



Red Hat OpenShift Data Science Self-managed 1.28

1.28 release notes

Features, Technology Previews, and known issues associated with this release

Red Hat OpenShift Data Science Self-managed 1.28 1.28 release notes

Features, Technology Previews, and known issues associated with this release

Legal Notice

Copyright © 2023 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux[®] is the registered trademark of Linus Torvalds in the United States and other countries.

Java[®] is a registered trademark of Oracle and/or its affiliates.

XFS[®] is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL[®] is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js[®] is an official trademark of Joyent. Red Hat is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack[®] Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Abstract

These release notes provide an overview of new features, enhancements, major technical changes, and any known bugs in version 1.28 of Red Hat OpenShift Data Science.

Table of Contents

CHAPTER 1. PROVIDING FEEDBACK ON RED HAT DOCUMENTATION	4
CHAPTER 2. OVERVIEW OF OPENSIFT DATA SCIENCE	5
CHAPTER 3. PRODUCT FEATURES	6
3.1. FEATURES FOR DATA SCIENTISTS	6
3.2. FEATURES FOR IT OPERATIONS ADMINISTRATORS	6
3.3. ENHANCEMENTS	7
3.4. LIMITED SUPPORT FEATURES	8
CHAPTER 4. BUG FIXES	9
4.1. EXPORTING AN ELYRA PIPELINE EXPOSED S3 STORAGE CREDENTIALS IN PLAIN TEXT	9
4.2. WHEN EDITING THE DETAILS OF A SHARED PROJECT, THE USER INTERFACE REMAINED IN A LOADING STATE WITHOUT REPORTING AN ERROR	9
4.3. DATA SCIENCE PIPELINE GRAPHS DID NOT DISPLAY NODE EDGES FOR RUNNING PIPELINES	9
4.4. NEWLY CREATED DATA CONNECTIONS WERE NOT DETECTED WHEN YOU ATTEMPTED TO CREATE A PIPELINE SERVER	10
4.5. WHEN SHARING A PROJECT WITH ANOTHER USER, THE OPENSIFT DATA SCIENCE USER INTERFACE TEXT WAS MISLEADING	10
4.6. USERS WITH "EDIT" PERMISSION COULD NOT CREATE A MODEL SERVER	10
4.7. OPENVINO MODEL SERVER RUNTIME DID NOT HAVE THE REQUIRED FLAG TO FORCE GPU USAGE	10
4.8. CHANGING THE HOST PROJECT WHEN CREATING A PIPELINE RAN RESULTED IN AN INACCURATE LIST OF AVAILABLE PIPELINES	10
CHAPTER 5. KNOWN ISSUES	11
5.1. DEFAULT SHARED MEMORY FOR JUPYTER NOTEBOOK MIGHT CAUSE A RUNTIME ERROR	11
5.2. DATA SCIENCE DASHBOARD DOES NOT DETECT AN EXISTING OPENSIFT PIPELINES INSTALLATION	12
5.3. ELYRA PIPELINE FAILS TO RUN IF WORKBENCH IS CREATED BY A USER WITH EDIT PERMISSIONS	12
5.4. DEPLOYING A CUSTOM MODEL-SERVING RUNTIME CAN RESULT IN AN ERROR MESSAGE	13
5.5. PIPELINES WITH NON-UNIQUE NAMES DO NOT APPEAR IN THE DATA SCIENCE PROJECT USER INTERFACE	13
5.6. UNINSTALL PROCESS FOR OPENSIFT DATA SCIENCE MIGHT BECOME STUCK WHEN REMOVING KFDEFS RESOURCES	13
5.7. DATA SCIENCE PIPELINE SETUP FAILS WITH AN SSL CERTIFICATION ERROR MESSAGE	14
5.8. RUNNING PIPELINES FROM ELYRA WITHIN JUPYTERLAB - NOT SUPPORTED IN A DISCONNECTED ENVIRONMENT	14
5.9. AFTER UPGRADE, THE DATA SCIENCE PIPELINES TAB IS NOT ENABLED ON THE OPENSIFT DATA SCIENCE DASHBOARD	14
5.10. INCORRECT CRON FORMAT DISPLAYED BY DEFAULT WHEN SCHEDULING A RECURRING PIPELINE RUN	15
5.11. YOU CANNOT CREATE A PIPELINE SERVER WHEN CUMULATIVE CHARACTER LIMIT IS EXCEEDED	15
5.12. A PIPELINE SERVER FAILS TO START UNLESS YOU SPECIFY AN AMAZON WEB SERVICES (AWS) SIMPLE STORAGE SERVICE (S3) BUCKET RESOURCE	15
5.13. ERROR APPEARS WHEN INSTALLING THE RED HAT OPENSIFT DATA SCIENCE OPERATOR TO AN OPENSIFT CLUSTER IN A DISCONNECTED ENVIRONMENT WHEN USING OC-MIRROR VERSION 4.12 OR OLDER	15
5.14. USER WITHOUT DASHBOARD PERMISSIONS IS ABLE TO CONTINUE USING THEIR RUNNING NOTEBOOKS AND WORKBENCHES INDEFINITELY	16
5.15. ATTEMPTING TO INCREASE THE SIZE OF A PERSISTENT VOLUME (PV) FAILS WHEN IT IS NOT CONNECTED TO A WORKBENCH	16
5.16. UNABLE TO SCALE DOWN A WORKBENCH'S GPUS WHEN ALL GPUS IN THE CLUSTER ARE BEING USED	16

5.17. ANACONDA PROFESSIONAL EDITION CANNOT BE VALIDATED AND ENABLED IN OPENSIFT DATA SCIENCE	16
5.18. UNCLEAR ERROR MESSAGE DISPLAYS WHEN USING INVALID CHARACTERS TO CREATE A DATA SCIENCE PROJECT	17
5.19. AN ERROR CAN OCCUR WHEN TRYING TO EDIT A WORKBENCH	17
5.20. AN IMAGEPULLBACKOFF ERROR MESSAGE IS NOT DISPLAYED WHEN REQUIRED DURING THE WORKBENCH CREATION PROCESS	17
5.21. WORKBENCHES FAIL TO START WHEN CUMULATIVE CHARACTER LIMIT IS EXCEEDED.	17
5.22. THE NOTEBOOK CREATION PROCESS FAILS FOR USERS WHO HAVE NEVER LOGGED IN TO THE DASHBOARD	17
5.23. THE MODEL MESH OAUTH-PROXY CONTAINER IS INTERMITTENTLY UNSTABLE	17
5.24. THE NVIDIA GPU OPERATOR IS INCOMPATIBLE WITH OPENSIFT 4.11.12	18
5.25. INCORRECT PACKAGE VERSION DISPLAYED DURING NOTEBOOK SELECTION	18
5.26. WHEN USING THE NVIDIA GPU OPERATOR, MORE NODES THAN NEEDED ARE CREATED BY THE NODE AUTOSCALER	18
5.27. THE APPLICATION LAUNCHER MENU INCORRECTLY DISPLAYS A LINK TO OPENSIFT CLUSTER MANAGER	18
5.28. NOTEBOOK SERVER ADMINISTRATION PAGE SHOWS USERS WHO HAVE LOST PERMISSION ACCESS	18
5.29. GPUS ON NODES WITH UNSUPPORTED TAINTS CANNOT BE ALLOCATED TO NOTEBOOK SERVERS	18
5.30. TENSORBOARD REQUIRES MANUAL STEPS TO VIEW	19
5.31. THE INTEL® ONEAPI AI ANALYTICS TOOLKITS QUICK START REFERENCES NONEXISTENT SAMPLE NOTEBOOKS	19
5.32. THE CRONJOB RESPONSIBLE FOR VALIDATING ANACONDA PROFESSIONAL EDITION'S LICENSE IS SUSPENDED AND DOES NOT RUN DAILY	19
5.33. THE NVIDIA GPU OPERATOR CARD ON THE DASHBOARD DISPLAYS BUTTON UNNECESSARILY	19
5.34. DASHBOARD DOES NOT DISPLAY ENABLED PAGE CONTENT AFTER ISV OPERATOR UNINSTALL	19
5.35. INCORRECT PACKAGE VERSIONS DISPLAYED DURING NOTEBOOK SELECTION	20
5.36. ERROR CAN OCCUR WHEN CREATING A NOTEBOOK INSTANCE	20
5.37. ACTIONS ON DASHBOARD NOT CLEARLY VISIBLE	20
5.38. LICENSE RE-VALIDATION ACTION APPEARS UNNECESSARILY	20
5.39. ERROR CAN OCCUR DURING PACHYDERM DEPLOYMENT	20
5.40. IBM WATSON STUDIO NOT AVAILABLE IN OPENSIFT DATA SCIENCE	20
5.41. UNNECESSARY WARNINGS ABOUT MISSING GRAPHICAL PROCESSING UNITS (GPUS)	20
5.42. USER SESSIONS PERSIST IN SOME COMPONENTS	21

CHAPTER 1. PROVIDING FEEDBACK ON RED HAT DOCUMENTATION

Let Red Hat know how we can make our documentation better. You can provide feedback directly from a documentation page by following the steps below.

1. Make sure that you are logged in to the Customer Portal.
2. Make sure that you are looking at the **Multi-page HTML** format of this document.
3. Highlight the text that you want to provide feedback on. The **Add Feedback** prompt appears.
4. Click **Add Feedback**.
5. Enter your comments in the **Feedback** text box and click **Submit**.



IMPORTANT

Some ad blockers might impede your ability to provide feedback on Red Hat documentation. If you are using a web browser that has an ad blocker enabled and you are unable to leave feedback, consider disabling your ad blocker. For more information about how to disable your ad blocker, see the documentation for your web browser.

Red Hat automatically creates a tracking issue each time you submit feedback. Open the link that is displayed after you click Submit and start watching the issue, or add more comments to give us more information about the problem.

Thank you for taking the time to provide your feedback.

CHAPTER 2. OVERVIEW OF OPENSIFT DATA SCIENCE

Using Red Hat OpenShift Data Science, users can integrate data, artificial intelligence and machine learning software to execute end-to-end machine learning workflows. OpenShift Data Science is supported in two configurations:

1. Installed as an Add-on to a Red Hat managed environment such as Red Hat OpenShift Dedicated and Red Hat OpenShift Service on Amazon Web Services (ROSA).
2. Installed as a self-managed Operator on a self-managed environment, such as Red Hat OpenShift Container Platform.

For data scientists, OpenShift Data Science includes Jupyter and a collection of default notebook images optimized with the tools and libraries required for model development, and the TensorFlow and PyTorch frameworks. Deploy and host your models, integrate models into external applications, and export models to host them in any hybrid cloud environment. You can also accelerate your data science experiments through the use of graphics processing units (GPUs).

For administrators, OpenShift Data Science enables data science workloads in an existing Red Hat OpenShift Dedicated or ROSA environment. Manage users with your existing OpenShift identity provider, and manage the resources available to notebook servers to ensure data scientists have what they require to create, train, and host models.

CHAPTER 3. PRODUCT FEATURES

Red Hat OpenShift Data Science provides several features for data scientists and IT operations administrators.

3.1. FEATURES FOR DATA SCIENTISTS

Model serving

As a data scientist, you can now deploy your trained machine-learning models to serve intelligent applications in production. Deploying or serving a model makes the model's functions available as a service endpoint that can be used for testing or integration into applications.

Work with data science pipelines

OpenShift Data Science now supports data science pipelines. Using Red Hat OpenShift Data Science pipelines, you can standardize and automate machine learning workflows to enable you to develop and deploy your data science models.

One-page Jupyter notebook server configuration

Choose from a default set of notebook images pre-configured with the tools and libraries you need for model development.

Collaborate on notebooks using Git

Use JupyterLab's Git interface to work collaboratively with application developers or add other models to your notebooks.

Deploy using application templates

Red Hat provides application templates designed for data scientists so that you can easily deploy your models and applications for testing purposes on Red Hat OpenShift Container Platform.

Try it out in the Red Hat Developer sandbox environment

You can try out OpenShift Data Science and access tutorials and activities in the [Red Hat Developer sandbox](#) environment.

Configure custom notebooks

In addition to notebook images provided and supported by Red Hat and independent software vendors (ISVs), you can configure custom notebook images that cater to your project's specific requirements.

3.2. FEATURES FOR IT OPERATIONS ADMINISTRATORS

General availability of OpenShift Data Science Self-managed

Starting with version 1.20, Red Hat OpenShift Data Science Self-managed is generally available. This allows OpenShift Data Science to be installed as a self-managed Operator on a self-managed environment, such as Red Hat OpenShift Container Platform.

Disconnected installation support

Red Hat OpenShift Data Science Self-managed now supports installation in a disconnected environment. Disconnected clusters are on a restricted network, typically behind a firewall. In this case, clusters cannot access the remote registries where Red Hat provided OperatorHub sources reside. Instead, the OpenShift Data Science Operator can be deployed to a disconnected environment using a private registry to mirror the images.

Manage users with an identity provider

OpenShift Data Science supports the same authentication systems as Red Hat OpenShift Container Platform. You can configure existing groups in your identity provider as administrators or users of OpenShift Data Science.

Manage resources with OpenShift Container Platform

Use OpenShift Container Platform to configure and manage the lifecycle of your container-based applications and their dependencies. OpenShift Container Platform deploys, configures, and manages containers. OpenShift Container Platform offers usability, stability, and customization of its components.

Control Red Hat usage data collection

Choose whether to allow Red Hat to collect data about OpenShift Data Science usage in your cluster. Usage data collection is enabled by default when you install OpenShift Data Science on your OpenShift Container Platform cluster.

Apply autoscaling to your cluster to reduce usage costs

Use the cluster autoscaler to adjust the size of your cluster to meet its current needs and optimize costs.

Customize PVC size to suit your workloads

Allocate the right amount of persistent storage for your data scientists by default to optimize resource costs and productivity.

Manage resource usage by stopping idle notebooks

Reduce resource usage in your OpenShift Data Science deployment by stopping notebook servers that have been idle (without logged in users) for a period of time.

Serving runtimes

OpenShift Data Science now provides support for model-serving runtimes. A model-serving runtime provides integration with a specified model server and the model frameworks that it supports. By default, Red Hat OpenShift Data Science includes the OpenVINO Model Server runtime. However, if this runtime doesn't meet your needs (it doesn't support a particular model framework, for example), you might want to add your own, custom runtimes. You specify a model-serving runtime when you configure a model server.

3.3. ENHANCEMENTS

This section describes enhancements to existing features in Red Hat OpenShift Data Science.

Support for creating OpenShift Data Science pipelines within JupyterLab

You can create data science pipelines within JupyterLab.

Support for data science project sharing

To enable you to work collaboratively on your data science projects with other users, you can share access to your project. After creating your project, you can then set the appropriate access permissions from the OpenShift Data Science user interface.

Support for older versions of notebook images

An older version of each notebook image is now available and supported. Two supported notebook images are now typically available at any given time. Notebook images are supported for a minimum of one year. Major updates to pre-configured notebook images occur approximately every six months.

Support for the TrustyAI notebook image

The TrustyAI notebook image is now available on Red Hat OpenShift Data Science. This notebook image is pre-built and usable immediately after OpenShift Data Science is installed or upgraded.

Support for data science projects

OpenShift Data Science now has the ability to create one or more projects to help organize your data science work for a project in one place. With data science projects, you can create a project work bench, add a data connection to a data source, and add cluster storage to a data science project.

Default persistent volume claim (PVC) size increased

The default size of a PVC provisioned for a data science user in an OpenShift Data Science cluster has been increased from 2 GB to 20 GB.

Improved resilience to OpenShift Dedicated node failure

OpenShift Data Science services now try to avoid being scheduled on the same node so that OpenShift Data Science components are more failure resistant.

Improved notebook controller

OpenShift Data Science now uses an improved notebook controller. This change enables future feature development and provides a better user experience in the event of errors, such as if a notebook server fails to launch correctly. If you have bookmarked the previous notebook controller URL, you will need to update your bookmark accordingly.

Support for AWS Security Token Service (STS) ROSA clusters

OpenShift Data Science now supports AWS (STS) enabled ROSA clusters.

3.4. LIMITED SUPPORT FEATURES

This section outlines features provided with limited support in Red Hat OpenShift Data Science.

CHAPTER 4. BUG FIXES



This section describes the fixes for notable user-facing issues in Red Hat OpenShift Data Science.

4.1. EXPORTING AN ELYRA PIPELINE EXPOSED S3 STORAGE CREDENTIALS IN PLAIN TEXT

In OpenShift Data Science 1.28.0, when you exported an Elyra pipeline from JupyterLab in Python DSL format or YAML format, the generated output contained S3 storage credentials in plain text. This issue has been resolved in OpenShift Data Science 1.28.1. However, after you upgrade to OpenShift Data Science 1.28.1, if your deployment contains a data science project with a pipeline server and a data connection, you must perform the following additional actions for the fix to take effect:

1. Refresh your browser page.
2. Stop any running workbenches in your deployment and restart them.

Furthermore, to confirm that your Elyra runtime configuration contains the fix, perform the following actions:

1. In the left sidebar of JupyterLab, click **Runtimes** ().
2. Hover the cursor over the runtime configuration that you want to view and click the **Edit** button ().
The **Data Science Pipelines runtime configuration** page opens.
3. Confirm that **KUBERNETES_SECRET** is defined as the value in the **Cloud Object Storage Authentication Type** field.
4. Close the runtime configuration without changing it.

4.2. WHEN EDITING THE DETAILS OF A SHARED PROJECT, THE USER INTERFACE REMAINED IN A LOADING STATE WITHOUT REPORTING AN ERROR

When a user with permission to edit a project attempted to edit its details, the user interface remained in a loading state and did not display an appropriate error message. Users with permission to edit projects cannot edit any fields in the project, such as its description. Those users can edit only components belonging to a project, such as its workbenches, data connections, and storage.

The user interface now displays an appropriate error message and does not try to update the project description.

4.3. DATA SCIENCE PIPELINE GRAPHS DID NOT DISPLAY NODE EDGES FOR RUNNING PIPELINES

If you ran pipelines that did not contain Tekton-formatted **Parameters** or **when** expressions in their YAML code, the OpenShift Data Science user interface did not display connecting edges to and from graph nodes. For example, if you used a pipeline containing the **runAfter** property or **Workspaces**, the user interface displayed the graph for the executed pipeline without edge connections. The OpenShift Data Science user interface now displays connecting edges to and from graph nodes.

4.4. NEWLY CREATED DATA CONNECTIONS WERE NOT DETECTED WHEN YOU ATTEMPTED TO CREATE A PIPELINE SERVER

If you created a data connection from within a Data Science project, and then attempted to create a pipeline server, the **Configure a pipeline server** dialog did not detect the data connection that you created. This issue is now fixed.

4.5. WHEN SHARING A PROJECT WITH ANOTHER USER, THE OPENSIFT DATA SCIENCE USER INTERFACE TEXT WAS MISLEADING

When you attempted to share a Data Science project with another user, the user interface text misleadingly implied that users could edit all of its details, such as its description. However, users can edit only components belonging to a project, such as its workbenches, data connections, and storage. This issue is now fixed and the user interface text no longer misleadingly implies that users can edit all of its details.

4.6. USERS WITH "EDIT" PERMISSION COULD NOT CREATE A MODEL SERVER

Users with "Edit" permissions can now create a Model Server without token authorization. Users must have "Admin" permissions to create a Model Server with token authorization.

4.7. OPENVINO MODEL SERVER RUNTIME DID NOT HAVE THE REQUIRED FLAG TO FORCE GPU USAGE

OpenShift Data Science includes the OpenVINO Model Server (OVMS) model-serving runtime by default. When you configured a new model server and chose this runtime, the **Configure model server** dialog enabled you to specify a number of GPUs to use with the model server. However, when you finished configuring the model server and deployed models from it, the model server did not actually use any GPUs. This issue is now fixed and the model server uses the GPUs.

4.8. CHANGING THE HOST PROJECT WHEN CREATING A PIPELINE RUN RESULTED IN AN INACCURATE LIST OF AVAILABLE PIPELINES

If you changed the host project while creating a pipeline run, the interface failed to make the pipelines of the new host project available. Instead, the interface showed pipelines that belong to the project you initially selected on the **Data Science Pipelines > Runs** page. This issue is now fixed. You no longer select a pipeline from the **Create run** page. The pipeline selection is automatically updated when you click the **Create run** button, based on the current project and its pipeline.

CHAPTER 5. KNOWN ISSUES

This section describes known issues in Red Hat OpenShift Data Science and any known methods of working around the issues described.

5.1. DEFAULT SHARED MEMORY FOR JUPYTER NOTEBOOK MIGHT CAUSE A RUNTIME ERROR

The default shared memory for a Jupyter notebook is set to 64 Mb and you cannot change this default value in the Notebook configuration. For example, PyTorch relies on shared memory and the default size of 64 Mb is not enough for large use cases, such as when training a model or when performing heavy data manipulations. Jupyter reports a "no space left on device" message and `/dev/smh` is full.

Workaround

1. In your data science project, create a workbench as described in [Creating a project workbench](#).
2. In the data science project page, in the **Workbenches** section, click the **Status** toggle for the workbench to change it from **Running** to **Stopped**.
3. Open your OpenShift Console and then select **Administrator**.
4. Select **Home** → **API Explorer**.
5. In the **Filter by kind** field, type **notebook**.
6. Select the **kubeflow v1** notebook.
7. Select the **Instances** tab and then select the instance for the workbench that you created in Step 1.
8. Click the **YAML** tab and then select **Actions** → **Edit Notebook**.
9. Edit the YAML file to add the following information to the configuration:
 - For the container that has the name of your Workbench notebook, add the following lines to the **volumeMounts** section:

```
- mountPath: /dev/shm
  name: shm
```

For example, if your workbench name is **myworkbench**, update the YAML file as follows:

```
spec:
  containers:
  - env
    ...
    name: myworkbench
    ...
    volumeMounts:
    - mountPath: /dev/shm
      name: shm
```

- In the volumes section, add the lines shown in the following example:

```
volumes:
  name: shm
  emptyDir:
    medium: Memory
```

Note: Optionally, you can specify a limit to the amount of memory to use for the **emptyDir**.

10. Click **Save**.
11. In the data science dashboard, in the **Workbenches** section of the data science project, click the **Status** toggle for the workbench. The status changes from **Stopped** to **Starting** and then **Running**.
12. Restart the notebook.



WARNING

If you later edit the notebook's configuration through the Data Science dashboard UI, your workaround edit to the notebook configuration will be erased.

5.2. DATA SCIENCE DASHBOARD DOES NOT DETECT AN EXISTING OPENSIFT PIPELINES INSTALLATION

When the OpenShift Pipelines operator is installed as a global operator on your cluster, the Data Science dashboard does not properly detect it.

An alert icon appears next to the **Data Science Pipelines** option in the left navigation bar. When you open **Data Science Pipelines**, you see the message: "To use pipelines, first install the Red Hat OpenShift Pipelines Operator." However, when you view the list of installed operators in the **openshift-operators** project, you see that OpenShift Pipelines is installed as a global operator on your cluster.

Workaround

Follow these steps as a user with **cluster-admin** permissions:

1. Log in to your cluster using the **oc** client.
2. Enter the following command to update **OdhDashboardConfig** in the **redhat-ods-applications** application namespace:

```
$ oc patch OdhDashboardConfig odh-dashboard-config -n redhat-ods-applications --
type=merge -p '{"spec": {"dashboardConfig": {"disablePipelines": false}}}'
```

5.3. ELYRA PIPELINE FAILS TO RUN IF WORKBENCH IS CREATED BY A USER WITH EDIT PERMISSIONS

If a user who has been granted edit permissions for a project creates a project workbench, that user sees the following behavior:

- During the workbench creation process, the user sees an **Error creating workbench** message related to the creation of Kubernetes role bindings.
- Despite the preceding error message, OpenShift Data Science still creates the workbench. However, the error message means that the user will not be able to use the workbench to run Elyra data science pipelines.
- If the user tries to use the workbench to run an Elyra pipeline, Jupyter shows an **Error making request** message that describes failed initialization.

Workaround: A user with administrator permissions (for example, the project owner) must create the workbench on behalf of the user with edit permissions. That user can then use the workbench to run Elyra pipelines.

5.4. DEPLOYING A CUSTOM MODEL-SERVING RUNTIME CAN RESULT IN AN ERROR MESSAGE

If you use the Open Shift Data Science dashboard to deploy a custom model-serving runtime, the deployment process can fail with an **Error retrieving Serving Runtime** message.

Workaround: There are two workarounds:

- Deploy the custom serving runtime by using the CLI.
- After you create the serving runtime in the dashboard, open the **Serving runtimes** page and then edit the YAML file for the runtime to specify **serverType**, as shown in this example for the OpenVino Model Server:

```
spec:
  builtInAdapter:
    serverType: ovms
```

5.5. PIPELINES WITH NON-UNIQUE NAMES DO NOT APPEAR IN THE DATA SCIENCE PROJECT USER INTERFACE

If you launch a notebook from a Jupyter application that supports Elyra, or if you use a workbench, when you submit a pipeline to be run, pipelines with non-unique names do not appear in the **Pipelines** section of the relevant data science project page or the **Pipelines** heading of the data science pipelines page.

Workaround: Give a unique name to each pipeline when submitting it from Elyra so that the pipeline shows up correctly in both the data science project and the data science pipelines pages.

5.6. UNINSTALL PROCESS FOR OPENSIFT DATA SCIENCE MIGHT BECOME STUCK WHEN REMOVING KFDEFS RESOURCES

The steps for uninstalling OpenShift Data Science self-managed are described in [Uninstalling OpenShift Data Science self-managed](#).

However, even when you follow this guide, you might see that the uninstall process does not finish successfully. Instead, the process stays stuck on the step of deleting **kfdefs** resources that are used by the KubeFlow Operator. As shown in the following example, **kfdefs** resources might exist in the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** namespaces:

```
$ oc get kfdefs.kfdef.apps.kubeflow.org -A
```

NAMESPACE	NAME	AGE
redhat-ods-applications	rhods-anaconda	3h6m
redhat-ods-applications	rhods-dashboard	3h6m
redhat-ods-applications	rhods-data-science-pipelines-operator	3h6m
redhat-ods-applications	rhods-model-mesh	3h6m
redhat-ods-applications	rhods-nbc	3h6m
redhat-ods-applications	rhods-osd-config	3h6m
redhat-ods-monitoring	modelmesh-monitoring	3h6m
redhat-ods-monitoring	monitoring	3h6m
rhods-notebooks	rhods-notebooks	3h6m
rhods-notebooks	rhods-osd-config	3h5m

Failed removal of the **kfdefs** resources might also prevent later installation of a newer version of OpenShift Data Science.

Workaround: To manually delete the **kfdefs** resources so that you can complete the uninstall process, see the "**Force individual object removal when it has finalizers**" section of the following Red Hat solution article: <https://access.redhat.com/solutions/4165791>.

5.7. DATA SCIENCE PIPELINE SETUP FAILS WITH AN SSL CERTIFICATION ERROR MESSAGE

If you try to set up a data science pipelines on an unsecured cluster, the setup fails with an error message related to SSL certification on the cluster.

Workaround: Before you create a workbench on the cluster, manually create the following environment variable:

```
PIPELINES_SSL_SA_CERTS=/var/run/secrets/kubernetes.io/serviceaccount/ca.crt
```

5.8. RUNNING PIPELINES FROM ELYRA WITHIN JUPYTERLAB - NOT SUPPORTED IN A DISCONNECTED ENVIRONMENT

To run a data science pipeline submitted from Elyra (within JupyterLab) requires internet access and is not supported in a disconnected environment.

5.9. AFTER UPGRADE, THE DATA SCIENCE PIPELINES TAB IS NOT ENABLED ON THE OPENSIFT DATA SCIENCE DASHBOARD

After you upgrade from OpenShift Data Science 1.26 to OpenShift Data Science 1.28, the Data Science Pipelines tab is not enabled on the OpenShift Data Science dashboard.

If you are upgrading from OpenShift Data Science 1.26 to OpenShift Data Science 1.28, follow these steps:

Workaround: As a workaround, perform the following steps as a user with **cluster-admin** permissions:

1. Upgrade from OpenShift Data Science 1.26 to OpenShift Data Science 1.27.
2. Upgrade from OpenShift Data Science 1.27 to OpenShift Data Science 1.28.

3. Log in to your cluster using the **oc** client.
4. Enter the following command to update **OdhDashboardConfig** in the **redhat-ods-applications** application namespace:

```
$ oc patch OdhDashboardConfig odh-dashboard-config -n redhat-ods-applications --
type=merge -p '{"spec": {"dashboardConfig": {"disablePipelines": false}}}'
```

5.10. INCORRECT CRON FORMAT DISPLAYED BY DEFAULT WHEN SCHEDULING A RECURRING PIPELINE RUN

When you schedule a recurring pipeline run by configuring a cron job, the OpenShift Data Science interface displays the following incorrect format by default:

```
cron = 0 0 0 * *
```

The correct format must comply with the [CRON expression format for the Go cron package](#). By default, the correct format is:

```
cron = 0 0 0 * * *
```

The following example shows a pipeline run that executes daily at 16:16 UTC using the correct format, scheduled using a cron job in the user interface:

```
cron = 0 16 16 * * *
```

5.11. YOU CANNOT CREATE A PIPELINE SERVER WHEN CUMULATIVE CHARACTER LIMIT IS EXCEEDED

When the cumulative character limit of a data science project name and a pipeline server name exceeds 62 characters, you are unable to successfully create a pipeline server.

Workaround: Rename your data science project so that it does not exceed 30 characters.

5.12. A PIPELINE SERVER FAILS TO START UNLESS YOU SPECIFY AN AMAZON WEB SERVICES (AWS) SIMPLE STORAGE SERVICE (S3) BUCKET RESOURCE

When you create a data connection for a data science project, the **AWS_S3_BUCKET** field is not designated as a mandatory field. However, if you do not specify a value for this field, and you attempt to configure a pipeline server, the pipeline server fails to start successfully.

5.13. ERROR APPEARS WHEN INSTALLING THE RED HAT OPENSIFT DATA SCIENCE OPERATOR TO AN OPENSIFT CLUSTER IN A DISCONNECTED ENVIRONMENT WHEN USING OC-MIRROR VERSION 4.12 OR OLDER

When installing the Red Hat OpenShift Data Science Operator to an OpenShift cluster in a disconnected environment when using **oc mirror** version 4.12 or older, the following error is displayed

when mirroring the required container images to a private container registry using the **oc mirror** command.

```
error: unable to push manifest to file:///modh/cuda-notebooks:latest: symlink
sha256:348fa993347f86d1e0913853fb726c584ae8b5181152f0430967d380d68d804f mirror-
rhods/oc-mirror-workspace/src/v2/modh/cuda-notebooks/manifests/latest.download: file exists
```

Workaround: When mirroring images to a private registry for a disconnected installation, ensure that you have installed **oc-mirror** OpenShift CLI (**oc**) plug-in, version 4.13 or greater. Versions of **oc-mirror** preceding version 4.13 do not allow you to mirror the full image set configuration provided. **oc-mirror** version 4.13 is compatible with previous versions of OpenShift.

For more information about the procedure involving this workaround, see [Mirroring images to a private registry for a disconnected installation](#).

5.14. USER WITHOUT DASHBOARD PERMISSIONS IS ABLE TO CONTINUE USING THEIR RUNNING NOTEBOOKS AND WORKBENCHES INDEFINITELY

When a Red Hat OpenShift Data Science administrator revokes a user's permissions, the user can continue to use their running notebooks and workbenches indefinitely.

Workaround: When the OpenShift Data Science administrator revokes a user's permissions, the administrator should also stop any running notebooks and workbenches for that user.

5.15. ATTEMPTING TO INCREASE THE SIZE OF A PERSISTENT VOLUME (PV) FAILS WHEN IT IS NOT CONNECTED TO A WORKBENCH

Attempting to increase the size of a Persistent Volume (PV) that is not connected to a workbench fails. When changing a data science project's storage, users can still edit the size of the PV in the user interface, but this action does not have any effect.

5.16. UNABLE TO SCALE DOWN A WORKBENCH'S GPUS WHEN ALL GPUS IN THE CLUSTER ARE BEING USED

It is not possible to scale down a workbench's GPUs if all GPUs in the cluster are being used. This issue applies to GPUs being used by one workbench, and GPUs being used by multiple workbenches.

Workaround: To workaround around this issue, perform the following steps:

1. Stop all active workbenches that are using GPUs.
2. Wait until the relevant GPUs are available again.
3. Edit the workbench and scale down the GPU instances.

5.17. ANACONDA PROFESSIONAL EDITION CANNOT BE VALIDATED AND ENABLED IN OPENSIFT DATA SCIENCE

Anaconda Professional Edition cannot be enabled as the dashboard's key validation for Anaconda Professional Edition is inoperable.

5.18. UNCLEAR ERROR MESSAGE DISPLAYS WHEN USING INVALID CHARACTERS TO CREATE A DATA SCIENCE PROJECT

When creating a data science project's data connection, workbench, or storage connection using invalid special characters, the following error message is displayed:

```
the object provided is unrecognized (must be of type Secret): couldn't get version/kind; json parse error: unexpected end of JSON input ({"apiVersion":"v1","kind":"Sec ...)
```

The error message fails to clearly indicate the problem.

5.19. AN ERROR CAN OCCUR WHEN TRYING TO EDIT A WORKBENCH

When editing a workbench, an error similar to the following can occur:

```
Error creating workbench
Operation cannot be fulfilled on notebooks.kubeflow.org "workbench-name": the object has been modified; please apply your changes to the latest version and try again
```

5.20. AN IMAGEPULLBACKOFF ERROR MESSAGE IS NOT DISPLAYED WHEN REQUIRED DURING THE WORKBENCH CREATION PROCESS

Pods can experience issues pulling container images from the container registry. If an error occurs, the relevant pod enters into an **ImagePullBackOff** state. During the workbench creation process, if an **ImagePullBackOff** error occurs, an appropriate message is not displayed.

Workaround: Check the event log for further information on the **ImagePullBackOff** error. To do this, click on the workbench status when it is starting.

5.21. WORKBENCHES FAIL TO START WHEN CUMULATIVE CHARACTER LIMIT IS EXCEEDED.

When the cumulative character limit of a data science project's title and workbench title exceeds 62 characters, workbenches fail to start.

5.22. THE NOTEBOOK CREATION PROCESS FAILS FOR USERS WHO HAVE NEVER LOGGED IN TO THE DASHBOARD

The dashboard's notebook **Administration** page displays users belonging to the user group and admin group in OpenShift. However, if an administrator attempts to start a notebook server on behalf of a user who has never logged in to the dashboard, the server creation process fails and displays the following error message:

```
Request invalid against a username that does not exist.
```

Workaround: Request that the relevant user logs into the dashboard.

5.23. THE MODEL MESH OAUTH-PROXY CONTAINER IS INTERMITTENTLY UNSTABLE

ModelMesh pods do not deploy correctly due to a failure of the ModelMesh **oauth-proxy** container. This issue occurs intermittently and only if authentication is enabled in the ModelMesh runtime environment. It is more likely to occur when additional ModelMesh instances are deployed in different namespaces.

5.24. THE NVIDIA GPU OPERATOR IS INCOMPATIBLE WITH OPENSIFT 4.11.12

Provisioning a GPU node on a OpenShift 4.11.12 cluster results in the **nvidia-driver-daemonset** pod getting stuck in a CrashLoopBackOff state. The NVIDIA GPU Operator is compatible with OpenShift 4.11.9 and 4.11.13.

5.25. INCORRECT PACKAGE VERSION DISPLAYED DURING NOTEBOOK SELECTION

The **Start a notebook server** page displays an incorrect version number for the Anaconda notebook image.

5.26. WHEN USING THE NVIDIA GPU OPERATOR, MORE NODES THAN NEEDED ARE CREATED BY THE NODE AUTOSCALER

When a pod cannot be scheduled due to insufficient available resources, the Node Autoscaler creates a new node. There is a delay until the newly created node receives the relevant GPU workload. Consequently, the pod cannot be scheduled and the Node Autoscaler's continuously creates additional new nodes until one of the nodes is ready to receive the GPU workload. For more information about this issue, see [When using the NVIDIA GPU Operator, more nodes than needed are created by the Node Autoscaler](#).

Workaround: Apply the **cluster-api/accelerator** label in **machineset.spec.template.spec.metadata**. This causes the autoscaler to consider those nodes as unready until the GPU driver has been deployed.

5.27. THE APPLICATION LAUNCHER MENU INCORRECTLY DISPLAYS A LINK TO OPENSIFT CLUSTER MANAGER

Red Hat OpenShift Data Science incorrectly displays a link to the OpenShift Cluster Manager from the application launcher menu. Clicking this link results in a "Page Not Found" error because the URL is not valid.

5.28. NOTEBOOK SERVER ADMINISTRATION PAGE SHOWS USERS WHO HAVE LOST PERMISSION ACCESS

If a user who previously started a notebook server in Jupyter loses their permissions to do so (for example, if an OpenShift Data Science administrator changes the user's group settings or removes the user from a permitted group), administrators continue to see the user's notebook servers on the server **Administration** page. As a consequence, an administrator is able to restart notebook servers that belong to the user who's permissions were revoked.

5.29. GPUS ON NODES WITH UNSUPPORTED TAINTS CANNOT BE ALLOCATED TO NOTEBOOK SERVERS

GPUs on nodes marked with any taint other than the supported `nvidia.com/gpu` taint cannot be selected when creating a notebook server. To avoid this issue, use only the `nvidia.com/gpu` taint on GPU nodes used with OpenShift Data Science.

5.30. TENSORBOARD REQUIRES MANUAL STEPS TO VIEW

When a user has TensorFlow or PyTorch notebook images and wants to use TensorBoard to display data, manual steps are necessary to include environment variables in the notebook environment, and to import those variables for use in your code.

Workaround: When you start your notebook server, use the following code to set the value for the `TENSORBOARD_PROXY_URL` environment variable to use your OpenShift Data Science user ID.

```
import os
os.environ["TENSORBOARD_PROXY_URL"]= os.environ["NB_PREFIX"]+"/proxy/6006/"
```

5.31. THE INTEL® ONEAPI AI ANALYTICS TOOLKITS QUICK START REFERENCES NONEXISTENT SAMPLE NOTEBOOKS

The Intel® oneAPI AI Analytics Toolkits quick start, located on the **Resources** page on the dashboard, requires the user to load sample notebooks as part of the instruction steps, but refers to notebooks that do not exist in the associated repository.

5.32. THE CRONJOB RESPONSIBLE FOR VALIDATING ANACONDA PROFESSIONAL EDITION'S LICENSE IS SUSPENDED AND DOES NOT RUN DAILY

The CronJob responsible for validating Anaconda Professional Edition's license is automatically suspended by the OpenShift Data Science operator. As a result, the CronJob does not run daily as scheduled. In addition, when Anaconda Professional Edition's license expires, Anaconda Professional Edition is not indicated as disabled on the OpenShift Data Science dashboard.

5.33. THE NVIDIA GPU OPERATOR CARD ON THE DASHBOARD DISPLAYS BUTTON UNNECESSARILY

GPUs are automatically available in Jupyter after the NVIDIA GPU Operator is installed. The **Enable** button, located on the NVIDIA GPU Operator card on the **Explore** page, is therefore redundant. In addition, clicking the **Enable** button moves the NVIDIA GPU Operator card to the **Enabled** page, even if the Operator is not installed.

5.34. DASHBOARD DOES NOT DISPLAY ENABLED PAGE CONTENT AFTER ISV OPERATOR UNINSTALL

After an ISV operator is uninstalled, no content is displayed on the **Enabled** page on the dashboard. Instead, the following error is displayed:

```
Error loading components
HTTP request failed
```

Workaround: Wait 30-40 seconds and then refresh the page in your browser.

5.35. INCORRECT PACKAGE VERSIONS DISPLAYED DURING NOTEBOOK SELECTION

In the OpenShift Data Science interface, the **Start a notebook server page** displays incorrect version numbers for the JupyterLab and Notebook packages included in the oneAPI AI Analytics Toolkit notebook image. The page might also show an incorrect value for the Python version used by this image.

Workaround: When you start your oneAPI AI Analytics Toolkit notebook server, you can check which Python packages are installed on your notebook server and which version of the package you have by running the **!pip list** command in a notebook cell.

5.36. ERROR CAN OCCUR WHEN CREATING A NOTEBOOK INSTANCE

When creating a notebook instance in Jupyter, a **Directory not found** error appears intermittently. This error message can be ignored by clicking **Dismiss**.

5.37. ACTIONS ON DASHBOARD NOT CLEARLY VISIBLE

The dashboard actions to re-validate a disabled application's license, and to remove a disabled application's card are not clearly visible to the user. These actions only appear when the user clicks on the application card's **Disabled** label. As a result, the intended workflows may not be clear to the user.

5.38. LICENSE RE-VALIDATION ACTION APPEARS UNNECESSARILY

The dashboard action to re-validate a disabled application's license appears unnecessarily for applications that do not have a license validation or activation system. In addition, when a user attempts to re-validate a license that cannot be re-validated, feedback is not displayed to state why the action cannot be completed.

5.39. ERROR CAN OCCUR DURING PACHYDERM DEPLOYMENT

When creating an instance of the Pachyderm operator, a webhook error appears intermittently, preventing the creation process from starting successfully. The webhook error is indicative that, either the Pachyderm operator failed a health check, causing it to restart, or that the operator process exceeded its container's allocated memory limit, triggering an Out of Memory (OOM) kill.

Workaround: Repeat the Pachyderm instance creation process until the error no longer appears.

5.40. IBM WATSON STUDIO NOT AVAILABLE IN OPENSIFT DATA SCIENCE

IBM Watson Studio is not available when OpenShift Data Science is installed on OpenShift Dedicated 4.9 or higher, because it is not compatible with these versions of OpenShift Dedicated. Contact [Marketplace support](#) for assistance manually configuring Watson Studio on OpenShift Dedicated 4.9 and higher.

5.41. UNNECESSARY WARNINGS ABOUT MISSING GRAPHICAL PROCESSING UNITS (GPUS)

The TensorFlow notebook image checks for graphical processing units (GPUs) whenever a notebook is run, and issues warnings about missing GPUs when none are present. These messages can safely be ignored, but you can disable them by running the following in a notebook when you start a notebook

server that uses the TensorFlow notebook image.

```
import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'
```

5.42. USER SESSIONS PERSIST IN SOME COMPONENTS

Although users of OpenShift Data Science and its components are authenticated through OpenShift, session management is separate from authentication. This means that logging out of OpenShift Dedicated or OpenShift Data Science does not affect a logged in Jupyter session running on those platforms. When a user's permissions change, that user must log out of all current sessions so that changes take effect.