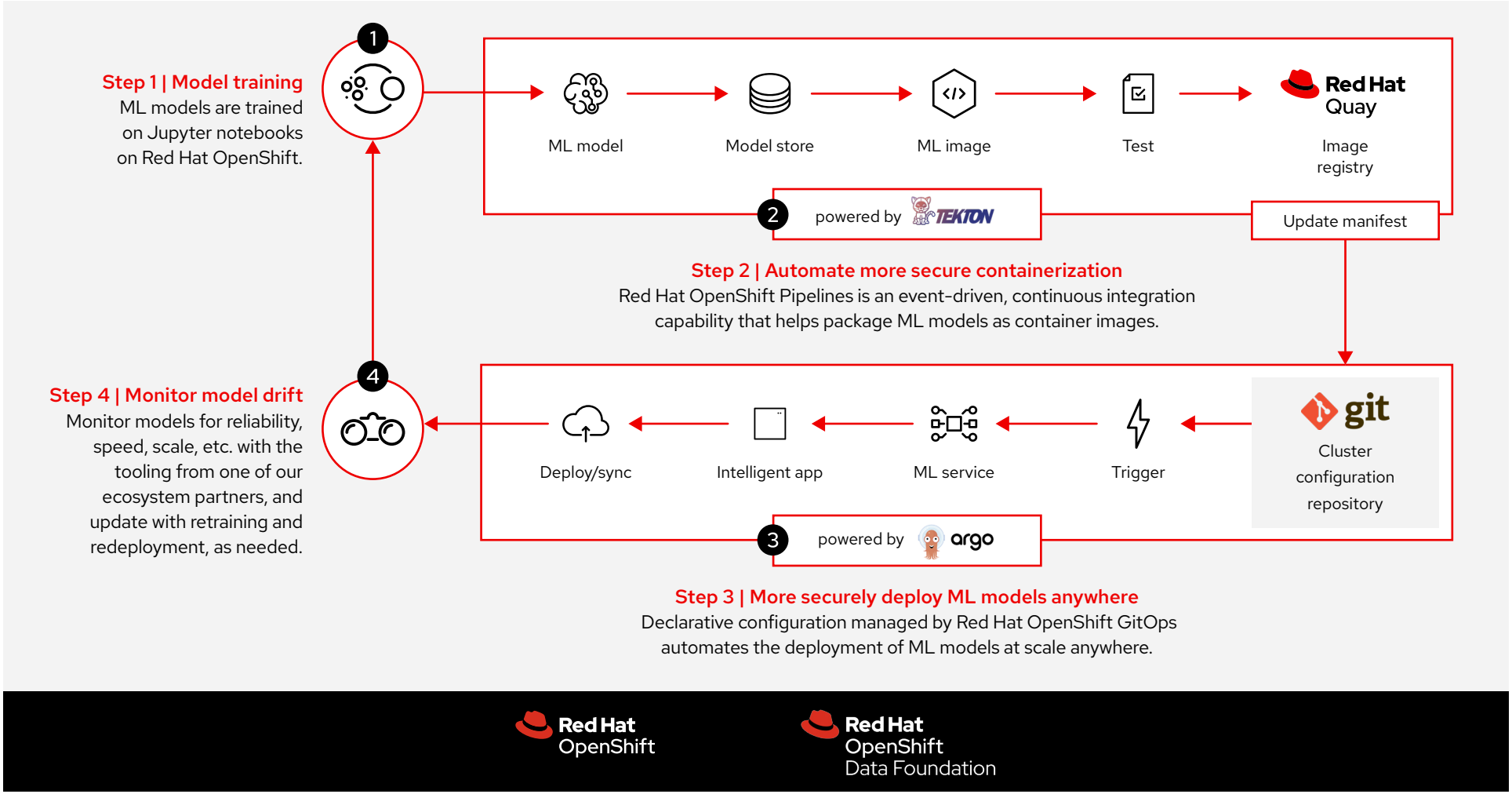
Red Hat OpenShift

MLOps: Machine learning operations with Red Hat OpenShift

Operationalizing machine learning (ML) models for DevOps and ML engineers

There's no single way to build and operationalize ML models, but there is a consistent need to gather and prepare data, develop models, turn models into intelligent applications, and derive revenue from those applications. Adopting MLOps practices means there's no time wasted building or deploying a model and keeping it up to date. Red Hat® OpenShift®, a leading Kubernetes hybrid cloud platform,[1] includes key capabilities to enable MLOps in a consistent way across datacenters, public cloud computing, and edge computing.



Red Hat OpenShift

Red Hat OpenShift Data Foundation

1. Forrester Research. "The Forrester Wave™: Multicloud Container Development Platforms Q3 2020." Sept. 2020.

Red Hat OpenShift

By applying DevOps and GitOps principles, organizations automate and simplify the iterative process of integrating ML models into software development processes, production rollout, monitoring, retraining, and redeployment for continued prediction accuracy.

1 Model training

ML models are trained on Jupyter notebooks on Red Hat OpenShift.

2 Automate containerization with security

Red Hat OpenShift Pipelines, is an event-driven, continuous integration capability that helps package ML models as container images by:

- **Saving** the models ready for deployment in a model store.
- **Converting** the saved models to container images with Red Hat OpenShift build.
- **Testing** the containerized model images to ensure they remain functional.
- **Storing** the containerized model images in a private, global container image registry like Red Hat Quay where the images are analyzed to identify potential issues, mitigating security risks and geo replication.

3 More securely deploy models anywhere

Declarative configuration managed by Red Hat OpenShift GitOps automates the deployment of ML models at scale, anywhere, by:

- **Configuring** Red Hat OpenShift environments for artificial intelligence (AI) inferencing, anywhere, via Git repositories. These configuration requirements are recorded and can be versioned at the source, reducing the propensity for errors and increasing developer and data scientist productivity.
- **Monitoring** the manifest with the latest container image to be used by the intelligent app. Red Hat Quay scans for vulnerabilities in container images for more secure deployment.
- **Triggering** the deployment of containers running the latest version of the models as ML services that are used by intelligent software applications via application programming interfaces (APIs).
- **Deploying** the latest versions of the containerized ML models and the associated intelligent application to Red Hat OpenShift at all the locations where ML inferencing is to be performed (i.e., datacenters, public cloud computing, and edge computing).

4 Monitor model drift

Monitor models for reliability, speed, scale, etc. with the tooling from one of our ecosystem partners, and update with retraining and redeployment, as needed.

Learn more

Webinar:
Accelerate MLOps and deliver intelligent apps ▸

Architecture blueprint:
Industrial manufacturing for edge computing ▸

Get started:
Red Hat OpenShift interactive learning portal ▸