Red Hat
OpenShift
Data Science

Red Hat OpenShift Data Science

A cloud service for rapid AI/ML development, deployment, and scaling

Highlights

Rapidly develop, train, test, and deploy containerized machine learning models without having to design and deploy Kubernetes infrastructure.

Conduct exploratory data science in Jupyter notebooks with access to core AI/ML libraries and frameworks, including TensorFlow and Pytorch.

Collaborate within a common platform to bring IT, data science, and app dev teams together.

Serve models for integration into intelligent applications; rebuild and deploy based on changes to the source notebook.

Open Data Hub is a blueprint for building an AI-as-a-Service (AIaaS) platform on Red Hat OpenShift. It inherits from upstream efforts such as Jupyter and Kubeflow, and is the foundation for Red Hat's internal data science and AI platform.

Accelerating artificial intelligence and machine learning deployments

Artificial intelligence (AI), machine learning (ML), and deep learning (DL) have rapidly become critical for businesses and organizations. According to IDC, "AI is profound and is affecting businesses and organizations across industries. AI is everywhere across the technology stack."[1] Deploying these technologies, however, can be complicated. As data scientists strive to build their models, they often encounter a lack of alignment between rapidly evolving tools, influencing productivity and collaboration among themselves, software developers, and IT operations. Scaling AI/ML deployments can be resource-limited and administratively complex while requiring expensive graphics processing unit (GPU) resources for hardware acceleration. Popular cloud platforms offer scalability and attractive toolsets, but those same tools often lock users in, limiting architectural and deployment choices.

Red Hat® OpenShift® Data Science is an AI platform offering based on the open source Open Data Hub project. Data scientists and developers can rapidly develop, train, test, and iterate ML/DL models with full support, allowing them to focus on their modeling and application development without waiting for infrastructure provisioning. Available as an add-on cloud service to Red Hat OpenShift Dedicated and Red Hat OpenShift Service on AWS or as a self-managed software product, OpenShift Data Science combines Red Hat components, open source software, and technology partner offerings with the flexibility to develop and serve models on-premise or in all three public clouds.

Red Hat OpenShift Data Science

Red Hat OpenShift Data Science offers organizations an efficient way to deploy an integrated set of common open source and third-party tools to perform AI/ML modeling. The platform makes it simpler to exploit hardware acceleration, including central processing unit (CPU) and NVIDIA GPU-supported hardware infrastructure without the need to stand up and perform daily management of Kubernetes on your own.

Red Hat OpenShift Data Science represents an alternative to prescriptive and opinionated AI/ML suites available from individual cloud providers. Adopters gain a collaborative open source toolset and a platform for building experimental models without worrying about the infrastructure or lock-in from public cloud-specific tools. They can then extend that base platform with partner tools to gain increased capability. Models can be served to production environments in a container-ready format, consistently, across hybrid cloud and edge environments.

1 *IDC Market Forecast. "Worldwide Artificial Intelligence Software Forecast, 2022–2026," Doc # US49571222, August 2022.*

Red Hat OpenShift Data Science

Red Hat Consulting offers several AI/ML consulting engagements for challenges like implementing MLOps and intelligent application development.

Red Hat OpenShift Data Science supports rapid model development with user-supplied data where the model outputs are:

- Hosted in the cloud service for testing or integration into a customer-defined intelligent application.
- Exported or deployed to other Red Hat OpenShift locations for integration into a customer-defined intelligent application.

Red Hat OpenShift Data Science provides IT operations with an environment that is easy to manage, with simple configurations on a security-focused and proven platform you can scale up or down with low effort. Capabilities like the ability to deploy custom notebook images to your data scientist help to maintain control while not sacrificing experimentation.

Upstream open source and commercial technology partner tools

Red Hat OpenShift Data Science provides a subset (Table 1) of the tools found in the upstream Open Data Hub project. Organizations can develop, test, and deploy models across any cloud environment, fully managed, and self-managed Red Hat OpenShift and centrally monitor their performance. Red Hat provides regular updates to open source tools (e.g., Jupyter, Pytorch, and Tensorflow), removing integration and testing burden. The offering also integrates several AI/ML technology partner offerings (Table 1). Additional commercial technology partner offerings can also be added from more than 30 AI technology partners who have certified their product on Red Hat OpenShift.

Table 1 Initial Red Hat OpenShift Data Science ecosystem

AI/ML modeling and visualization tools	JupyterLab UI with out-of-the-box notebook images and common Python libraries and packages; TensorFlow; PyTorch, CUDA; Kubeflow notebook controller for managing multiple notebook sessions, Anaconda (Professional is optional); Intel AI Analytics Toolkit, IBM Watson Studio (optional)
Data engineering	Starburst (Galaxy is optional); Pachyderm (optional)
Data ingestion and storage	Red Hat OpenShift Streams for Apache Kafka (optional add-on); Amazon Simple Storage Service (S3)
GPU support	NVIDIA (with GPU operator)
Model serving and monitoring	Model serving (model mesh with user interface), model monitoring, Source-to-Image (OpenShift) Red Hat OpenShift API Management (optional add-on), Intel Distribution of the OpenVINO toolkit

Conclusion

Red Hat OpenShift Data Science delivers common data science tools as the foundation of a hybrid AI and MLOps platform integrated with partner offerings. The platform simplifies the development, training, testing, and deployment of AI/ML models, complete with a shared user interface for navigation, onboarding, and exploring partner options. Organizations can rapidly develop their AI/ML models, expanding further by adding open source tools and Red Hat technology partner solutions.

About Red Hat

Red Hat helps customers standardize across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with award-winning support, training, and consulting services.

facebook.com/redhatinc
@RedHat
linkedin.com/company/red-hat

North America
1 888 REDHAT1
www.redhat.com

Europe, Middle East, and Africa
00800 7334 2835
europe@redhat.com

Asia Pacific
+65 6490 4200
apac@redhat.com

Latin America
+54 11 4329 7300
info-latam@redhat.com

redhat.com
#220003_0223