

## Setting Up Model Registry

Date published: 2020-07-16

Date modified: 2025-05-29



# Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

# Contents

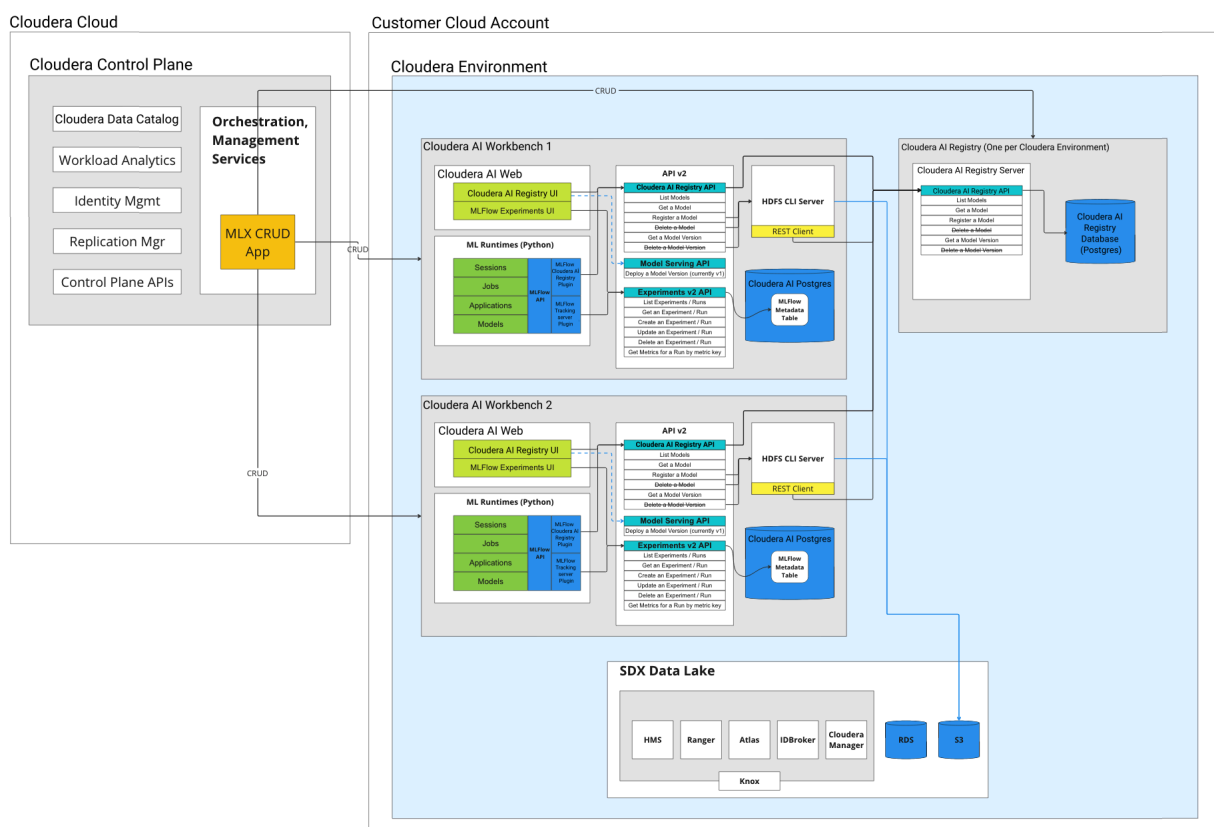
- Setting up Cloudera AI Registry..... 4**
  - Creating a Cloudera AI Registry..... 5
  - Creating a Cloudera AI Registry on an Azure UDR Private Cluster..... 6
  - Setting up access for Cloudera AI Registry in a RAZ-enabled environment..... 7
  - Setting up access for Cloudera AI Registry in a non-RAZ-enabled environment..... 9
  - Synchronizing Cloudera AI Registry with a workbench..... 10
  - Viewing details for Cloudera AI Registries..... 11
  - Cloudera AI Registry permissions..... 12
  - Model access control..... 12

## Setting up Cloudera AI Registry

Cloudera AI Registry is the core enabler for MLOps, or DevOps for machine learning.

Cloudera AI Registry stores and manages machine learning models and associated metadata, such as the model's version, dependencies, and performance. The registry enables MLOps and facilitates the development, deployment, and maintenance of machine learning models in a production environment.

Cloudera AI Registry in Public Cloud



Cloudera AI Registry includes functionality for the following tasks:

- Storing and organizing different versions of a machine learning model and its associated metadata.
- Tracking the lineage of a model, including who created it, when it was created, and any changes made to it over time.
- Providing APIs for accessing and deploying models, as well as for querying and searching the registry.
- Integrating with CI/CD pipelines and other tools used in the MLOps workflow.

Cloudera AI Registry instances help organizations improve the quality and reliability of their machine learning models by providing a centralized location for storing and managing models, as well as enabling traceability and reproducibility of model development. They also make deploying and managing models in a production environment easier by providing a single source for model versions and dependencies.

The Cloudera AI Registry integrates MLFlow and maintains compatibility with the open source ecosystem.

### Limitations

- Upgrade to the General Availability (GA) version of Cloudera AI Registry might not be supported. Alternatively, upgrade to the GA version of Cloudera AI Registry might require reinstalling Cloudera AI Registry which could

result in loss of Cloudera AI Registry data configured with the technical preview (TP) version of Cloudera AI Registry.

## Creating a Cloudera AI Registry

Before you can start using Cloudera AI Registry you must create a AI registry for your environment.


### Procedure

1. In the **Cloudera** console, click the **Cloudera AI** tile.  
The **Cloudera AI Workbenches** page displays.
2. Click AI Registries in the left navigation pane.
3. Click Create AI Registry.  
Cloudera AI displays the Create AI Registry dialog box.
4. Choose your environment from the Environment Name drop down list.
5. Depending on your environment, complete one of the following:
  - a) If your environment is in AWS, **AI Registry** displays the following dialog box:

### Create AI Registry

✕

\* Environment Name ⓘ

 a' -2)

▼

☐ Enable Public IP Address for Load Balancer

Subnets for Load Balancer ⓘ

Cancel

Create

1. From the Environment Name dropdown list, select your environment.
2. Enable Public IP Address for Load Balancer: By default, AI Registry service uses a private load balancer for cluster ingress. If you use a public load balancer instead, select the Enable Public IP Address for Load

Balancer option. If you use a private load balancer for cluster ingress, you must have a VPN connection between your corporate network and the Virtual Private Cloud (VPC) in which the AI Registry is deployed.

3. Subnets for Load Balancer: If your public subnets are protected by a firewall, select the specific subnets to be used by the load balancer.
  4. Click Create to create the Cloudera AI Registry.
- b) If your environment is in Azure, Cloudera AI Registry displays the following dialog box:

## Create AI Registry

✕

---

**\* Environment Name** ⓘ

▲
dr [redacted] is)
▼

☐ Enable Public IP Address for Load Balancer

Cancel

Create

1. Choose the Azure environment for the Cloudera AI Registry.
2. Enable Public IP Address for Load Balancer: By default, AI Registry service uses a private load balancer for cluster ingress. If you use a public load balancer instead, select the Enable Public IP Address for Load Balancer option. If you use a private load balancer for cluster ingress, you must have a VPN connection between your corporate network and the Virtual Private Cloud (VPC) in which the AI Registry is deployed.
3. Click Create to create the Cloudera AI Registry.

## Creating a Cloudera AI Registry on an Azure UDR Private Cluster

Use the following template Cloudera CLI command to create a UDR private cluster on Azure with a Cloudera AI Registry.

You must replace the following template items with your own information.

- <environment CRN>
- <environment name> (in two places)
- <subnet>

Model registries are also supported on Azure private clusters with UDR. For more information about UDR, see the [Preview Feature](#) documentation.

If you have not yet downloaded the Cloudera CLI tool, see the [documentation](#).

Use the latest version of the Cloudera CLI.

### Cloudera CLI command to create a Cloudera AI Registry

This Cloudera CLI command performs has three key sections:

1. Enables support for private clusters in Azure ( "privateCluster": true, )

2. Enables UDR for the private cluster ("outboundTypes": ["OUTBOUND\_TYPE\_UDR"],)
3. Specifies the subnet for the UDR-enabled private cluster ("subnets")

```
cdp ml create-model-registry --cli-input-json {
  "environmentCrn": "<environment CRN>",
  "environmentName": "<environment name>",
  "privateCluster": true,
  "usePublicLoadBalancer": false,
  "outboundTypes": [
    "OUTBOUND_TYPE_UDR"
  ],
  "provisionK8sRequest": {
    "network": {
      "topology": {
        "subnets": [
          "<subnet>" # subnet with a default route configuration to
forward the traffic to the network appliance or firewall. This is required t
o enable UDR.
        ]
      }
    }
  }
}
```

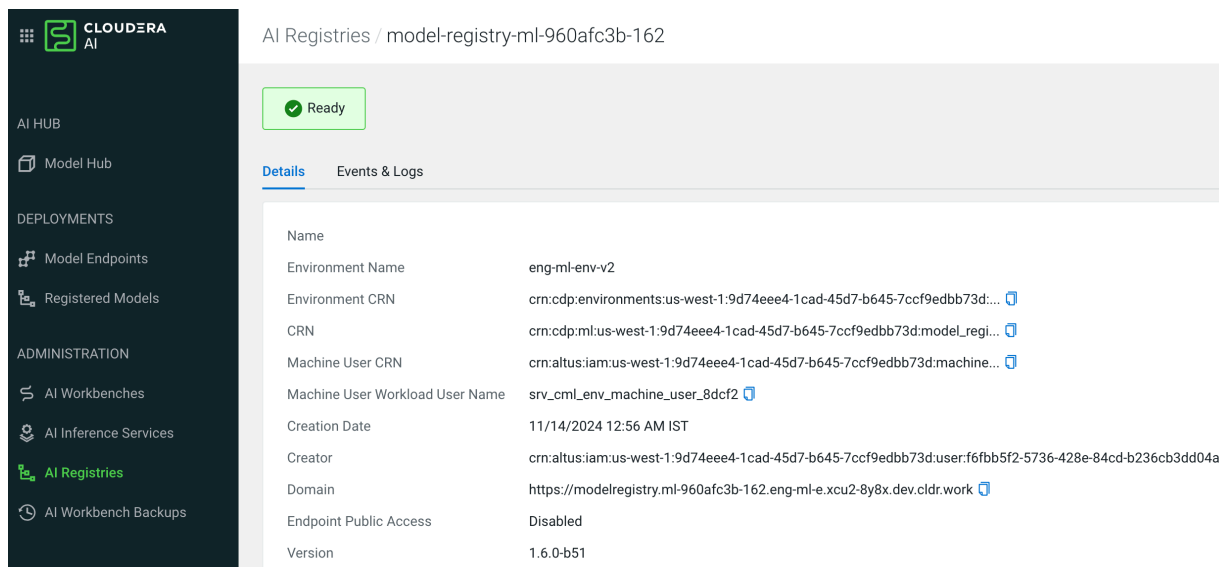
## Setting up access for Cloudera AI Registry in a RAZ-enabled environment

In a [RAZ-enabled environment](#) you need to set up the S3-Ranger policy by manually adding the machine user name in the S3 Ranger policy.

To set up the S3-Ranger policy, complete the following:

1. On the AI Registries Details page, find and copy the Machine User Workload User Name in the Machine User Workload User Name field.

For example, in the following screenshot, the Machine User Workload User Name field contains `srv_cml_env_machine_user_8dcf2`. Copy the Machine User Workload User Name which is `8dcf2`.



AI Registries / model-registry-ml-960afc3b-162

Ready

[Details](#) [Events & Logs](#)

Name	
Environment Name	eng-ml-env-v2
Environment CRN	crn:cdp:environments:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:...
CRN	crn:cdp:ml:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:model_regi...
Machine User CRN	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:machine...
Machine User Workload User Name	srv_cml_env_machine_user_8dcf2
Creation Date	11/14/2024 12:56 AM IST
Creator	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:user:f6fbb5f2-5736-428e-84cd-b236cb3dd04a
Domain	https://modelregistry.ml-960afc3b-162.eng-ml-e.xcu2-8y8x.dev.cldr.work
Endpoint Public Access	Disabled
Version	1.6.0-b51

- Depending on your environment, select `cm_s3` (AWS) or `cm_adls` (Azure).

The screenshot shows the Ranger web interface. The top navigation bar includes the Ranger logo and tabs for 'Access Manager', 'Audit', 'Security Zone', and 'Settings'. The 'Service Manager' tab is selected. The main content area displays a grid of service managers. Each service manager is represented by a folder icon, a name, and a list of components. Each component has a status icon (eye, lock, or red square). The services listed are: cm\_hdfs, YARN (cm\_yarn), KAFKA (cm\_kafka), ATLAS (cm\_atlas), OZONE (cm\_ozone), S3 (cm\_s3), cm\_hbase, KNOX (cm\_knox), NIFI, ADLS (cm\_adls), SCHEMA-REGISTRY, Hadoop SQL, SOLR (cm\_solr), NIFI-REGISTRY, KUDU (cm\_kudu), and KAFKA-CONNECT (cm\_kafka\_connect). A 'Last Response Time' is shown as 03/07/2023 03:46:15 AM.

- | List of Policies : cm_s3   |                    |               |         |               |       |                          |   |   |
|--|--------------------|---------------|---------|---------------|-------|--------------------------|---|---|
| <div> <input type="text" value="Search for your policy..."/> <span>?</span> <span>Add New Policy</span> </div> |                    |               |         |               |       |                          |   |   |
| Policy ID ▾  | Policy Name        | Policy Labels | Status  | Audit Logging | Roles | Groups                   | Users                                       | Action  |
| 69   | all - bucket, path | --            | Enabled | Enabled       | --    | c_ranger_admins_69afd1d8 | rangerraz<br>srv_cm1_env_machine_user_70378 | <input type="button" value="View"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/> |



5. Enter the Machine User Workload User Name in the Select User field in the allow conditions section.

For example, using the Machine User Workload User Name from Step 2, add the value which is 82a49.

## Setting up access for Cloudera AI Registry in a non-RAZ-enabled environment

In a non-RAZ-enabled environment you need to add the Machine User CRN to the IDBroker mapping in order to access the S3/ADLS buckets.

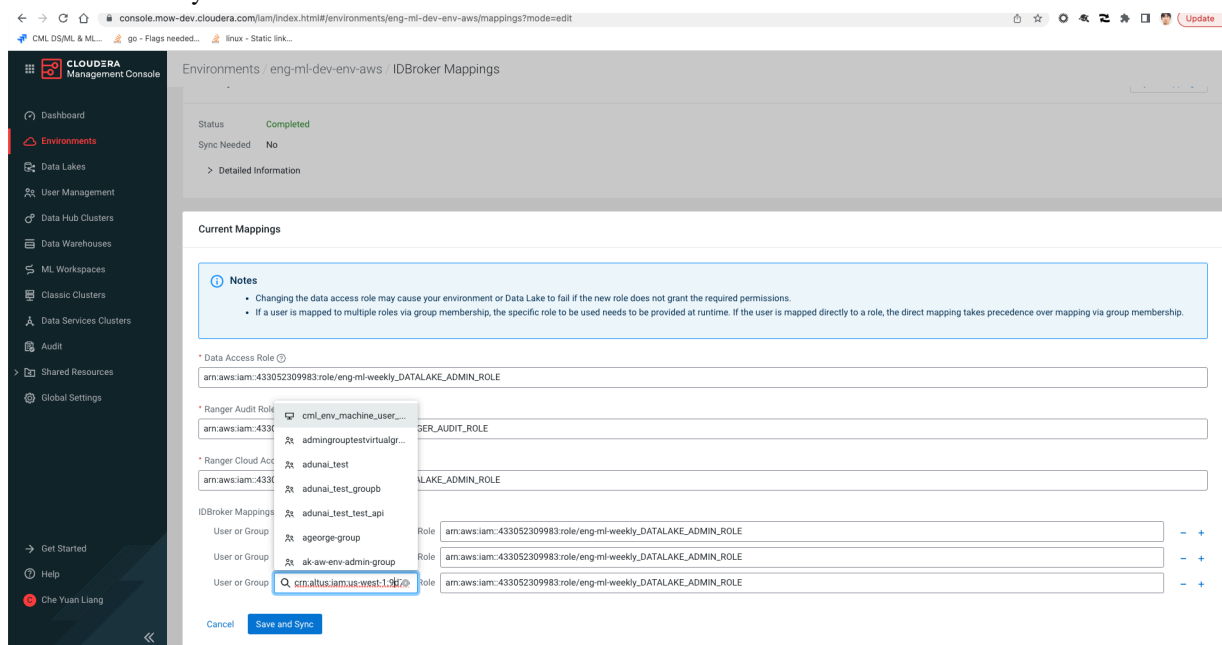
To add the Machine User CRN to the IDBroker mapping complete the following:

1. Locate the Machine User CRN in the AI Registries Details page.

AI Registries / model-registry-ml-960afc3b-162	
Status	Ready
Details	Events & Logs
Name	
Environment Name	eng-ml-env-v2
Environment CRN	crn:cdp:environments:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:...
CRN	crn:cdp:ml:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:model_regi...
Machine User CRN	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:machine...
Machine User Workload User Name	srv_cml_env_machine_user_8dcf2
Creation Date	11/14/2024 12:56 AM IST
Creator	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:user:f6fbb5f2-5736-428e-84cd-b236cb3dd04a
Domain	https://modelregistry.ml-960afc3b-162.eng-ml-e.xcu2-8y8x.dev.cldr.work
Endpoint Public Access	Disabled
Version	1.6.0-b51

2. Copy the Machine User CRN mapping. When retrieving the user from the model registry, do not include the srv\_ prefix.
3. Navigate to the Environment Manage access idbroker page and add or choose the Machine User CRN mapping to the Data Access Role field.

#### 4. Click Save and Sync.



## Synchronizing Cloudera AI Registry with a workbench

If you deploy a Cloudera AI Registry in an environment that contains one or more Cloudera AI Workbench, you must synchronize Cloudera AI Registry with the workbenches.



**Important:** If your Cloudera AI Workbench version is 2.0.46 or higher, or Cloudera AI Inference service version is 1.2.0 or higher, and if you deploy a Cloudera AI Registry in an environment that contains one or more Cloudera AI Workbenches, the Cloudera AI Registry is auto-discovered and periodically synchronized by Cloudera AI Inference service and Cloudera AI Workbenches and no manual synchronization is required.

Cloudera AI Workbench is auto-synchronized every five minutes and Cloudera AI Inference service is auto-synchronized every 30 seconds.

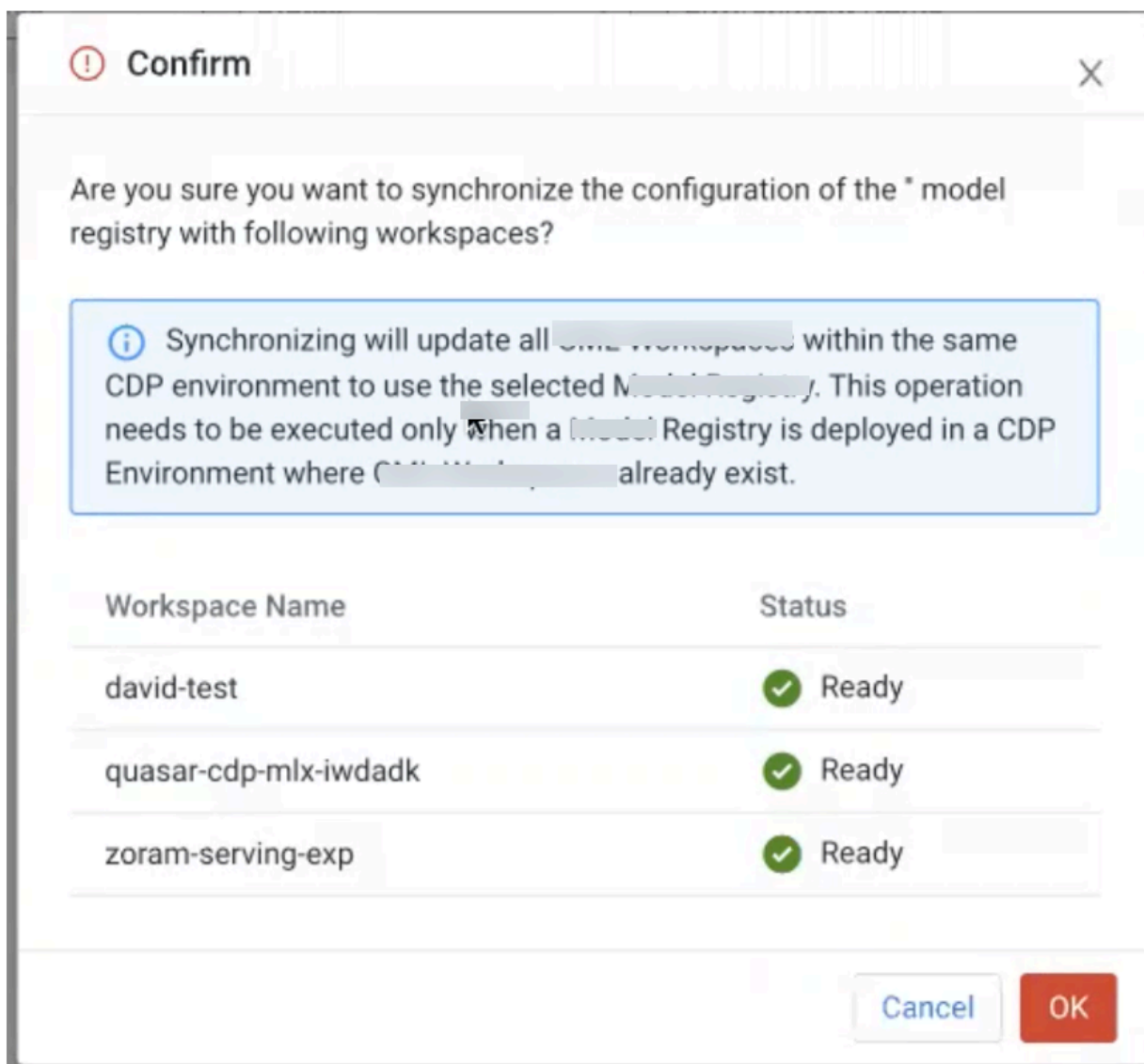
The outbound network access on [AWS](#) and [Azure](#) must be configured correctly for the auto synchronization to work.

### Procedure

1. Click AI Registries to display the AI Registries window.
2. Choose the Cloudera AI Registry you want to synchronize with the workbenches in the environment.

- From the Actions menu, click Synchronize.

AI Registries displays the Confirm dialog box listing all of the workbenches in the environment.



- Click OK.

## Viewing details for Cloudera AI Registries

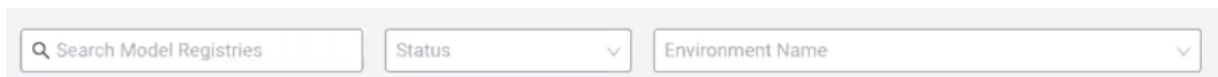
You can view the information for your registered models using **AI Registries**.

### Procedure

- In the Cloudera Console, click the Cloudera AI tile.  
The **Cloudera AI Workbenches** page displays.
- Select AI Registries from the left navigation pane.

On the main AI Registries page, you can see all the models currently registered, their respective owners, location of creation, and the last updated time, if known.

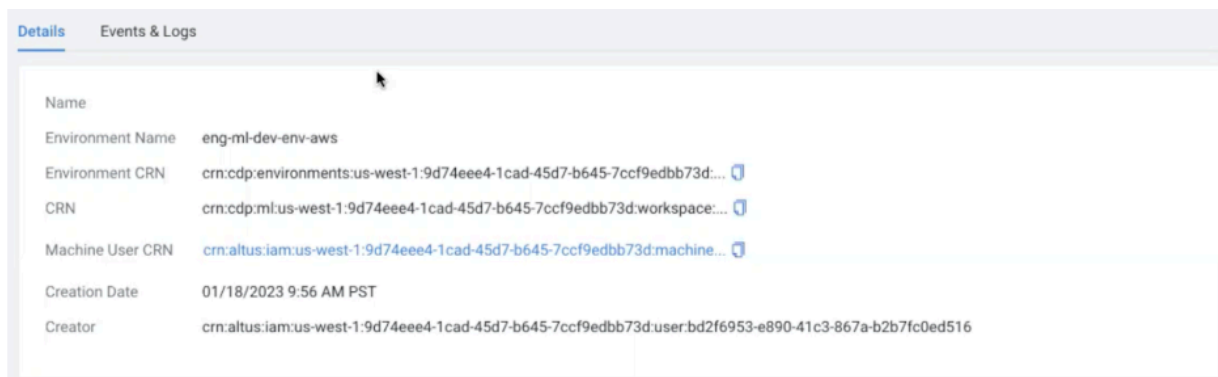
- Use the filter bar at the top of the window to filter the list of AI registries by name, status, and environment name.



The filter bar consists of three components: a search input field labeled 'Search Model Registries' with a magnifying glass icon, a status dropdown menu labeled 'Status', and an environment name dropdown menu labeled 'Environment Name'.

- Select a AI Registry to see its description.

Cloudera AI displays the Details page which lists the environment name, environment CRN, CRN, machine user CRN, creator, and creation date.



The Details page shows the following information:

Details	
Name	
Environment Name	eng-ml-dev-env-aws
Environment CRN	crn:cdp:environments:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:...
CRN	crn:cdp:ml:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:workspace:...
Machine User CRN	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:machine:...
Creation Date	01/18/2023 9:56 AM PST
Creator	crn:altus:iam:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:user.bd2f6953-e890-41c3-867a-b2b7fc0ed516

- You can also click the Events & Logs tab to display information on the events and logs for the Cloudera AI Registry.

## Cloudera AI Registry permissions

Cloudera AI Registry's permissions for the following actions are separate from workbench permissions, but they are inherited from environment level workbench permissions.

- create
- delete
- getKubeconfig
- grant/list/revoke access

Therefore, if you have the MLAdmin role on an environment, you can perform these actions for Cloudera AI Registry instances, but an MLUser cannot.

Remote access to a Cloudera AI Registry works similarly to workbench remote access. In addition to AI Registry, but an MLUser cannot.

Remote access to a AI Registry works similarly to workbench remote access. In addition to downloading the `kubeconfig` file, you need to use `Grant/List/RevokeModelRegistryAccess` endpoints to manage what cloud user identity can access the Kubernetes cluster using your cloud credential.

## Model access control

Access to models is dependent on the user permissions, as described here.

- Only administrators are able to see all models, including those not created by them.
- A user, who is an owner of a given model, can set the visibility of the model to public or private on the **AI Registries UI**.
- When a user registers a new model, the user becomes the owner, and no other user has access.
- A user can make the visibility of a model public, and then all users receive viewer access to the model by default.