

Building Enterprise-class AI Infrastructure

Building a robust and secure AI infrastructure is crucial for enterprises aiming to leverage the power of artificial intelligence effectively. Run:ai offers a comprehensive platform designed to address the unique challenges faced by IT, DevOps and Platform teams in building an AI platform at scale while ensuring compliance, security, and control.



Compliance, Security, and Control

Provide a platform that enables secure and controlled access to AI resources. With features such as data isolation between users and prevention of unauthorized access to other IDEs and workloads, organizations can easily onboard new users and ensure compliance with stringent security policies. Seamlessly integrate with existing identity providers, such as Single Sign-On (SSO), streamlining access management and enhancing security.



Support multiple teams and align resources with business priorities

Cater to the needs of multiple teams within an organization while aligning resources with business priorities. Platform teams can set policies to enforce resource restrictions or defaults per workload or project, ensuring efficient resource allocation. Advanced scheduling capabilities, quota management, and over-quota support empower teams to optimize resource utilization and enhance productivity.

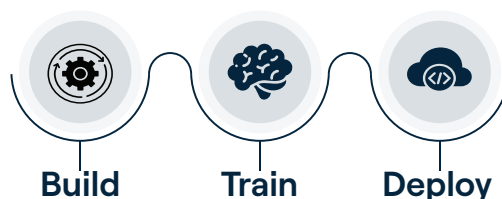


Resource Management at Scale

Resource management becomes automated and optimized. Intelligently allocate resources, ensuring maximum utilization and minimizing manual intervention. Guaranteed quotas can be set on the department or project level, ensuring resource availability and allocation control. Utilize analytics for capacity planning and visibility, to gain valuable insights for data-driven decision-making.

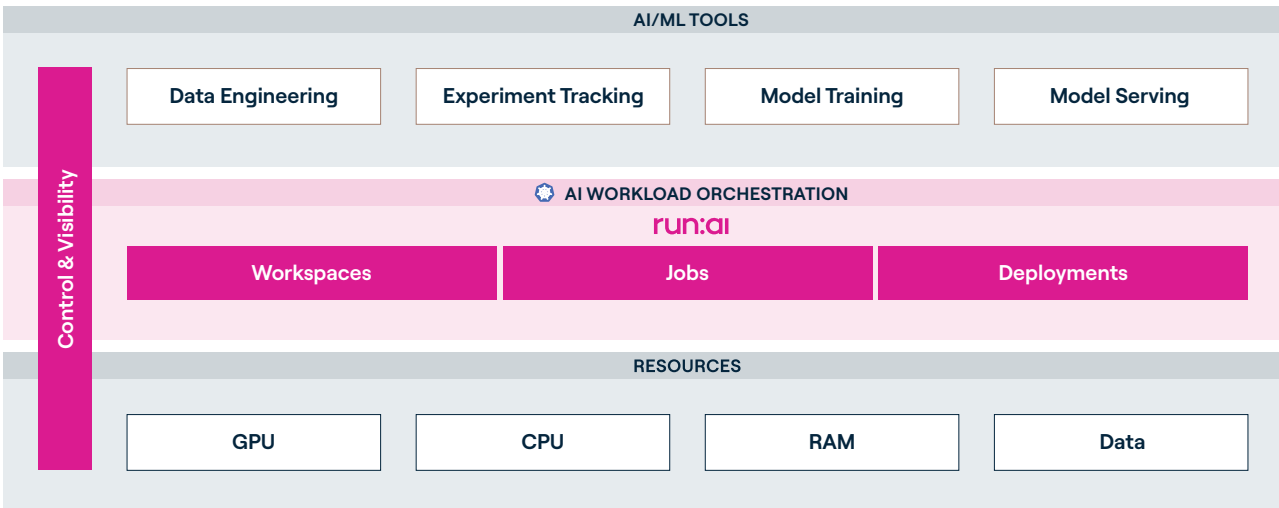
One Platform for the Full AI Lifecycle

ONE PLATFORM



Platform Overview

The Run:ai Atlas platform sits in between the infrastructure and the AI workloads that require access to these valuable resources. Platform teams gain centralized control and visibility across all AI infrastructure, whether on-premises or cloud. AI/ML teams get streamlined and self-service access to all the compute they need, when they need it, using the tools they prefer.



Feature Highlights



Multi-Cluster Dashboard

Provides a holistic, aggregated view across clusters and sites, including information about Cluster and Node utilization, available resources, and allocated resources.



AI Workload Scheduler

Run:ai's K8s Scheduler uses multiple queues to manage batch tasks, with customizable rules and policies for each queue based on business priorities. Combined with over-quota and fairness policies, resource allocation is automated and optimized for maximum cluster utilization.



Workspaces

Workspaces enable data scientists to self-provision the model development tools (like Jupyter Notebook, W&B, MLflow) together with the compute and data they need, in a simplified, streamlined and secured manner.

Customers Accelerating AI with Run:ai Atlas

SONY

BNY MELLON

ZEBRA

XIAOMI