

**Publication date:**

11 May 2021

**Author:**

Bradley Shimmin, Chief Analyst, AI & Data Analytics

# Red Hat introduces Red Hat OpenShift Data Science with a platform-savvy take on machine learning in the enterprise

Table of Contents



[Omdia view2](#)  
[Appendix5](#)

# Omdia view

---

## Summary

Building on its experience in developing and supporting containerized, platform-agnostic software, Red Hat has introduced Red Hat OpenShift Data Science, an enterprise AI development cloud service that emphasizes security, portability, and scale without constraining the use of disparate data science technologies. Though later to the commercial machine learning operationalization (MLOps) marketplace, Red Hat's distinctive IT- and developer-centric point of view coupled with a strong dedication to open source software and ecosystem partnerships bodes well for the company over the long term. However, with more to learn and accomplish, it may take the vendor some time to disrupt what is already a crowded marketplace.

## Does the market need an IT-savvy ML platform?

In late April 2021, Red Hat introduced Red Hat OpenShift Data Science, a machine learning (ML) workflow platform designed to support the development, training, and testing of ML models, ensuring those models are packaged for export within a container-based format. Available immediately in beta with a planned launch July 2021, Red Hat's new platform can be purchased as an add-on to Red Hat OpenShift managed cloud services running initially on Amazon Web Services (AWS) in two flavors—Red Hat OpenShift Dedicated and Red Hat OpenShift Service on AWS.

Why would a renowned Linux and cloud-native platform player leap into the enterprise ML development marketplace? On the surface, this move seems unusual, given the company's long-standing history of speaking directly to developers and infrastructure engineers. Data science emphasizes experimentation and exploration, something far removed from the highly programmatic and operational worldview of enterprise IT practitioners. Yet, this disconnect represents the very reason Red Hat "should" take on ML development. The chasm between data scientists, developers, and IT professionals represents a fundamental challenge for all enterprise practitioners wishing to build AI outcomes. It may only take an experienced data scientist (and colleagues) a month or two to build a predictive ML model, for example, but with no direct reach into IT operations to assist with testing and deployment, it may take far longer for that final model to reach production—if it does at all.

This chasm and disconnect have created a highly competitive marketplace for MLOps platforms that strives to operationalize the complete ML development lifecycle. Omdia recently reviewed 10 leading players in this field (see the *Omdia Universe: Selecting an Enterprise MLOps Platform, 2021*) and found two important takeaways. First, MLOps vendors are building for the cloud, actually multiple clouds, with cross-cloud platform support for deployed models serving as a key differentiator. Second, enterprise AI practitioners live and breathe open source software. Solutions built on ML are built not on ML platforms but rather with open source tools like MLflow and Seldon, using open source software libraries like TensorFlow and PyTorch. Additionally, they are built to run within containerized endpoints managed via popular open source orchestrators like Kubernetes.

## An open workflow is the key

Red Hat's answer to this challenge revolves around the notion of workflow. Rather than build a solution-complete MLOps platform such as those available from pure play vendors DataRobot, cnvrg.io, Iguazio, etc.,

Red Hat has instead built an ML “workflow” platform that enables users to assemble their MLOps solutions using the tools with which they are most familiar. Users then run that solution on top of the company’s open hybrid cloud platform, Red Hat OpenShift, gaining access to a host of benefits such as ready access to AI acceleration hardware, hybrid and edge deployment options, as well as data governance and security capabilities. Or, they can simply package their ML models for deployment on any platform using Red Hat’s Source-to-Image (S2i) toolkit in Red Hat OpenShift.

True to Red Hat’s roots, the company’s approach with Red Hat OpenShift Data Science revolves around the open source ecosystem. The product is built on top of the Open Data Hub project (a project Red Hat has been supporting for several years), which itself built on top of the open source project Kubeflow and uses several popular open source projects, including:

- **Airflow:** workflow management
- **Kafka:** data streaming
- **Spark:** data processing
- **Superset:** data exploration
- **Argo:** workflows for Kubernetes
- **Grafana:** data visualization
- **JupyterHub:** Jupyter notebook server
- **Prometheus:** monitoring
- **Seldon:** MLOps deployment

These projects play an important role within Red Hat’s implementation of Open Data Hub. However, they are not required, nor are they the only tools that can be employed by users in building their MLOps workflows within the confines of Red Hat OpenShift Data Science. Red Hat intends for its new solution to serve as the pivot point for an ecosystem of both open source and commercial tools. Out of the box, Red Hat OpenShift Data Science starts with JupyterLab and associated frameworks like TensorFlow and PyTorch. It also comes pre-integrated with several tools that are available from an initial set of partners, including:

- **Starburst Galaxy:** data integration across hybrid cloud scenarios
- **Anaconda Commercial Edition:** virtual project environments plus package control and versioning
- **IBM Watson Studio:** a full data science development environment
- **NVIDIA:** direct access to the NVIDIA GPU-enabled hardware
- **Seldon Deploy:** ML model packaging and deployment

Red Hat intends to expand this roster throughout the remainder of this year, opening access to software from technology partners through Red Hat Marketplace. An important aspect of this approach is that Red Hat will be able to certify that all the software accessed or purchased through this marketplace will work on Red Hat OpenShift. Doing so future proofs customer investments in terms of portability, building trust that their software can deploy this software on any platform equipped with OpenShift, whether on-premises or in the cloud. Note also that customers can still deploy from all certified marketplace partners. So, while not integrated into OpenShift Data Science, Red Hat is not limiting customers in terms of the software they can run alongside its new offering.

That is the key to Red Hat OpenShift Data Science. With this offering, Red Hat is not trying to compete head-on with established ML development solutions such as AWS SageMaker Suite. Instead, the company is attempting to build the means for enterprise AI practitioners to use Red Hat's open development and deployment workflow components to build their rendition of SageMaker Suite and to do so not just on top of the AWS platform alone, but anywhere that OpenShift runs.

## Next steps and opportunities

With the Red Hat OpenShift Data Science, Red Hat is seeking to build a broader unified platform upon which developers can build cloud-native applications. In support of this endeavor, Red Hat launched two additional solutions alongside OpenShift Data Science:

- **Red Hat OpenShift Streams for Apache Kafka**, a fully managed cloud service of Apache Kafka for the creation of real-time data streams and app messaging
- **Red Hat OpenShift API Management**, a fully managed application programming interface (API) and API management solution for microservices-based development that is tightly integrated with OpenShift.

Taken together, these three services form a suite of core services for developing cloud-native data, application, and ML services, which can be deployed across a wide range of premises, cloud, and edge configurations. This makes Red Hat look a bit like the multiple public cloud platforms upon which OpenShift runs in terms of providing a full development stack, a similarity that will likely increase as Red Hat builds its Red Hat Marketplace partner ecosystem.

Of course, it will take some time for Red Hat to mature and extend its portfolio of Red Hat cloud services. Red Hat OpenShift Data Science currently only runs as a cloud-born service on AWS. Support for cloud platforms from Google, Microsoft, even IBM is forthcoming. On-premises deployments are also on the company's roadmap for the near future. Outside of initial points of integration with IBM Watson Studio, Red Hat has not yet exploited the numerous data and analytics opportunities currently on offer within IBM's Cloud Pak portfolio, particularly Cloud Pak for Data.

Furthermore, within the product, numerous technological holes await support from Red Hat and its emerging partner ecosystem. For example, the company is intending to build in data versioning, governance, and lineage capabilities soon by integrating Pachyderm (another open source project). Red Hat intends to provide support for advanced functionality like AutoML entirely through integration with established players such as H2O.ai and PerceptiLabs.

Regardless, in introducing OpenShift Data Science, Red Hat does an excellent job of building a unique approach to the problem of closing the chasm between data scientists and IT professionals. The company has built an open source, managed cloud service equipped with a solid set of core MLOps workflow services, spanning the development, training, and deployment of ML models using cloud-native technologies. For potential enterprise buyers looking for a hybrid, multi-cloud platform that favors open source software and cloud-native deployment methodologies, Red Hat OpenShift Data Science represents a compelling means to avoid being locked into a monolithic software stack or tied to a single cloud platform.

# Appendix

---

## Further reading

[Omdia Universe: Selecting an Enterprise MLOps Platform, 2021](#) (April 2021)

## Author

Bradley Shimmin, Chief Analyst, AI Platforms, Analytics and Data Management

[askananalyst@omdia.com](mailto:askananalyst@omdia.com)

## Citation policy

Request external citation and usage of Omdia research and data via [citations@omdia.com](mailto:citations@omdia.com).

## Omdia consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help you. For more information about Omdia's consulting capabilities, please contact us directly at [consulting@omdia.com](mailto:consulting@omdia.com).

## Copyright notice and disclaimer

The Omdia research, data and information referenced herein (the "Omdia Materials") are the copyrighted property of Informa Tech and its subsidiaries or affiliates (together "Informa Tech") or its third party data providers and represent data, research, opinions, or viewpoints published by Informa Tech, and are not representations of fact.

The Omdia Materials reflect information and opinions from the original publication date and not from the date of this document. The information and opinions expressed in the Omdia Materials are subject to change without notice and Informa Tech does not have any duty or responsibility to update the Omdia Materials or this publication as a result.

Omdia Materials are delivered on an "as-is" and "as-available" basis. No representation or warranty, express or implied, is made as to the fairness, accuracy, completeness, or correctness of the information, opinions, and conclusions contained in Omdia Materials.

To the maximum extent permitted by law, Informa Tech and its affiliates, officers, directors, employees, agents, and third party data providers disclaim any liability (including, without limitation, any liability arising from fault or negligence) as to the accuracy or completeness or use of the Omdia Materials. Informa Tech will not, under any circumstance whatsoever, be liable for any trading, investment, commercial, or other decisions based on or made in reliance of the Omdia Materials.

## CONTACT US

[omdia.com](https://www.omdia.com)

[askananalyst@omdia.com](mailto:askananalyst@omdia.com)