordination
- Implemented Python-based service mesh for agent communication
- Built comprehensive observability with metrics, logging, and tracing
- Achieved sub-second response times for infrastructure provisioning
- Scaled to handle thousands of concurrent AI agent operations

Technologies

Python & Backend Development: Python, FastAPI, Async Programming, Data Pipelines, Performance Optimization, Testing Frameworks

Cloud Platforms & Infrastructure: AWS Services, GCP Platform, Kubernetes, Docker, Terraform, Infrastructure as Code

Distributed Systems & Scale: API Platform Design, Microservices Architecture, Service Mesh, Load Balancing, Caching Strategies, Performance Tuning

AI/ML Infrastructure: LLM Deployment, AI Agent Orchestration, Model Serving, GPU/Accelerator Optimization, ML Pipeline Design

Observability & Reliability: SLI/SLO Design, Monitoring Systems, Distributed Tracing, Incident Response, Error Budgets, Chaos Engineering

Technical Leadership: System Architecture, Technical Mentorship, Cross-team Collaboration, Documentation, Code Review, Best Practices