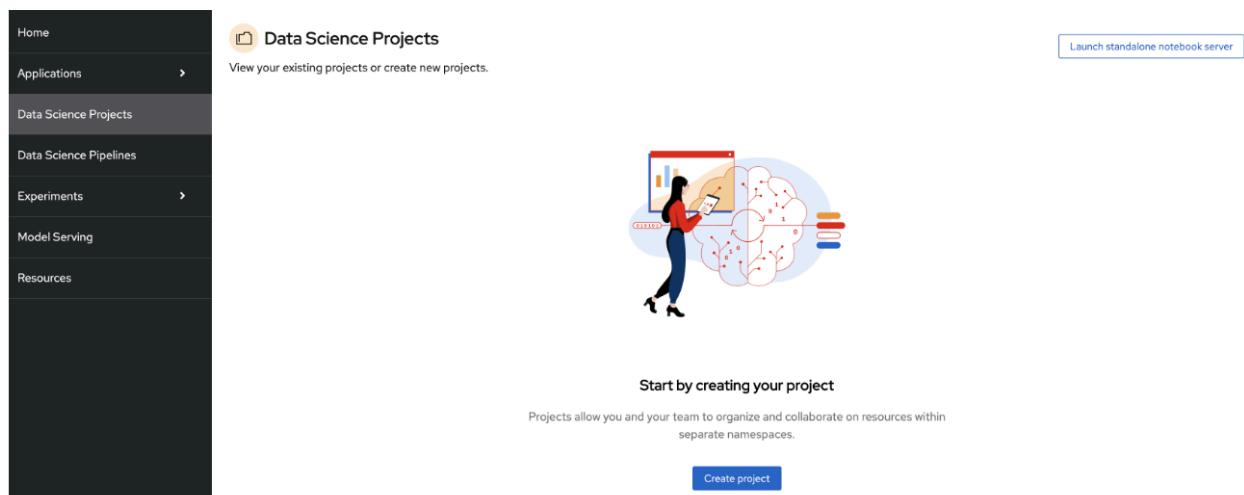


Roey Werner
Consultant

Openshift AI Workshop Short Demo

Setting up your data science project:

On the navigation menu, select **Data Science Projects**. This page lists any existing projects that you have access to. From this page, you can select an existing project (if any) or create a new one.



The screenshot displays the Red Hat OpenShift AI console interface. On the left is a dark navigation sidebar with the following menu items: Home, Applications, Data Science Projects (highlighted), Data Science Pipelines, Experiments, Model Serving, and Resources. The main content area is titled 'Data Science Projects' with a brief description: 'View your existing projects or create new projects.' In the top right corner of the main area is a button labeled 'Launch standalone notebook server'. Below the title is an illustration of a person interacting with a large brain-like network diagram. Further down, the text 'Start by creating your project' is followed by a paragraph: 'Projects allow you and your team to organize and collaborate on resources within separate namespaces.' At the bottom of this section is a blue button labeled 'Create project'.



Create project

Name *

The resource name will be **fraud-detection**.

[Edit resource name](#) ?

Description

Create

[Cancel](#)

Enter a display name and description.

- **Workbenches** are instances of your development and experimentation environment. They typically contain IDEs, such as JupyterLab, RStudio, and Visual Studio Code.
- **Pipelines** contain the data science pipelines that are executed within the project.
- **Models** allow you to quickly serve a trained model for real-time inference. You can have multiple model servers per data science project. One model server can host multiple models.
- **Cluster storage** is a persistent volume that retains the files and data you're working on within a workbench. A workbench has access to one or more cluster storage instances.
- **Data connections** contain configuration parameters that are required to connect to a data source, such as an S3 object bucket.
- **Permissions** define which users and groups can access the project.



Create Workbench:

Click the Workbenches tab, and then click the Create workbench button.

Overview Workbenches Pipelines Models Cluster storage Data connections Permissions



Start by creating a workbench

A workbench is an isolated area where you can work with models in your preferred IDE, such as a Jupyter notebook. You can add accelerators and data connections, create pipelines, and add cluster storage in your workbench.

Create workbench



Name *

Fraud Detection

Description

My Fraud Detection workbench

Fill out the name and description.

Red Hat provides several supported notebook images. In the **Notebook image** section, you can choose one of these images or any custom images that an administrator has set up for you. The **Tensorflow** image has the libraries needed for this tutorial.

Select the latest **Tensorflow** image.



Notebook image

Image selection *

TensorFlow

Version selection *

2024.1

i A new image version is available. Select the latest image version to use Elyra for pipelines.
Hover an option to learn more information about the packages included.

Choose a small deployment.

Deployment size

Container size

Small

Leave the default environment variables no GPU and storage options



Cluster storage

 Cluster storage will mount to /

☒ Create new persistent storage

This creates storage that is retained when logged out.

Name *

Fraud Detection

Description

Persistent storage size

-

20

+

 GiB

☐ Use existing persistent storage

This reuses a previously created persistent storage.

Under **Data connections**, select **Create new data connection**



Add data connection



Name *

MyStorage

Access key *

youraccesskey

Secret key *

.....



Endpoint *

https://storage-endpoint:storage-cluster.com:1234

Region

us-east-1

Bucket

my-storage

Connected workbench



Add data connection

Cancel

Click Create.

Verify That the notebook is running.



Data Science Projects > Fraud Detection

Fraud Detection

My Fraud Detection project

Overview Workbenches Pipelines Models Cluster storage Data connections Permissions



Workbenches

Create workbench

Name	Notebook image	Container size	Status	
> Fraud Detection My Fraud Detection Workbench	Tensor Flow	Small	<input checked="" type="checkbox"/> Running	Open

Note

If you made a mistake, you can edit the workbench to make changes.

Status



Running

Open



Edit workbench

Importing files from git:

The Jupyter environment is a web-based environment, but everything you do inside it happens on **Red Hat OpenShift AI** and is powered by the **OpenShift** cluster. This means that, without having to install and maintain anything on your own computer, and without disposing of valuable local resources such as CPU, GPU and RAM, you can conduct your Data Science work in this powerful and stable managed environment.

Procedure

1. Click the **Open** link next to your workbench. If prompted, log in and allow the Notebook to authorize your user.

Data Science Projects > Fraud Detection

Fraud Detection



My Fraud Detection project

Overview Workbenches Pipelines Models Cluster storage Data connections Permissions

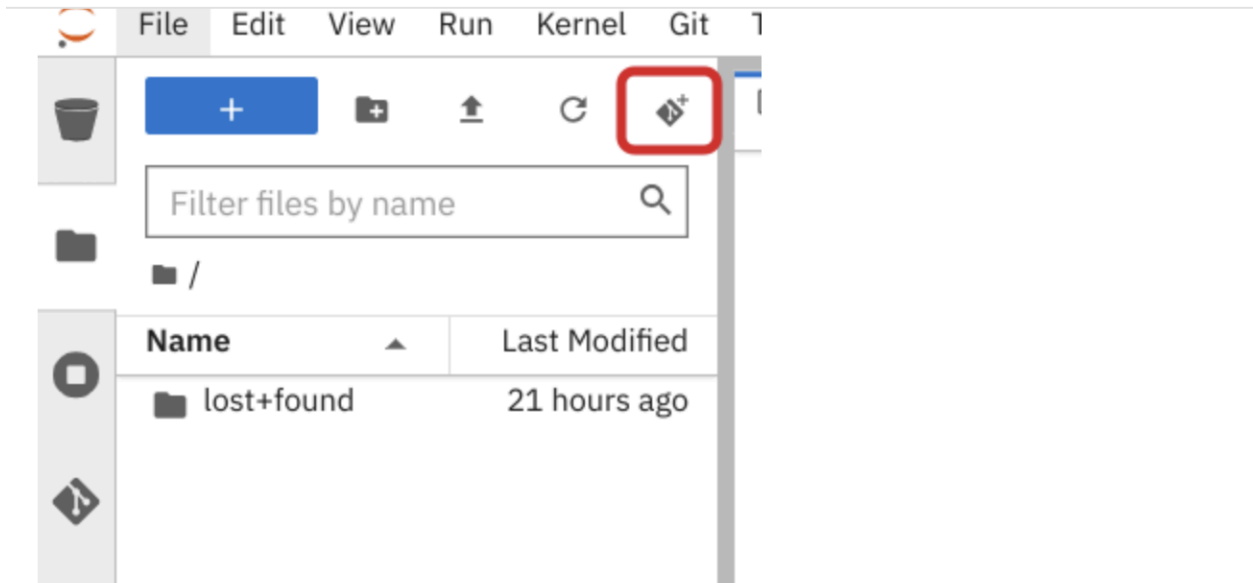


Workbenches ?

Create workbench

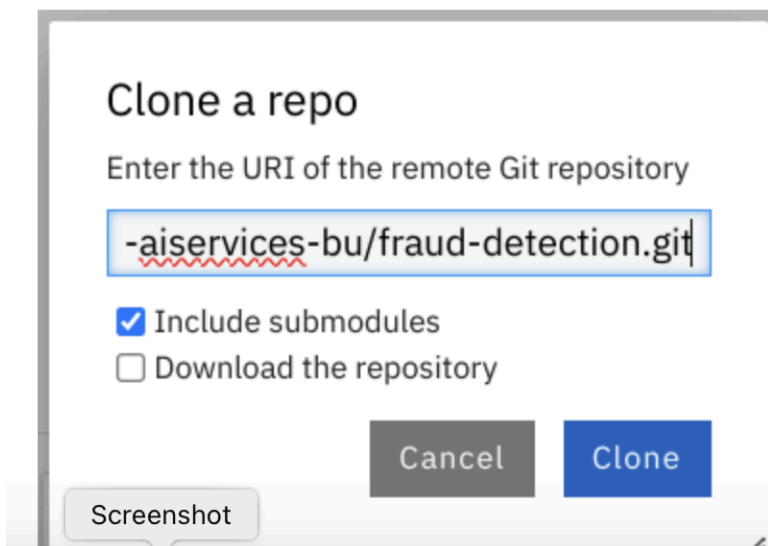
Name ↑	Notebook image	Container size	Status ↑	
> Fraud Detection ? My Fraud Detection Workbench	Tensor Flow	Small	 Running	Open 

1. Your Jupyter environment window opens.
This file-browser window shows the files and folders that are saved inside your own personal space in OpenShift AI.
2. Bring the content of this tutorial inside your Jupyter environment:
 - i. On the toolbar, click the **Git Clone** icon:



ii. Enter the following tutorial Git **https** URL:

`https://github.com/rh-aisservices-bu/fraud-detection.git`



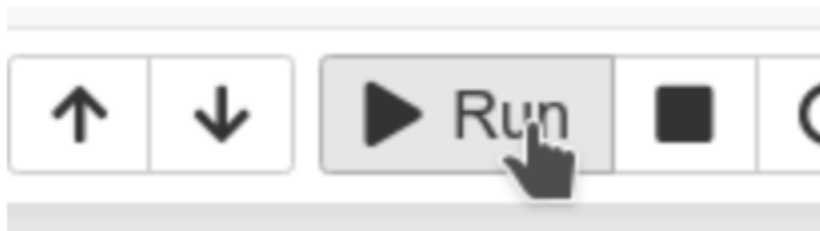
<https://github.com/rh-aisservices-bu/fraud-detection.git>

- i. Check the **Include submodules** option, and then click **Clone**.
- ii. In the file browser, double-click the newly-created **fraud-detection** folder.

Running code in a notebook:

You can run a code cell from the notebook interface or from the keyboard:

- **From the user interface:** Select the cell (by clicking inside the cell or to the left side of the cell) and then click **Run** from the toolbar.



- **From the keyboard:** Press **CTRL + ENTER** to run a cell or press **SHIFT + ENTER** to run the cell and automatically select the next one.

Try it:

Procedure

1. In your Jupyter environment, locate the `0_sandbox.ipynb` file and double-click it to launch the notebook. The notebook opens in a new tab in the content section of the environment.



2. Experiment by, for example, running the existing cells, adding more cells and creating functions.
You can do what you want - it's your environment and there is no risk of breaking anything or impacting other users. This environment isolation is also a great advantage brought by OpenShift AI.
3. Optionally, create a new notebook in which the code cells are run by using a Python 3 kernel:
 - a. Create a new notebook by either selecting **File** → **New** → **Notebook** or by clicking the Python 3 tile in the Notebook section of the launcher window:

Object storage integration:

In this section we will learn how to list, upload and download files from the jupyter notebook:

- Clone the repo: <https://github.com/Roeywer/ocp-ai.git>
- Open bucket-connection-demo.ipynb file and run it to learn about S3 bucket integration.