

MLOps & AI Infrastructure

Model lifecycle management and AI infrastructure—reliable, observable, scalable

Overview

We deliver MLOps and AI infrastructure: model lifecycle management, monitoring, drift detection, CI/CD for ML, inference pipelines—in cloud and hybrid environments with cost control and scalability.

Business problems we address

- Models work in dev but fail or degrade in production—no model lifecycle management or drift detection.
- No CI/CD for ML; deploy, version, and roll back are manual or inconsistent across teams.
- Inference pipelines and scaling for peak demand cause outages or cost overruns; cost control is missing.
- Monitoring and drift detection are absent—no visibility into latency, throughput, or model degradation.
- Cloud vs on-prem is unclear; hybrid environments lack a single, scalable, governable platform.

How we deliver

We review your current platform and workloads (including cost and cloud vs hybrid requirements), then design and build MLOps and AI infrastructure: model lifecycle management, CI/CD for ML, inference pipelines, monitoring and drift detection, and cost control and scalability. We deploy in your environment (cloud or hybrid) and hand over runbooks and ownership so your teams can operate and iterate with confidence.

Delivery steps

1. Platform & workload review

Current infra, bottlenecks, cost; cloud vs hybrid requirements.

2. Design & build MLOps

Model lifecycle, CI/CD for ML, inference pipelines, monitoring, drift detection.

3. Launch & optimize

Go-live; cost control, scalability tuning, runbooks and ownership.

Security & governance

Infrastructure is built with security and governance in mind: identity and access control for models and pipelines, secrets management, network isolation, and audit logging. We align to your compliance requirements (e.g. SOC2, HIPAA, residency) and document the control environment so security and audit can validate.

Who should use this service

Ideal for VP Engineering, ML platform leads, and DevOps/MLOps teams who need model lifecycle management, CI/CD for ML, inference pipelines, and monitoring with drift detection—in cloud or hybrid environments with cost control and scalability. We serve enterprises that run or plan to run multiple models and need a single, secure, scalable foundation.

Example use case

A retail client needed 10x scale for Black Friday with cost control. We implemented model lifecycle management, CI/CD for ML, inference pipelines with auto-scaling, and monitoring with drift detection—in a hybrid cloud setup. Zero downtime, 3x throughput, and predictable cost during peak.

Key capabilities

- Model lifecycle management
- Monitoring and drift detection
- CI/CD for ML
- Inference pipelines
- Cloud and hybrid environments
- Cost control and scalability

Moaisus Global Solutions — Enterprise AI, securely and at scale.