

Ruan du Toit

Senior DevOps and Platform Engineer · AI infrastructure and agentic operations

Ely, Cambridgeshire, UK · +44 7506 679669 · hello@ruandutoit.online · ruandutoit.online · linkedin.com/in/ruandutoit-online

Right to work: UK Indefinite Leave to Remain. No visa sponsorship required.

PROFILE

Senior DevOps engineer with 13 years running production infrastructure for an ISP, e-commerce clients and UN public data platforms. I lead the DevOps team at UNEP-WCMC and was previously Staff Engineer and Platform Team Manager at Agile Analog.

Strongest in Terraform, Ansible, CI/CD and Linux across AWS, Azure and Proxmox. I write production tooling in Go and Python and have built with LLMs since 2023. I use AI coding agents daily and test their output before it ships. Where data is sensitive, I run self-hosted models.

EXPERIENCE

Senior DevOps Engineer and Team Lead

UNEP-WCMC, Cambridge | Jan 2024 – present

Lead a DevOps team of two at the UN Environment Programme's biodiversity centre, owning almost 60 cloud hosts across AWS, Azure and Linode, a 5-node Proxmox cluster, and the platforms behind UN Biodiversity Lab (UNBL), Protected Planet and Species+.

- Re-platformed the public GIS service from Windows Server to Ubuntu on AWS, taking ArcGIS Enterprise and PostgreSQL up several major versions while keeping every public URL working.
- Rebuilt UNBL, the UN's spatial-data platform for national biodiversity planning, on Azure with Terraform and Bicep: staging and production, auto-scaling, MongoDB backup and restore, and cost controls.
- Mentored a junior engineer, ran the estate alone between hires and onboarded a new engineer. Created the team's hiring test, one-command developer setup and auditable zero-trust access (Teleport, NetBird).
- Introduced fleet-wide Ansible automation with AWX on managed Kubernetes (Linode LKE) and Vault-backed secrets; the first complete server inventory uncovered servers nobody was tracking.
- Stood up a VPC-isolated Species+ stack alongside live production with no downtime, and a national environmental information system on Azure, both in Terraform.
- Wrote the org-wide GitHub Actions library for Kamal 2 deployments, used by 8+ production projects and hardened against CI injection, running on two redundant sets of four self-hosted runners on Kubernetes with custom Ruby and Kamal images.
- Upgraded the Proxmox cluster and self-hosted GitLab (with PostgreSQL) across major versions with no data loss.
- Developed an in-house operations console (Python and Django, later rewritten in Go and Vue) covering fleet logs, web analytics, scheduled Ansible runs, backups and cloud billing, with Slack integration.
- Delivered self-service VMs on Proxmox: staff sign in with Active Directory and get virtual machines on demand, with Slack integration.
- Led the response to 15+ production and security incidents (including a botnet attack), introduced written incident reports, and rebuilt monitoring with Zabbix (replacing Nagios), Prometheus, Grafana and Graylog.
- Cut idle AWS spend, led a UK-GDPR data-disposal audit of a legacy account holding ~444 TB, and wrote a read-only audit tool covering five AWS accounts.
- Set up an on-premises LLM platform on pooled GPUs and a RAG prototype over UNBL's dataset catalogue, keeping all data on UNEP-WCMC hardware.
- Deployed OpenUniverse, my own AI operations platform, with the organisation's permission: its agents run on GLM-4.7 and investigate ops tickets and incidents.

Staff Engineer (DevSecOps) and Platform Team Manager

Agile Analog, Cambridge | Sep 2021 – Jun 2023

Joined as Platform Engineer in 2021, promoted to Senior DevSecOps Engineer in 2022 and to Staff Engineer and Platform Team Manager in 2023. Left in June 2023 when most of the Platform team was made redundant.

- Line-managed the Platform team and ran its planning, Jira projects and internal helpdesk.
- Built isolated AWS project environments with Terraform, VPC peering and FortiGate access, monitored through Security Hub and CloudTrail alerts.
- Ran configuration and secrets management with Puppet, Foreman and HashiCorp Vault, with Zabbix monitoring.
- Cut idle EC2 spend with a custom AMI that supported hibernation, and introduced Kasm Workspaces for high-security work.

DevOps Engineer

RSAWEB, Cape Town | Sep 2017 – Sep 2021

Cloud Engineer, then DevOps Engineer from 2018, in the Core Network team of an ISP serving hundreds of thousands of customers.

- Ran core network services (FreeRADIUS, PowerDNS with DNSSEC, Sandvine PacketLogic) and penetration-tested internal core services.
- Automated Linux server builds with Chef and Terraform across AWS and VMware, built CI/CD in Jenkins and GitHub Actions, and ran a shared-hosting Docker cluster.
- Operated MySQL clusters behind ProxySQL and Percona, plus MongoDB, with Zabbix and TICK-stack monitoring.

DevOps Engineer

InventCommerce, Cape Town | Nov 2013 – Sep 2017

- Ran hosting infrastructure and architecture planning for e-commerce clients including Pearson, Danone and Edcon.
- Ran Jenkins CI, advised clients on Docker, and load-tested client platforms with JMeter.

PERSONAL PROJECTS

Built in my own time: I set the design and test every change, and AI coding agents write most of the code. Earlier LLM work since May 2023 includes a ChatGPT shell plugin, LLM-generated scripts in GitLab CI and a multi-agent DevOps platform.

OpenUniverse: AI operations platform

Personal project | Feb 2026 – present | private alpha at openuniverse.app

A self-hosted platform where AI agents investigate incidents and tickets, then act on servers through SSH, Ansible and OpenTofu.

- Sole developer for seven months, directing AI coding agents: Go backend, Vue 3 frontend and 1,000+ automated tests.
- Designed a provider-agnostic agent system (Anthropic, OpenAI, OpenRouter or self-hosted models) with an MCP server exposing 82 infrastructure tools to LLM agents, every call logged with secrets redacted.
- Built allow/ask/deny guardrails with human-in-the-loop approval in Slack, and prompt caching that cut input-token cost by ~90%.

Hostirr: compute marketplace

Personal project | Jul 2026 – present | invite-only alpha at hostirr.ai

A marketplace for renting out idle machines as sandboxed Linux environments and app hosting, billed by the second.

- Built in Go and Nuxt since July 2026, with gVisor and Kata microVM sandboxes and gRPC over mTLS to host agents.
- Built an LLM deploy pipeline that reads a repo, writes its Dockerfile and fixes failed builds automatically.

SKILLS

Infrastructure as code: Terraform, Ansible, AWX, Bicep, cloud-init, Puppet, Chef

Containers and virtualisation: Docker, Kubernetes (Linode LKE, Kustomize, cert-manager, kubectrl), Proxmox VE, VMware

CI/CD: GitHub Actions (reusable workflows, self-hosted runners), GitLab CI, Kamal, Jenkins

Cloud: AWS (EC2, S3, VPC, IAM, RDS, Route 53, CloudTrail, Security Hub), Azure, Linode

AI engineering: LLM agents, human-in-the-loop guardrails, MCP, multi-agent workflows, Claude Code, Anthropic and OpenAI APIs, RAG, self-hosted LLM inference (llama.cpp on GPUs), prompt caching

Observability: Zabbix, Prometheus, Grafana, Graylog

Security and identity: HashiCorp Vault, Teleport, NetBird, Active Directory / LDAP, SAST (gosec), UK-GDPR

Networking and edge: Cloudflare (DNS, Tunnels, Workers, WAF, API), NGINX, Traefik, WireGuard, FreeRADIUS, PowerDNS

Languages: Go, Python, Bash, SQL

Data: PostgreSQL, PostGIS, MySQL, MongoDB

EDUCATION AND TRAINING

Cisco Certified Network Associate (CCNA) · 2005 (lapsed)

IT Engineering, CTI, Cape Town · 2003

Spoken languages: English, Afrikaans · Driving licence