# Red Hat OpenShift Data Science self-managed 2.4

# Release notes

Features, enhancements, resolved issues, and known issues associated with this release

# Red Hat OpenShift Data Science self-managed 2.4 Release notes

Features, enhancements, resolved issues, and known issues associated with this release

## Legal Notice

## Abstract

These release notes provide an overview of new features, enhancements, resolved issues, and known issues in version 2.4 of Red Hat OpenShift Data Science.

# Table of Contents

# CHAPTER 1. OVERVIEW OF OPENSHIFT 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) and Habana Gaudi devices.

For administrators, OpenShift Data Science enables data science workloads in an existing Red Hat OpenShift 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 2. NEW FEATURES AND ENHANCEMENTS

This section describes new features and enhancements in Red Hat OpenShift Data Science 2.4.

## 2.1. NEW FEATURES

**Redesigned Red Hat OpenShift Data Science Operator**

This release introduces a redesigned version of the Red Hat OpenShift Data Science Operator. The redesigned Operator gives you the flexibility to install only the components that you need for your specific workloads, instead of installing the entire OpenShift Data Science product. To learn more about how to customize your installation, see Overview of installing and deploying OpenShift Data Science. This feature was first introduced in OpenShift Data Science 2.4.

## 2.2. ENHANCEMENTS

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

**Support for duplication of custom model-serving runtimes**

In earlier releases, you could duplicate a provided model-serving runtime. You can now also duplicate a custom model-serving runtime. This enhancement was first introduced in OpenShift Data Science 2.4.

**Support for pipeline actions from the Runs home page**

In earlier releases, you could clone or stop a pipeline run from the pipeline **Run details** page only. You can now also clone or stop a pipeline run from the **Actions** menu on the **Runs** home page. This enhancement was first introduced in OpenShift Data Science 2.4.

**Improved user experience for custom notebooks**

If you have administrator privileges, you can add custom notebook images to your OpenShift Data Science environment by clicking **Settings → Notebook image settings → Import new image** In this release, the UI has been updated to improve the user experience of adding a custom notebook as follows:

- It is now easier to find your resource in the OpenShift Container Platform console:

  - The resource name has changed from **byon-########** to **custom-{your name}**.

  - You can use the new copy feature to copy the name, and then paste the name into the **Search** page in the OpenShift Container Platform console.

- Simple registry failures are detected earlier, preventing the creation of invalid custom images.

- When you add metadata to your images, the software and package flows are improved:

  - You can use keyboard navigation.

  - Pressing the Enter key brings you to the next page instead of completing the current page, which reduces the number of manual clicks.

This enhancement was first introduced in OpenShift Data Science 2.4.

**New method to send documentation feedback**

In earlier releases, you could use the Direct Documentation Feedback feature in the Customer Portal to send feedback on the documentation. The Direct Documentation Feedback feature is no longer available in the Customer Portal.

In this release, the new method to send documentation feedback is as follows:

1. Open a Jira ticket.

2. In the **Summary** field, enter a descriptive title.

3. In the **Description** field, enter your suggestion for improvement. Include links to the relevant parts of the documentation.

4. Click **Create** at the end of the dialog. A documentation issue will be created and routed to the appropriate documentation team.

This enhancement was first introduced in OpenShift Data Science 2.4.

# CHAPTER 3. TECHNOLOGY PREVIEW FEATURES

**IMPORTANT**

This section describes Technology Preview features in Red Hat OpenShift Data Science 2.4. Technology Preview features are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope .

**Accelerator profiles**

An *accelerator* is a specialized hardware component that increases the efficiency, speed, and scalability of compute-intensive tasks. For Red Hat OpenShift Data Science, the supported accelerators are NVIDIA graphics processing units (GPUs) and Habana Gaudi devices (Gaudi 1 and Gaudi 2). An *accelerator profile* defines the specification of an accelerator. Before you can use an accelerator in OpenShift Data Science, your OpenShift instance must contain the associated accelerator profile. Administrators can configure Red Hat OpenShift Data Science to enable users to select a specific type of accelerator that is most appropriate for a workload, in addition to specifying the preferred number of accelerators. For information on how to manually create an accelerator profile, see Working with accelerator profiles . The OpenShift Data Science upgrade process automatically creates an accelerator profile for existing NVIDIA GPUs. This profile can be modified after the upgrade. The accelerator profiles feature is currently available in Red Hat OpenShift Data Science 2.4 as a Technology Preview feature. This feature was first introduced in OpenShift Data Science 2.4.

**HabanaAI notebook image**

The HabanaAI notebook image optimizes high-performance deep learning (DL) with Habana Gaudi devices. Habana Gaudi devices accelerate DL training workloads and maximize training throughput and efficiency. Notebook images available on Red Hat OpenShift Data Science are pre-built and ready for you to use immediately after you install or upgrade OpenShift Data Science. You can use this feature with HabanaAI Operator 1.10 only. The HabanaAI notebook image feature is currently available in Red Hat OpenShift Data Science 2.4 as a Technology Preview feature. This feature was first introduced in OpenShift Data Science 2.4.

**Distributed workloads**

Distributed workloads enable data scientists to use multiple cluster nodes in parallel for faster, more efficient data processing and model training. The CodeFlare framework simplifies task orchestration and monitoring, and offers seamless integration for automated resource scaling and optimal node utilization with advanced GPU support. Designed for data scientists, the CodeFlare framework enables direct workload configuration from Jupyter Notebooks or Python code, ensuring a low barrier of adoption, and streamlined, uninterrupted workflows. Distributed workloads significantly reduce task completion time, and enable the use of larger datasets and more complex models. The distributed workloads feature is currently available in Red Hat OpenShift Data Science 2.4 as a Technology Preview feature. This feature was first introduced in OpenShift Data Science 2.4.

# CHAPTER 4. LIMITED AVAILABILITY FEATURES

This section describes Limited Availability features in Red Hat OpenShift Data Science 2.4. Limited Availability means that you can install and receive support for the feature only with specific approval from Red Hat. Without such approval, the feature is unsupported. This applies to all features described in this section.

**Composite model-serving runtime based on Caikit and TGIS**

A composite model-serving runtime that is based on Caikit and TGIS is available as a Limited Availability feature. You can install the runtime only using Red Hat OpenShift Data Science Operator version 2.1 or later. The runtime consists of the following components:

- TGIS

- Caikit

- Caikit-nlp

- KServe

- Service Mesh

- Serverless

# CHAPTER 5. RESOLVED ISSUES

This section describes notable issues that have been resolved in Red Hat OpenShift Data Science 2.4.

**DATA-SCIENCE-PIPELINES-OPERATOR-294 – Scheduled pipeline run that uses data-passing might fail to pass data between steps, or fail the step entirely**

A scheduled pipeline run that uses an S3 object store to store the pipeline artifacts might fail with an error such as the following:

> Bad value for --endpoint-url "cp": scheme is missing. Must be of the form http://<hostname>/ or https://<hostname>/

This issue occurred because the S3 object store endpoint was not successfully passed to the pods for the scheduled pipeline run. This issue is now resolved.

**RHODS-4769 – 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 could not be selected when creating a notebook server. This issue is now resolved.

**RHODS-6346 – 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 was 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 failed to clearly indicate the problem. The error message now indicates that invalid characters were entered.

**RHODS-6950 – Unable to scale down workbench GPUs when all GPUs in the cluster are being used**

In earlier releases, it was not possible to scale down workbench GPUs if all GPUs in the cluster were being used. This issue applied to GPUs being used by one workbench, and GPUs being used by multiple workbenches. You can now scale down the GPUs by selecting **None** from the **Accelerators** list.

**RHODS-8939 – Default shared memory for a Jupyter notebook created in a previous release causes a runtime error**

Starting with release 1.31, this issue is resolved, and the shared memory for any new notebook is set to the size of the node.

For a Jupyter notebook created in a release earlier than 1.31, the default shared memory for a Jupyter notebook is set to 64 MB and you cannot change this default value in the notebook configuration.

To fix this issue, you must recreate the notebook or follow the process described in the Knowledgebase article How to change the shared memory for a Jupyter notebook in Red Hat OpenShift Data Science .

**RHODS-9030 – Uninstall process for OpenShift Data Science might become stuck when removing kfdefs resources**

The steps for uninstalling OpenShift Data Science self-managed are described in Uninstalling Red Hat OpenShift Data Science self-managed.

However, even when you followed this guide, you might have seen that the uninstall process did not finish successfully. Instead, the process stayed on the step of deleting **kfdefs** resources that were 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 have also prevented later installation of a newer version of OpenShift Data Science. This issue no longer occurs in Red Hat OpenShift Data Science 2.4.

RHODS-9764 – **Data connection details get reset when editing a workbench**

When you edited a workbench that had an existing data connection and then selected the **Create new data connection** option, the edit page might revert to the **Use existing data connection** option before you had finished specifying the new connection details. This issue is now resolved.

# CHAPTER 6. KNOWN ISSUES

This section describes known issues in Red Hat OpenShift Data Science 2.4 and any known methods of working around these issues.

**DATA-SCIENCE-PIPELINES-165 – Poor error message when S3 bucket is not writable**

When you set up a data connection and the S3 bucket is not writable, and you try to upload a pipeline, the error message **Failed to store pipelines** is not helpful.

**Workaround**

> Verify that your data connection credentials are correct and that you have write access to the bucket you specified.

**KUBEFLOW-177 – Bearer token from application not forwarded by OAuth-proxy**

You cannot use an application as a custom workbench image if its internal authentication mechanism is based on a bearer token. The OAuth-proxy configuration removes the bearer token from the headers, and the application cannot work properly.

**ODH-DASHBOARD-1335 – Rename Edit permission to Contributor**

The term *Edit* is not accurate:

- For *most* resources, users with the **Edit** permission can not only edit the resource, they can also create and delete the resource.

- Users with the **Edit** permission cannot edit the project.

The term *Contributor* more accurately describes the actions granted by this permission.

**ODH-DASHBOARD-1758 – Error duplicating OOTB custom serving runtimes several times**

If you duplicate a model-serving runtime several times, the duplication fails with the **Serving runtime name "<name>" already exists** error message.

**Workaround**

> Change the **metadata.name** field to a unique value.

**ODH-DASHBOARD-1771 – JavaScript error during Pipeline step initializing**

Sometimes the pipeline **Run details** page stops working when the run starts.

**Workaround**

> Refresh the page.

**ODH-DASHBOARD-1781 – Missing tooltip for Started Run status**

Data science pipeline runs sometimes don't show the tooltip text for the status icon shown.

**Workaround**

> For more information, view the pipeline **Run details** page and see the run output.

**ODH-DASHBOARD-1908 – Cannot create workbench with an empty environment variable**

When creating a workbench, if you click **Add variable** but do not select an environment variable type from the list, you cannot create the workbench. The field is not marked as required, and no error message is shown.

## ODH-DASHBOARD-1928 – Custom serving runtime creation error message is unhelpful

When you try to create or edit a custom model-serving runtime and an error occurs, the error message does not indicate the cause of the error.

Example error message: **Request failed with status code 422**

### Workaround

Check the YAML code for the serving runtime to identify the reason for the error.

## ODH-DASHBOARD-1991 – ovms-gpu-ootb is missing recommended accelerator annotation

When you add a model server to your project, the **Serving runtime** list does not show the **Recommended serving runtime** label for the NVIDIA GPU.

### Workaround

Make a copy of the model-server template and manually add the label.

## ODH-DASHBOARD-2140 – Package versions displayed in dashboard do not match installed versions

The dashboard might display inaccurate version numbers for packages such as JupyterLab and Notebook. The package version number can differ in the image if the packages are manually updated.

### Workaround

To find the true version number for a package, run the **pip list** command and search for the package name, as shown in the following examples:

```
$ pip list | grep jupyterlab
jupyterlab                3.5.3
$ pip list | grep notebook
notebook                  6.5.3
```

## RHODS-12432 – Deletion of the notebook-culler ConfigMap causes Permission Denied on dashboard

If you delete the **notebook-controller-culler-config** ConfigMap in the **redhat-ods-applications** namespace, you can no longer save changes to the **Cluster Settings** page on the OpenShift Data Science dashboard. The save operation fails with an **HTTP request has failed** error.

### Workaround

Complete the following 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 the **OdhDashboardConfig** custom resource in the **redhat-ods-applications** application namespace:

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

RHODS-12717 - Pipeline server creation might fail on OpenShift Container Platform with Open Virtual Network on OpenStack

When you try to create a pipeline server on OpenShift Container Platform with Open Virtual Network on OpenStack, the creation might fail with a **Pipeline server failed** error. See OCPBUGS-22251.

RHODS-12798 - Pods fail with "unable to init seccomp" error

Pods fail with **CreateContainerError** status or **Pending** status instead of **Running** status, because of a known kernel bug that introduced a **seccomp** memory leak. When you check the events on the namespace where the pod is failing, or run the **oc describe pod** command, the following error appears:

> runc create failed: unable to start container process: unable to init seccomp: error loading seccomp filter into kernel: error loading seccomp filter: errno 524

**Workaround**

Increase the value of **net.core.bpf_jit_limit** as described in the following Red Hat solution article: https://access.redhat.com/solutions/7030968.

RHODS-12899 - OpenVINO runtime missing annotation for NVIDIA GPUs

Red Hat OpenShift Data Science currently includes an out-of-the-box serving runtime that supports NVIDIA GPUs: **OpenVINO model server (support GPUs)** You can use the accelerator profile feature introduced in OpenShift Data Science 2.4 to select a specific accelerator in model serving, based on configured accelerator profiles. If the cluster had NVIDIA GPUs enabled in an earlier OpenShift Data Science release, the system automatically creates a default NVIDIA accelerator profile during upgrade to OpenShift Data Science 2.4. However, the **OpenVINO model server (supports GPUs)**runtime has not been annotated to indicate that it supports NVIDIA GPUs. Therefore, if a user selects the **OpenVINO model server (supports GPUs)**runtime and selects an NVIDIA GPU accelerator in the model server user interface, the system displays a warning that the selected accelerator is not compatible with the selected runtime. In this situation, you can ignore the warning. The accelerator profiles feature is currently available in Red Hat OpenShift Data Science as a Technology Preview feature. See Technology Preview features.

RHODS-12903 - Successfully-submitted Elyra pipeline fails to run

If you use a private TLS certificate, and you successfully submit an Elyra-generated pipeline against the data science pipeline server, the pipeline steps fail to execute, and the following error messages are shown:

> File "/opt/app-root/src/bootstrapper.py", line 747, in <module>
> main()
> File "/opt/app-root/src/bootstrapper.py", line 730, in main
> Actions
> ...
> WARNING: Retrying (Retry (total-4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.HTTPSConnection obj
> In this situation, a new runtime image should be created, to include the correct CA bundle, as well as all the required pip packages.

**Workaround**

Contact Red Hat Support for detailed steps to resolve this issue.

RHODS-12904 - Pipeline submitted from Elyra might fail when using private certificate

If you use a private TLS certificate, and you submit a pipeline from Elyra, the pipeline might fail with a **certificate verify failed** error message. This issue might be caused by either or both of the following situations:

- The object storage used for the pipeline server is using private TLS certificates.

- The data science pipeline server API endpoint is using private TLS certificates.

**Workaround**

Provide the workbench with the correct Certificate Authority (CA) bundle, and set various environment variables so that the correct CA bundle is recognized. Contact Red Hat Support for detailed steps to resolve this issue.

**RHODS-12906** – **Cannot use ModelMesh with object storage that uses private certificates**

Sometimes, when you store models in an object storage provider that uses a private TLS certificate, the model serving pods fail to pull files from the object storage, and the **signed by unknown authority** error message is shown.

**Workaround**

Manually update the secret created by the data connection so that the secret includes the correct CA bundle. Contact Red Hat Support for detailed steps to resolve this issue.

**RHODS-12928** – **Using unsupported characters can generate Kubernetes resource names with multiple dashes**

When you create a resource and you specify unsupported characters in the name, then each space is replaced with a dash and other unsupported characters are removed, which can result in an invalid resource name.

**RHODS-12937** – **Previously deployed model server might no longer work after upgrade in disconnected environment**

In disconnected environments, after upgrade to Red Hat OpenShift Data Science 2.4, previously deployed model servers might no longer work. The model status might be incorrectly reported as **OK** on the dashboard.

**Workaround**

Update the **inferenceservices** resource to replace the **storage** section with the **storageUri** section. In the following instructions, replace *<placeholders>* with the values for your environment.

1. Remove the **storage** parameter section from the existing **inferenceservices** resource:

   ```
   "storage":
       "key": "<your_key>",
       "path": "<your_path>"
   ```

   Example:

   ```
   "storage":
       "key": "aws-connection-minio-connection",
       "path": "mnist-8.onnx"
   ```

2. Add the **storageUri** section to the **inferenceservices** resource, with the specified format **s3://bucket-name/path/to/object**, as shown in the following example:

Example:

> storageUri: 's3://bucket/mnist-8.onnx'

3. Capture the secret key name as follows:

> secret_key=$(oc get secret -n *<project_name>* | grep -i aws-connection | awk '{print $1}')

4. Update the annotation as follows:

> oc annotate $(oc get inferenceservices -n *<project_name>* -o name) -n *<project_name>*
> serving.kserve.io/secretKey="$secret_key"

## RHODS-12946 – Cannot install from PyPI mirror in disconnected environment or when using private certificates

In disconnected environments, Red Hat OpenShift Data Science cannot connect to the public-facing PyPI repositories, so you must specify a repository inside your network. If you are using private TLS certificates, and a data science pipeline is configured to install Python packages, the pipeline run fails.

### Workaround

Add the required environment variables and certificates to your pipeline, as described in the following Red Hat solution article: https://access.redhat.com/solutions/7045831.

## RHODS-12986 – Potential reconciliation error after upgrade to Red Hat OpenShift Data Science 2.4

After you upgrade to Red Hat OpenShift Data Science 2.4, a reconciliation error might appear in the Red Hat OpenShift Data Science Operator pod logs and in the **DataScienceCluster** custom resource (CR) conditions.

Example error:

```
2023-11-23T09:45:37Z    ERROR    Reconciler error    {"controller": "datasciencecluster",
"controllerGroup": "datasciencecluster.opendatahub.io", "controllerKind": "DataScienceCluster",
"DataScienceCluster": {"name":"default-dsc"}, "namespace": "", "name": "default-dsc", "reconcileID":
"0c1a32ca-7ffd-4310-8259-f6baabf3c868", "error": "1 error occurred:\n\t* Deployment.apps \"rhods-
prometheus-operator\" is invalid: spec.selector: Invalid value:
v1.LabelSelector{MatchLabels:map[string]string{\"app.kubernetes.io/part-of\":\"model-mesh\",
\"app.opendatahub.io/model-mesh\":\"true\", \"k8s-app\":\"rhods-prometheus-operator\"},
MatchExpressions:[]v1.LabelSelectorRequirement(nil)}: field is immutable\n\n"}
```

### Workaround

Restart the Red Hat OpenShift Data Science Operator pod.

## RHOAIENG-11 – Separately installed instance of CodeFlare Operator not supported

In Red Hat OpenShift Data Science, the CodeFlare Operator is included in the base product and not in a separate Operator. Separately installed instances of the CodeFlare Operator from Red Hat or the community are not supported.

### Workaround

Delete any installed CodeFlare Operators, and install and configure Red Hat OpenShift Data Science, as described in the following Red Hat solution article: https://access.redhat.com/solutions/7043796.

## RHOAIENG-12 – Cannot access Ray dashboard from some browsers

In some browsers, users of the distributed workloads feature might not be able to access the Ray dashboard, because the browser automatically changes the prefix of the dashboard URL from **http** to **https**. The distributed workloads feature is currently available in Red Hat OpenShift Data Science as a Technology Preview feature. See Technology Preview features.

**Workaround**

Change the URL prefix from **https** to **http**.

## RHOAIENG-52 – Token authentication fails in clusters with self-signed certificates

If you use self-signed certificates, and you use the Python **codeflare-sdk** in a notebook or in a Python script as part of a pipeline, token authentication will fail.

## RHOAIENG-133 – Existing workbench cannot run Elyra pipeline after notebook restart

If you use the Elyra JupyterLab extension to create and run data science pipelines within JupyterLab, and you configure the pipeline server *after* you created a workbench and specified a notebook image within the workbench, you cannot execute the pipeline, even after restarting the notebook.

**Workaround**

1. Stop the running notebook.

2. Edit the workbench to make a small modification. For example, add a new dummy environment variable, or delete an existing unnecessary environment variable. Save your changes.

3. Restart the notebook.

4. In the left sidebar of JupyterLab, click **Runtimes**.

5. Confirm that the default runtime is selected.

## RHOAIENG-807 – Accelerator profile toleration removed when restarting a workbench

If you create a workbench that uses an accelerator profile that in turn includes a toleration, restarting the workbench removes the toleration information, which means that the restart cannot complete. A freshly created GPU-enabled workbench might start the first time, but never successfully restarts afterwards because the generated pod remains forever pending.

## NOTEBOOKS-218 – Data science pipelines saved from the Elyra pipeline editor reference an incompatible runtime

When you save a pipeline in the Elyra pipeline editor with the format **.pipeline** in OpenShift Data Science version 1.31 or earlier, the pipeline references a runtime that is incompatible with OpenShift Data Science version 1.32 or later.

As a result, the pipeline fails to run after you upgrade OpenShift Data Science to version 1.32 or later.

**Workaround**

After you upgrade to OpenShift Data Science to version 1.32 or later, select the relevant runtime images again.

NOTEBOOKS-210 - **A notebook fails to export as a PDF file in Jupyter**

When you export a notebook as a PDF file in Jupyter, the export process fails with an error.

DATA-SCIENCE-PIPELINES-OPERATOR-349 - **The Import Pipeline button is prematurely accessible**

When you import a pipeline to a workbench that belongs to a data science project, the **Import Pipeline** button is prematurely accessible before the pipeline server is fully available.

**Workaround**

Refresh your browser page and import the pipeline again.

DATA-SCIENCE-PIPELINES-OPERATOR-362 - **Pipeline server fails that uses object storage signed by an unknown authority**

Data science pipeline servers fail if you use object storage signed by an unknown authority. As a result, you cannot currently use object storage with a self-signed certificate. This issue has been observed in a disconnected environment.

**Workaround**

Configure your system to use object storage with a self-signed certificate, as described in the following Red Hat solution article: https://access.redhat.com/solutions/7040631.

ODH-DASHBOARD-1776 - **Error messages when user does not have project administrator permission**

If you do not have administrator permission for a project, you cannot access some features, and the error messages do not explain why. For example, when you create a model server in an environment where you only have access to a single namespace, an **Error creating model server** error message appears. However, the model server is still successfully created.

RHODS-11791 - **Usage data collection is enabled after upgrade**

If you previously had the **Allow collection of usage data** option deselected (that is, disabled), this option becomes selected (that is, enabled) when you upgrade OpenShift Data Science.

**Workaround**

Manually reset the **Allow collection of usage data** option. To do this, perform the following actions:

1. In the OpenShift Data Science dashboard, in the left menu, click **Settings → Cluster settings**.
   The **Cluster Settings** page opens.

2. In the **Usage data collection** section, deselect **Allow collection of usage data**.

3. Click **Save changes**.

ODH-DASHBOARD-1741 - **Cannot create a workbench whose name begins with a number**

If you try to create a workbench whose name begins with a number, the workbench does not start.

**Workaround**

Delete the workbench and create a new one with a name that begins with a letter.

## RHODS-6913 (ODH-DASHBOARD-1699) – Workbench does not automatically restart for all configuration changes

When you edit the configuration settings of a workbench, a warning message appears stating that the workbench will restart if you make any changes to its configuration settings. This warning is misleading because in the following cases, the workbench does not automatically restart:

- Edit name

- Edit description

- Edit, add, or remove keys and values of existing environment variables

**Workaround**

Manually restart the workbench.

## KUBEFLOW-157 – Logging out of JupyterLab does not work if you are already logged out of the OpenShift Data Science dashboard

If you log out of the OpenShift Data Science dashboard before you log out of JupyterLab, then logging out of JupyterLab is not successful. For example, if you know the URL for a Jupyter notebook, you are able to open this again in your browser.

**Workaround**

Log out of JupyterLab before you log out of the OpenShift Data Science dashboard.

## RHODS-9789 – Pipeline servers fail to start if they contain a custom database that includes a dash in its database name or username field

When you create a pipeline server that uses a custom database, if the value that you set for the **dbname** field or **username** field includes a dash, the pipeline server fails to start.

**Workaround**

Edit the pipeline server to omit the dash from the affected fields.

## RHODS-9412 – 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.

**RHODS-8921** – 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.

**RHODS-8865** – 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.

**RHODS-7718** – 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.

**RHODS-6907** – 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.

**RHODS-6539** – Anaconda Professional Edition cannot be validated and enabled in OpenShift Data Science

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

**RHODS-6955** – 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

**RHODS-6383** – 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.

RHODS-6373 – 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.

RHODS-6356 – 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.

RHODS-6216 – The ModelMesh 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.

RHODS-5906 – The NVIDIA GPU Operator is incompatible with OpenShift 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.

RHODS-5763 – Incorrect package version displayed during notebook selection

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

RHODS-5543 – 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.

RHODS-5216 – The application launcher menu incorrectly displays a link to OpenShift 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.

**RHODS-5251** – 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.

**RHODS-4799** – 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/"
```

**RHODS-4718** – 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.

**RHODS-4627** – 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.

**RHODS-4502** – The NVIDIA GPU Operator tile 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 tile on the **Explore** page, is therefore redundant. In addition, clicking the **Enable** button moves the NVIDIA GPU Operator tile to the **Enabled** page, even if the Operator is not installed.

**RHODS-3985** – 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.

**RHODS-3984** – 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.

**RHODS-2956 – 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**.

**RHODS-2881 – Actions on dashboard not clearly visible**

The dashboard actions to revalidate a disabled application license and to remove a disabled application tile are not clearly visible to the user. These actions appear when the user clicks on the application tile's **Disabled** label. As a result, the intended workflows might not be clear to the user.

**RHODS-2879 – License revalidation action appears unnecessarily**

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

**RHODS-2650 – 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.

**RHODS-2096 – IBM Watson Studio not available in OpenShift 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.

# CHAPTER 7. PRODUCT FEATURES

Red Hat OpenShift Data Science provides a rich set of features for data scientists and IT operations administrators. To learn more, see Introduction to Red Hat OpenShift Data Science .