

Felipe dos Santos Miranda

Senior DevOps / Platform / SRE Engineer · MLOps · AIOps · GenAI/LLMOps

Caxambu, Minas Gerais, Brazil · Open to remote, worldwide · felipemiranda@outlook.com · linkedin.com/in/fsmiranda · +55 35 99124-1076

I build and operate cloud-native infrastructure that teams can trust — and lately, the platforms that put machine learning and LLMs into production. Thirteen years across big tech, investment banking, 5G telco and AI products: from moving a global telco’s core billing off the mainframe, to owning ML lifecycles end to end, to running agentic AIOps in production.

ROLE DevOps / Platform / SRE — Senior	EXPERIENCE 13+ years	FOCUS Platform · MLOps · AIOps · GenAI
CLOUDS AWS · Google Cloud · Azure	CURRENTLY System Engineer @ Apple	LOCATION Brazil — remote worldwide
LANGUAGES PT native · EN professional	STATUS Open to opportunities	

WHAT I DO

Platform Engineering

Building the paved road so other teams ship without filing tickets.

Terraform

Kubernetes

Flux

Argo CD

GitLab CI

HeIm

MLOps

Owning the path from a model that works to a model that serves.

SageMaker

PyTorch

YOLOv8

OpenCV

IoT Greengrass

Reliability & Observability

Instrumenting systems before the incident, not during it.

Prometheus

Grafana

OpenTelemetry

AWS X-Ray

Sentry

ELK

AIOps & GenAI

Agentic workflows wired into CI/CD and monitoring — running in production today.

Anthropic Claude SDK

LLMops

RAG

Python

Agentic CI/CD

IMPACT IN PRODUCTION

VODAFONE

The UK’s 999 emergency service, re-engineered for 5G

~~1,000 - 2,000 ms~~ → **<100 ms**

Emergency-call re-routing time

<10 s

AML (Advanced Mobile Location) response

A life-critical, heavily regulated system where a slow re-route is a person waiting for an ambulance. I designed the end-to-end Kubernetes architecture — non-production and production — and containerized the legacy stack, writing every manifest and a Helm chart per microservice.

Kubernetes

HeIm

Argo CD

OpenTelemetry

5G

Regulated

VODAFONE

Self-service VM provisioning for the whole engineering org

~~2 weeks~~ → **instant**

Time to provision, after approval

Provisioning a virtual machine meant a ticket and roughly two weeks of waiting. I designed and built the platform that removed the queue entirely.

Terraform

Jenkins

cloud-init

Node.js

Self-service

A global telco's billing core, off the mainframe

Mainframe / COBOL → Linux / Java

600+

Core billing platform

Red Hat Linux servers operated on vSphere

The largest billing application of a North American global telco, running on Mainframe and COBOL. I worked the migration to Linux and Java while operating the estate it ran on.

Linux

Java

COBOL

vSphere

Terraform

Chef

HOW I WORK

automate_over_toil = true

If it hurt twice, it becomes a pipeline. Manual work is a bug report about the platform.

infrastructure_as_code = always

Nothing reaches production unless it is versioned, reviewable and reproducible.

observability_first = true

Instrument before the incident. Debugging blind at 3 a.m. is a choice made months earlier.

own_code_and_infra = true

Horizontal teams work. Whoever writes the service should be able to operate it.

blameless_postmortems = true

The failure is in the system, not the person. Otherwise people stop reporting the near-misses.

ai_as_leverage = true

LLMs take repetitive work off the path. They do not replace judgement, and I do not pretend otherwise.

EXPERIENCE

System Engineer @ Apple

Jun 2025 - Present · Cupertino, CA · Remote

Building and operating internal cloud-native platforms at the intersection of DevOps, Cloud and applied Generative AI, on a horizontal team where every engineer owns both code and infrastructure.

- Build and operate the AWS infrastructure powering internal platforms and services; manage IaC with Terraform for reproducible, auditable environments.
- Design and run CI/CD across GitLab CI and GitHub Actions with GitOps-based continuous delivery (Flux) reconciling applications on Kubernetes (EKS).
- Implement monitoring and observability — AWS Managed Prometheus & Grafana, X-Ray, Sentry — to keep business-critical systems reliable and highly available.
- Lead LLM / GenAI adoption across the SDLC, creating and sharing Claude skills, subagents and hooks; act as the team's LLM/GenAI specialist for embedding LLMs into production.
- Pioneer AIOps with agentic workflows on the Anthropic SDK integrated into CI/CD and monitoring, automating alert investigation and proactive issue detection.

STACK

AWS

Kubernetes (EKS)

Terraform

GitOps (Flux)

GitLab CI

GitHub Actions

Prometheus

Grafana

X-Ray

Sentry

Python

Django

Anthropic Claude SDK

DevOps / MLOps Engineer @ CellarEye

Mar 2023 - Jun 2025 · California, United States · Remote

Owned the path from model to production on a Computer Vision team building an AI-powered app that digitizes wine-cellar management (900+ bottles per collector).

- Built the end-to-end ML lifecycle — data processing in PyTorch, object detection with YOLOv8, image processing with OpenCV — for real-time bottle recognition and cellar mapping.
- Implemented MLOps on AWS SageMaker (training, tuning, deployment) and AWS IoT Greengrass for edge / on-device inference.
- Designed CI/CD and managed Kubernetes on AWS for backend and iOS delivery; ran the infrastructure for scalability, security, high availability and disaster recovery.

STACK

Python

PyTorch

YOLOv8

OpenCV

AWS SageMaker

IoT Greengrass

Kubernetes

CI/CD

MLOps

Computer Vision

Edge

iOS

DevOps Specialist @ Vodafone

Apr 2021 - Mar 2023 · Portugal

Migrated the UK’s 999 emergency-call service to 5G — re-engineering and containerizing a life-critical, highly regulated system on Kubernetes.

- Designed the end-to-end Kubernetes architecture (non-prod + production) for high availability and fault tolerance; containerized legacy apps with all manifests and a Helm chart per microservice.
- Built the CI/CD pipelines and delivery toolchain from scratch (Jenkins, Nexus, GitLab, Argo CD / GitOps); implemented observability and distributed tracing (Prometheus, Grafana, Alertmanager, OpenTelemetry).
- Designed and built a self-service VM provisioning platform — Node.js/Angular portal on top of Terraform, Jenkins, cloud-init and Python.

IMPACT Cut emergency-call re-routing from ~1,000–2,000 ms to under 100 ms, brought AML response under 10 seconds, and reduced VM provisioning from ~2 weeks to instant after approval.

STACK

KubernetesHelmTerraformJenkinsGitLabArgo CDcloud-initPythonPrometheusGrafanaOpenTelemetryElasticsearchRedisvSphere

DevOps Engineer @ TempoTem

Jun 2020 - Apr 2021 · Barueri, São Paulo, Brazil

Built and operated high-availability, 100% AWS cloud-native e-commerce environments.

STACKAWSEKSCloudFormationJenkinsBitbucketPrometheusGrafanaMagento

DevOps Engineer @ Veloe

Feb 2019 - Jun 2020 · Barueri, São Paulo, Brazil

Operated a 100% AWS production environment for Java, Node.js, React and Python applications.

STACKAWSEKSDockerAzure DevOpsCloudFormationDynatraceCloudFront

DevOps Engineer @ BTG Pactual

May 2018 - Feb 2019 · São Paulo, Brazil

Introduced DevOps culture and built architectures for high-availability internal systems at one of Latin America's largest investment banks; 100% AWS cloud-native.

STACKAWSEKSCloudFormationJenkinsGitHub EnterprisePrometheusGrafana

DevOps Engineer @ IBM

Aug 2013 - Jul 2017 · Hortolândia, São Paulo, Brazil

Operated 600+ Red Hat Linux servers on vSphere (production and development) plus Mainframe MVS for a large global telco client.

STACKRed Hat LinuxvSphereMainframe MVSJenkinsDockerTerraformChefJavaCOBOL

System Administrator @ Earlier experience

2008 - 2013 · Brazil / Portugal

Manzano Pneumáticos, Concrete Solutions, Novo Motor and Credimais — Linux server administration, networking and firewall (Squid proxy, VPN, Active Directory), Apache/Magento performance tuning, and web-portal deployments on AWS EC2.

STACKLinuxApacheSquidVPNActive DirectoryAWS EC2

SKILLS

CLOUD	AWS, EKS, EC2, RDS, MSK, OpenSearch, ElastiCache, SageMaker, IoT Greengrass, CloudWatch, X-Ray, Google Cloud, Azure
CONTAINERS & ORCHESTRATION	Kubernetes, Helm, Docker, OpenShift
INFRASTRUCTURE AS CODE	Terraform, Ansible, CloudFormation, cloud-init
CI/CD & GITOPS	GitLab CI, GitHub Actions, Jenkins, Argo CD, Flux
OBSERVABILITY	Prometheus, Grafana, Alertmanager, OpenTelemetry, AWS X-Ray, Sentry, ELK / Elasticsearch
MLOPS / AIOPS / GENAI	AWS SageMaker, PyTorch, Model deployment, Agentic workflows (Anthropic Claude SDK), LLMops, RAG-oriented design
PROGRAMMING	Python, Bash, JavaScript / Node.js

DATABASES

PostgreSQL / RDS

DynamoDB

Redis

Elasticsearch

MongoDB

PRACTICES

SRE

High availability

Disaster recovery

Incident response

Cost optimization

CERTIFICATIONS

- AI Fluency for Builders — Anthropic
 - Teaching the AI Fluency Framework
 - Introduction to Claude Cowork
- AI Fluency for Students — Anthropic
 - AI Capabilities and Limitations
 - Big Data — Fundamentals

15 certifications in total — full list available on LinkedIn.

EDUCATION

Computer Science

IBTA — Instituto Brasileiro de Tecnologia Avançada

2009

LANGUAGES

Portuguese

Native

English

Professional working

Spanish

Elementary

Available for senior platform, SRE and MLOps roles.