CDP Public Cloud / Data Catalog

Data Sharing

Date published: 2019-11-14
Date modified: 2026-03-31



https://docs.cloudera.com/

Legal Notice

© Cloudera Inc. 2026. 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

Creating external users.....4

Creating a new Data Share.....6

Modifying Data Shares.....15

Revoking user credentials.....25

Regenerating user credentials.....27

Deleting a Data Share.....29

Verifying external access to a Data Share.....33

Creating external users

Register external users in Cloudera Data Catalog so that they can be granted access to Data Shares as Data Consumers.

About this task

External users are users or applications that are not part of the Cloudera ecosystem but need to consume data from Cloudera through Data Shares. Before a user can be added to a Data Share, they must be registered in the Manage Users page. Registration generates a credentials file (Client ID and Secret) that the user needs to authenticate with their Iceberg REST Catalog compatible compute engine.

Before you begin

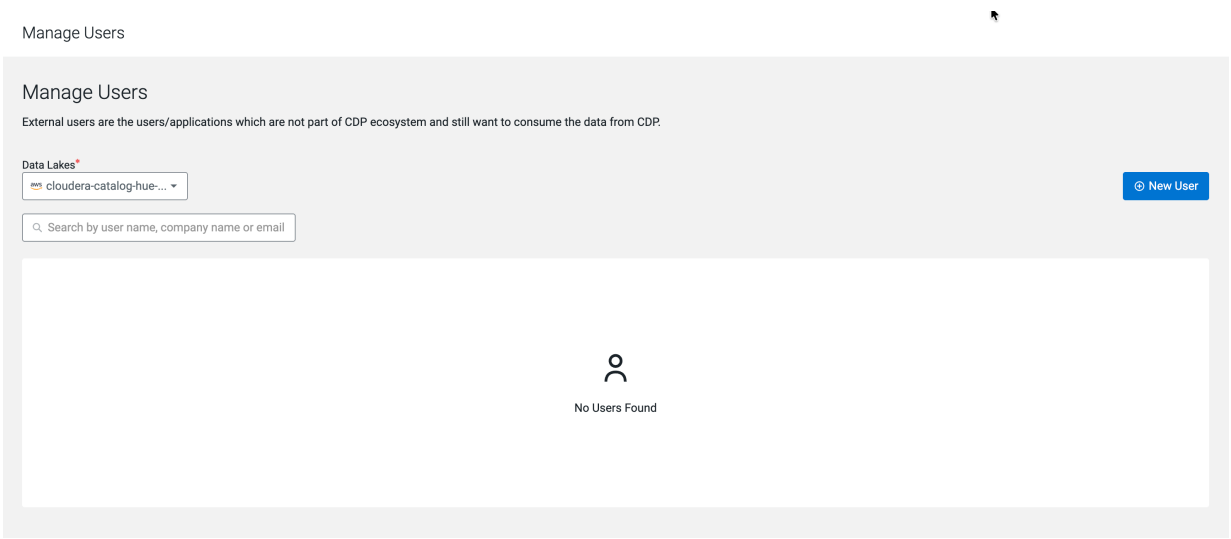
You must have the DataShareAdmin resource role for the relevant environment.

Procedure

1. In Cloudera Data Catalog, navigate to Manage Users.
2. Select the target Data Lake from the dropdown at the top of the page.

   The user list shows all existing external users for the selected Data Lake, including their Client ID, email address, company name, associated shares, and registration date.

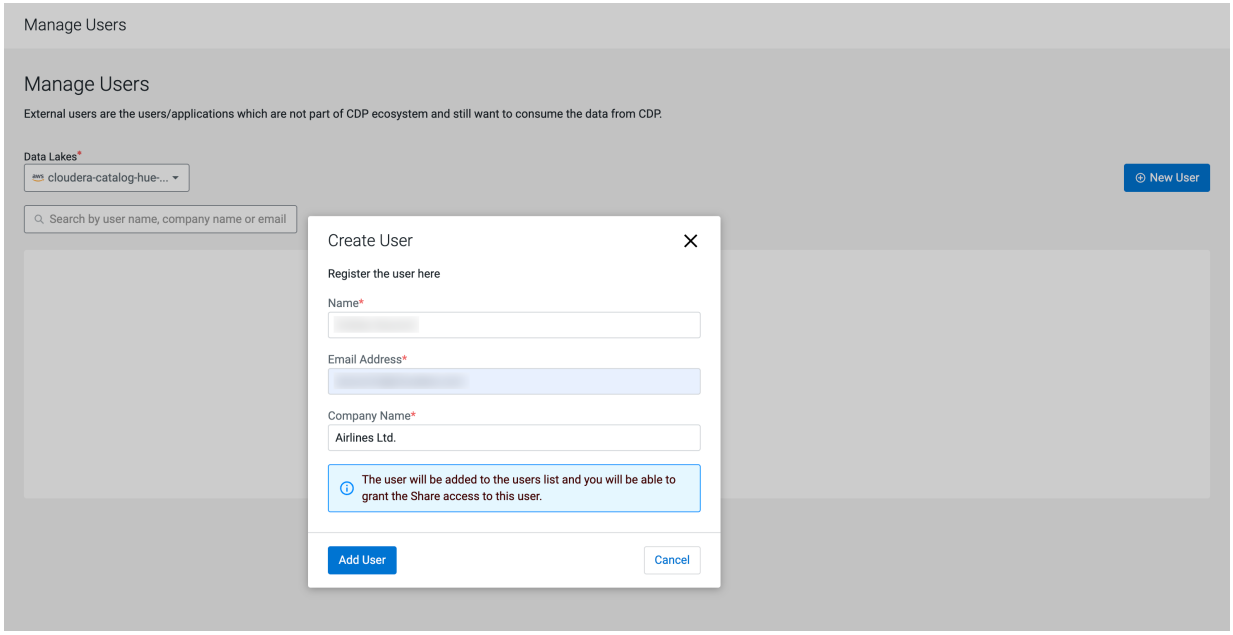
   Figure 1: Manage Users page



3. Click New User in the top right corner.

   The Create User dialog opens.

   Figure 2: Create User dialog



4. Fill in the required fields: Name, Email Address, and Company Name.
5. Click Add User.

Note:

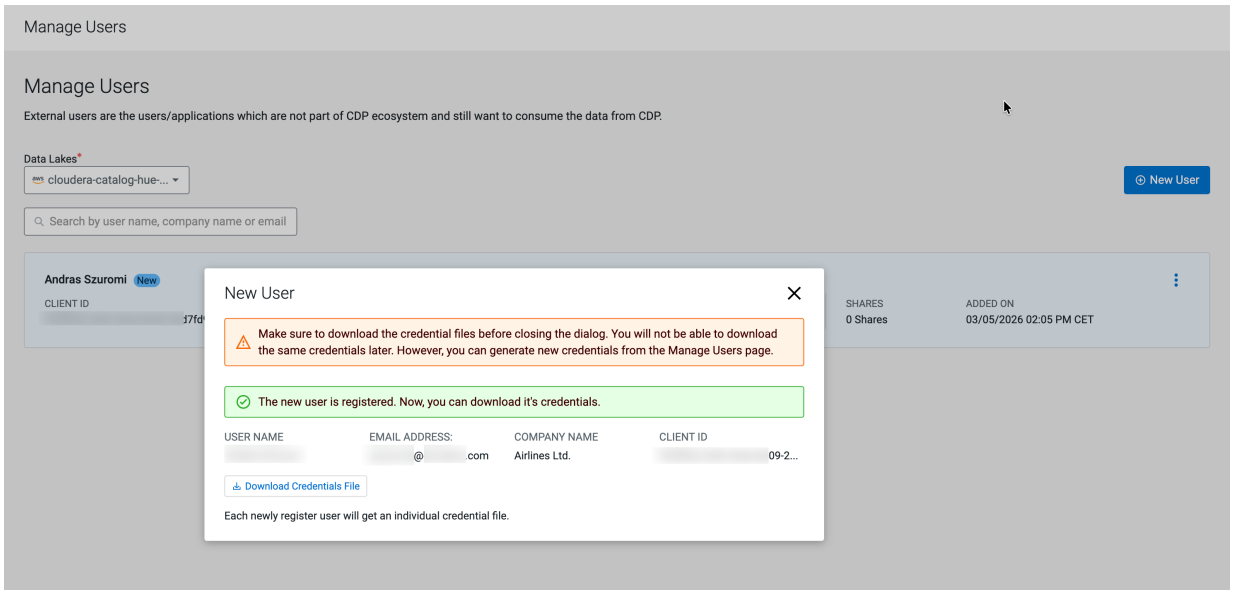
- You can also create external users with the cdp datacatalog create-external-users CDP CLI command.
- You can also register a new user during the Data Share creation.

A success message confirms that the new user is registered and their credentials are ready to download.

6. Click Download Credentials File to save the credentials before closing the dialog.

Important: Download the credentials file before closing this dialog. The same credentials cannot be downloaded again. If needed, you can generate new credentials later from the Manage Users page.

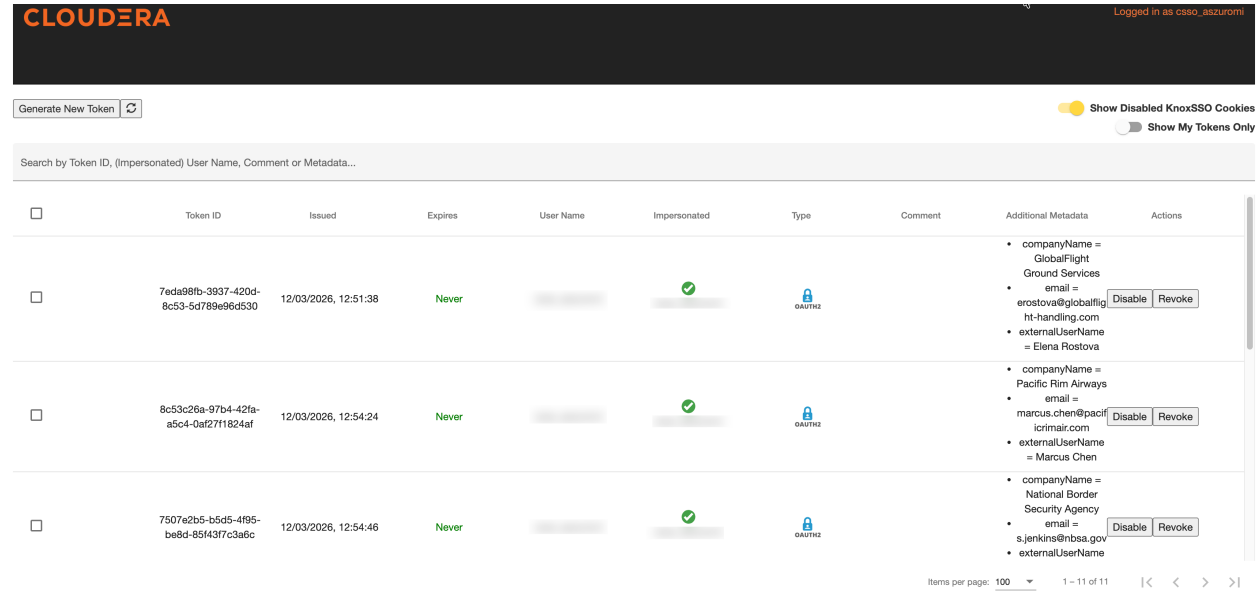
Figure 3: New User credentials download dialog



Results

The external user is registered and appears in the Manage Users list. You can now add this user to a Data Share.

External user Client IDs are also visible in Knox under Environment *[***YOUR_DATALAKE***]* Token Integration Token Management



What to do next

Securely share this downloaded credential file (for example, via secure email) to the external consumer so they can authenticate their compute engine by exchanging their Client ID and Secret for a Knox Access token.

Important: knox.token.ttl controls the lifetime of your Knox access token received in exchange for the Client ID and Secret of your external users. Once it expires, external users need to request a new access token. For more information, see Editing Knox topologies.

Related Information

Editing Knox topologies

create-external-users

list-external-users

Creating a new Data Share

Create and publish a Data Share in Cloudera Data Catalog to grant external users read-only access to Iceberg tables.

About this task

A Data Share is a logical unit that groups one or more Iceberg tables and grants specific external users read-only access to them for a defined period. The New Share wizard guides you through six steps: **Share Name**, **Assets**, **Users**, **Terms and Conditions**, **Additional Information**, and **Review**. All steps can be revisited after creation, you can also populate and customize a Share after initially creating it.

Before you begin

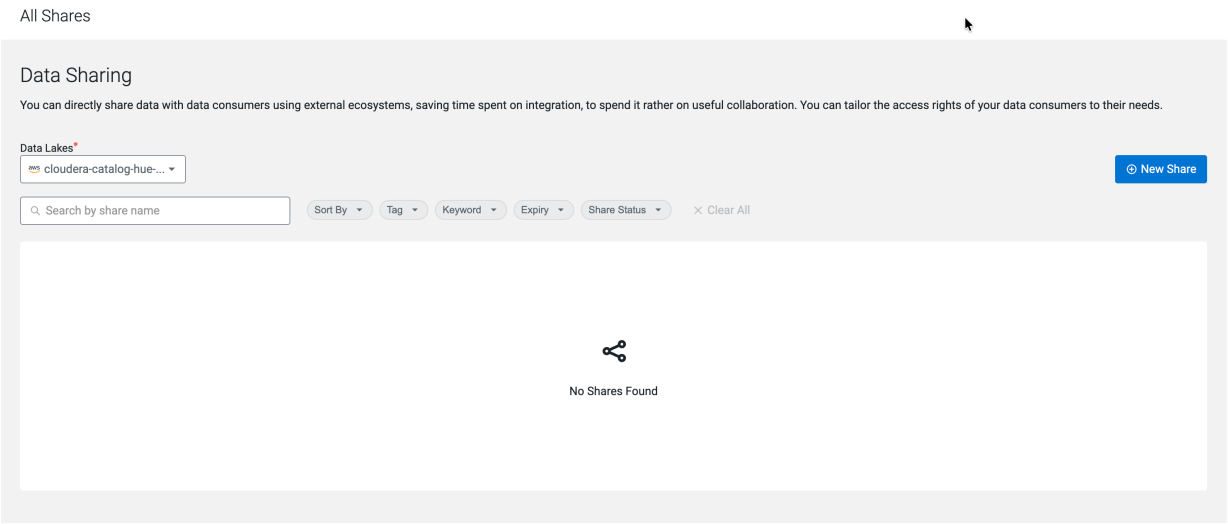
- You must have the DataShareAdmin resource role for the relevant environment.
- External users who will be granted access must already be registered on the Manage Users page. For more information, see Creating external users.

Procedure

1. In Cloudera Data Catalog, go to Data Sharing and click New Share.

   Select the target Data Lake from the dropdown if it is not already selected. The New Share wizard opens.

   Figure 4: All Shares page



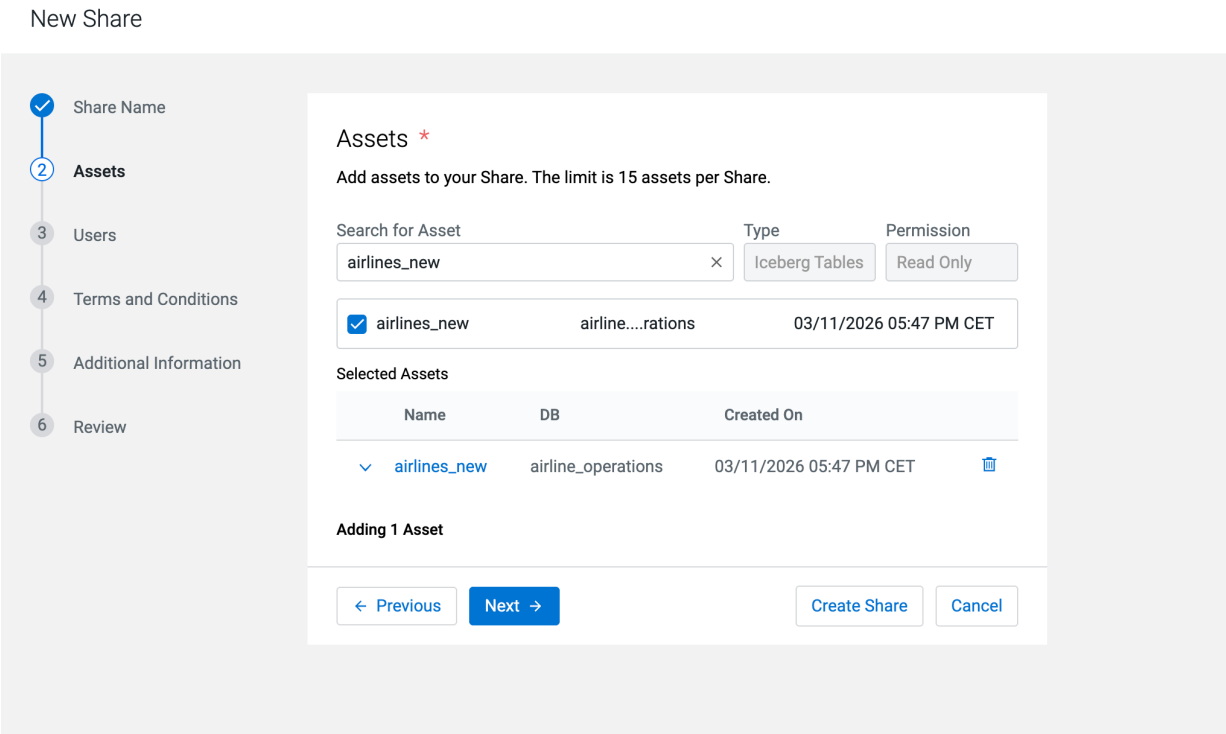
2. On the Share Name step, enter a name for the Data Share and click Next.
3. On the Assets step, search for and select the Iceberg tables to include in the Data Share, then click Next.

   Type a name in the Search for Asset field and check the checkbox next to the asset to add it to the Selected Assets list.

   Note:

   - Currently, the Type is fixed at Iceberg Tables and Permission is fixed at Read Only.
   - You can add up to 15 assets per Data Share.
   - If a recently created table does not appear in the search results, it might not be indexed yet. The asset list is populated by Atlas, which can take a few minutes to index new tables after the Impala or Hive CREATE TABLE command completes. Wait a moment, refresh the page, and search again.

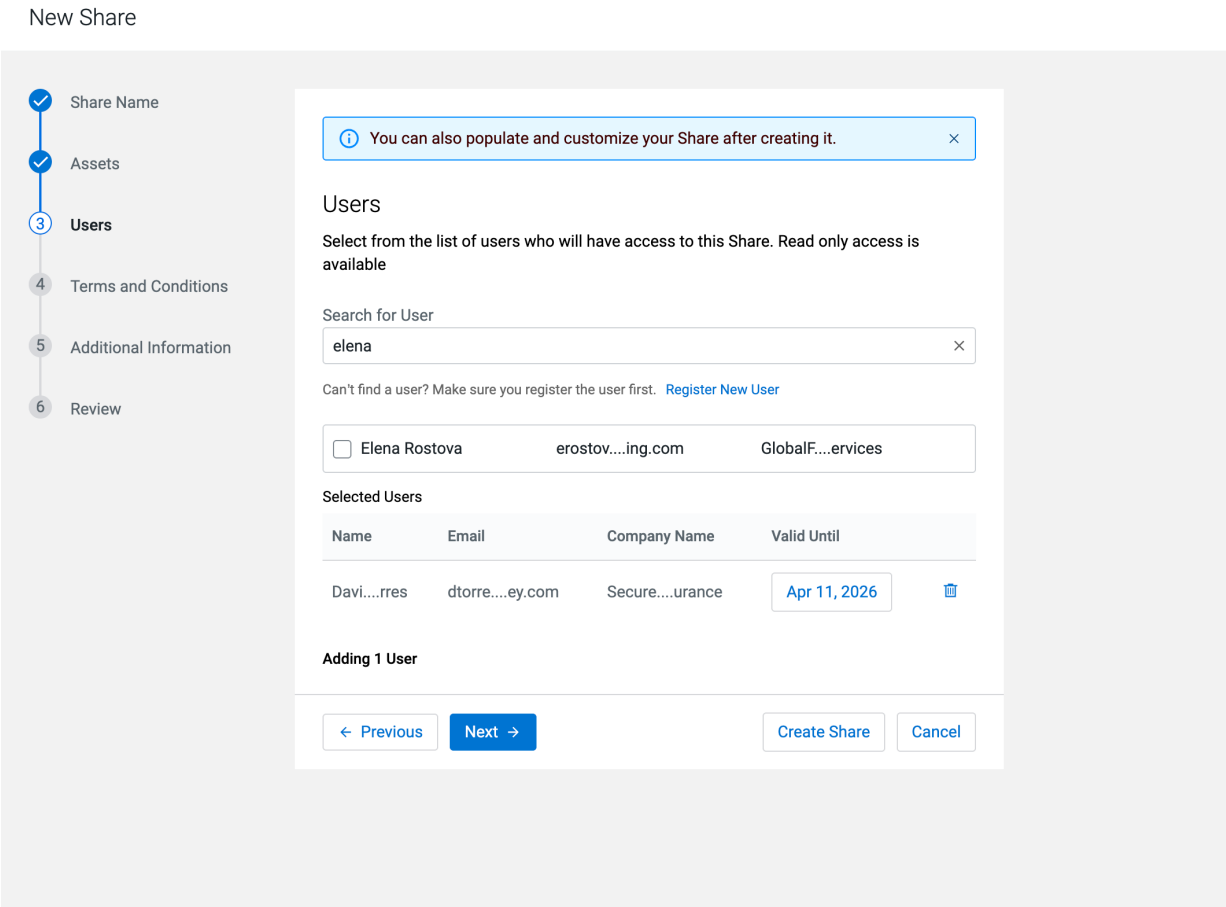
   Figure 5: Asset selection



4. On the Users step, search for and select the external users who will have access to this Data Share, then click Next.

   Type a name in the Search for User field and check the checkbox next to the user to add them to Selected Users. The Valid Until date defaults to 30 days from creation and can be adjusted per user by clicking the date. If the required user is not listed, click Register New User to register them first.

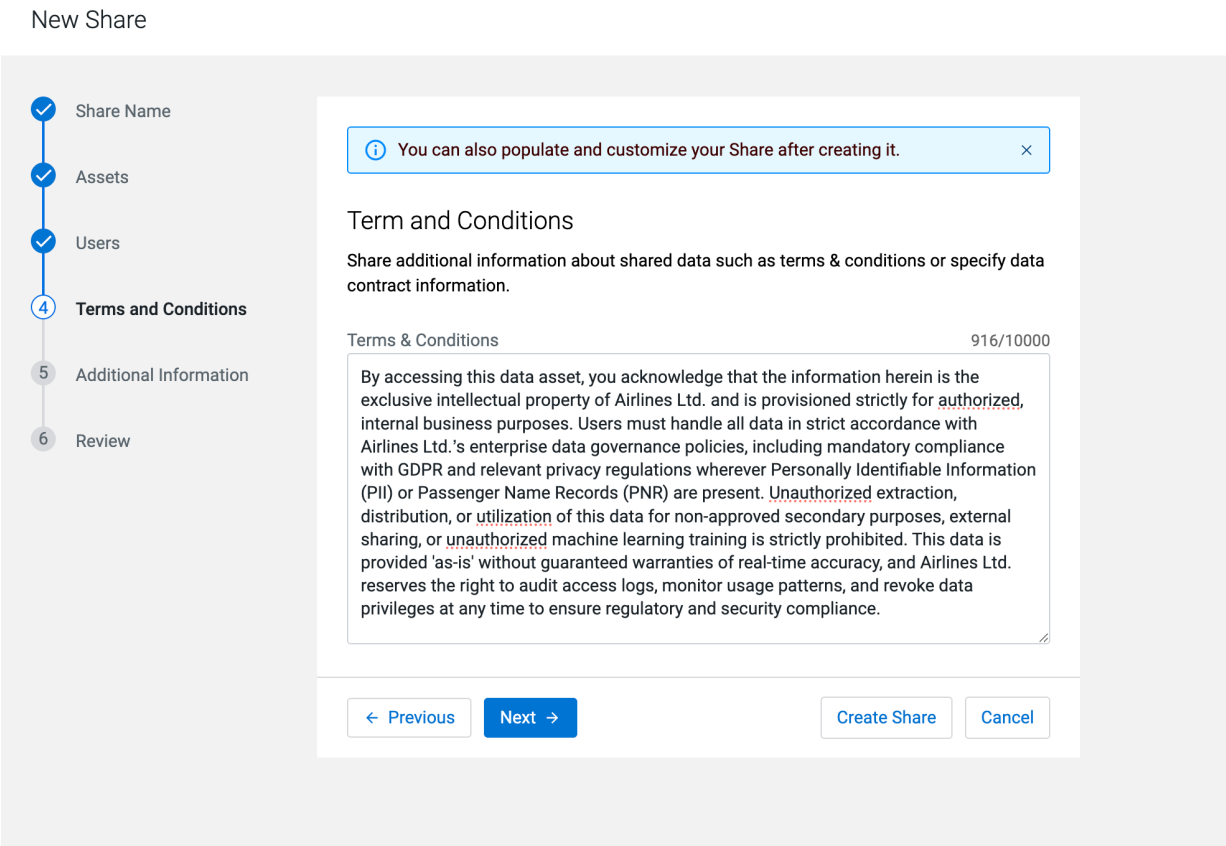
Figure 6: User selection



5. On the Terms and Conditions step, optionally enter a terms and conditions text, then click Next.

   Use this field to communicate data usage obligations, compliance requirements, or contractual constraints to Data Consumers. The field accepts up to 10,000 characters.

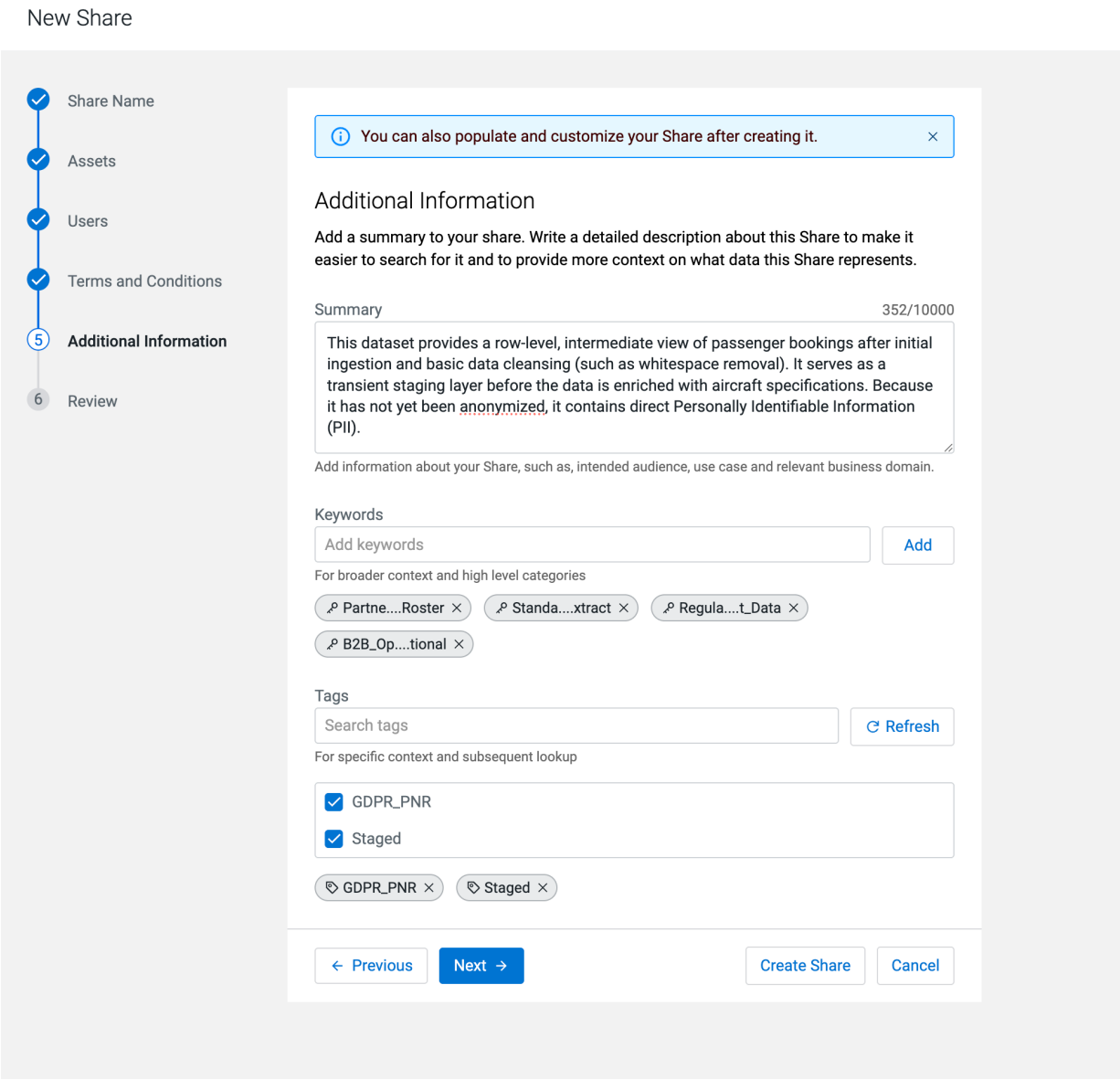
Figure 7: Terms and Conditions step



6. On the Additional Information step, optionally enter a Summary, add Keywords, and select Tags, then click Next.

   The summary (up to 10,000 characters) describes the Data Share content and intended audience, making it easier to discover. Keywords provide broad, high-level categorization. Tags apply Atlas tags (classifications) for specific context and subsequent lookup. Both keywords and Atlas classifications are internal to Cloudera Data Catalog search for Data Shares.

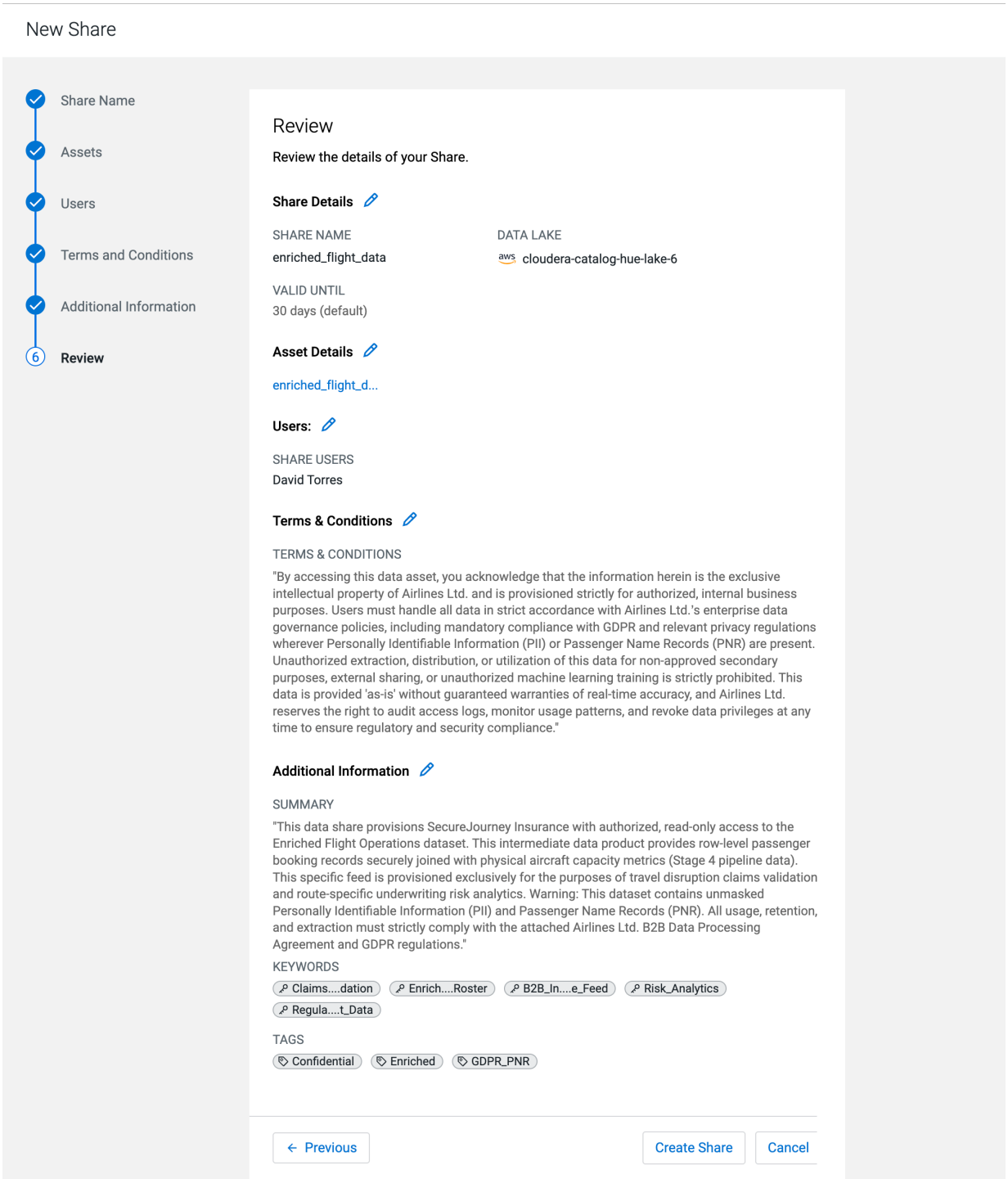
Figure 8: Additional Information step



7. On the Review step, verify all Data Share details and click Create Share.

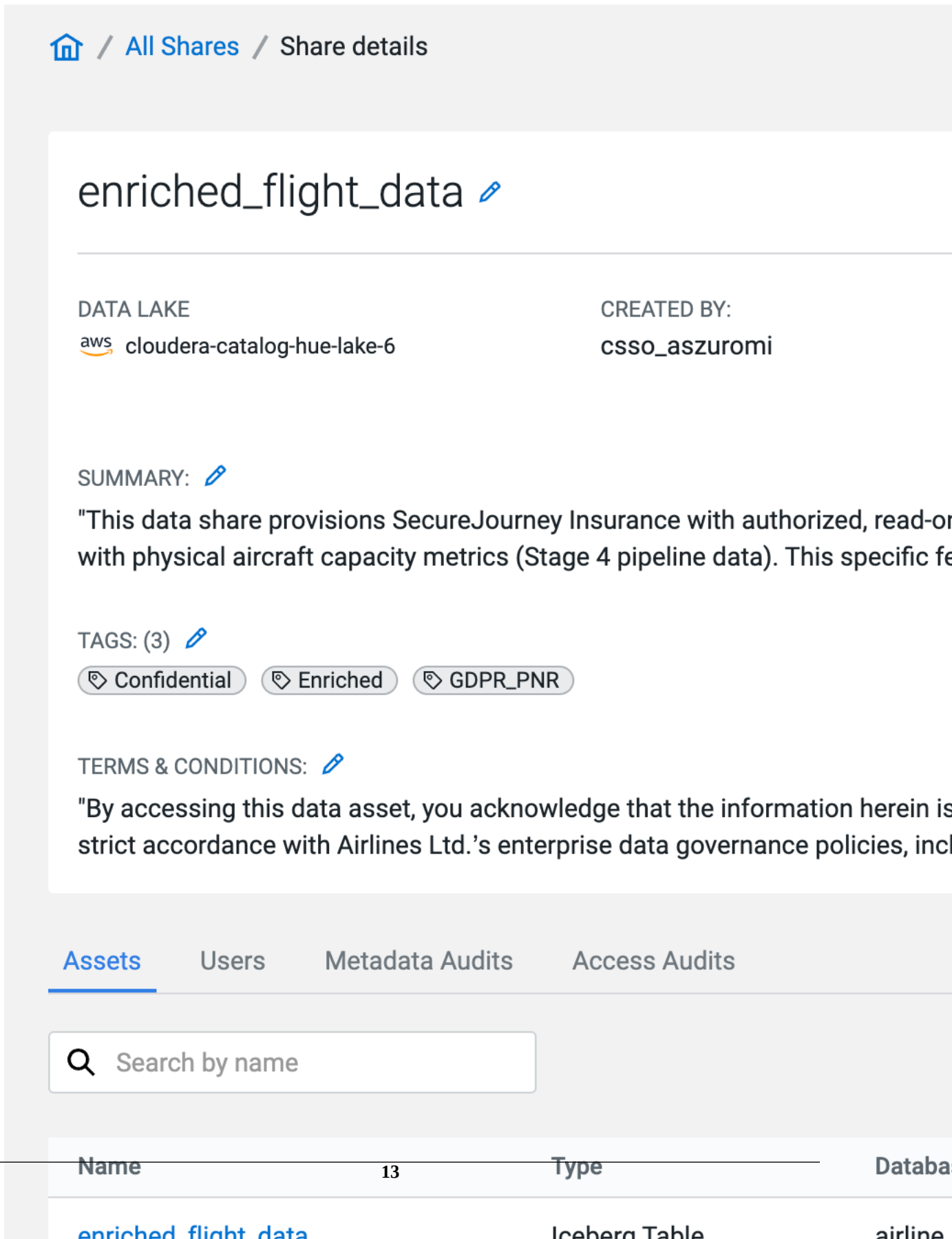
Note: Use the icons next to each section to jump back and correct any details before creating.

Figure 9: New Share review step



Data Shares are created in Not Shared status. Assets are not yet accessible to users until you publish the Data Share.

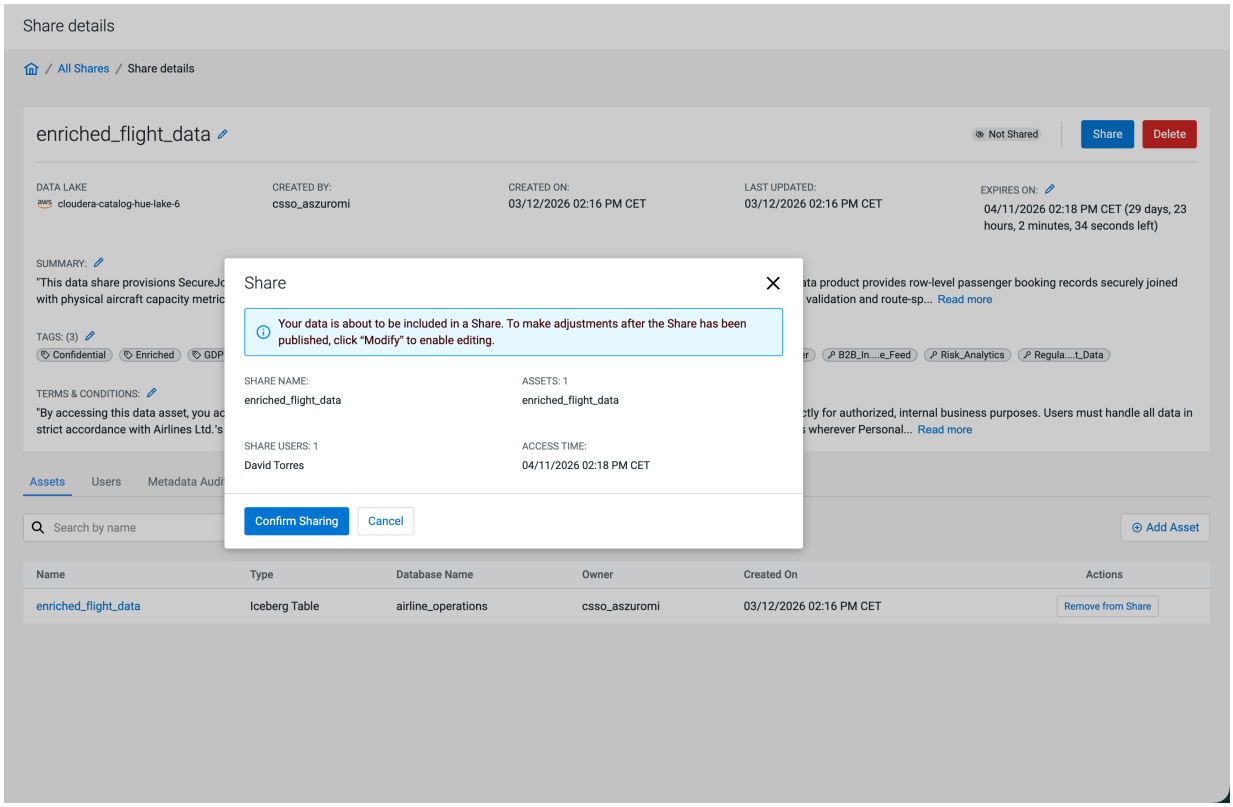
Share details



8. On the Share details page, click Share.

   The Share confirmation dialog opens, showing the Data Share name, included assets, assigned users, and the access expiry time.

Figure 10: Confirm Sharing dialog



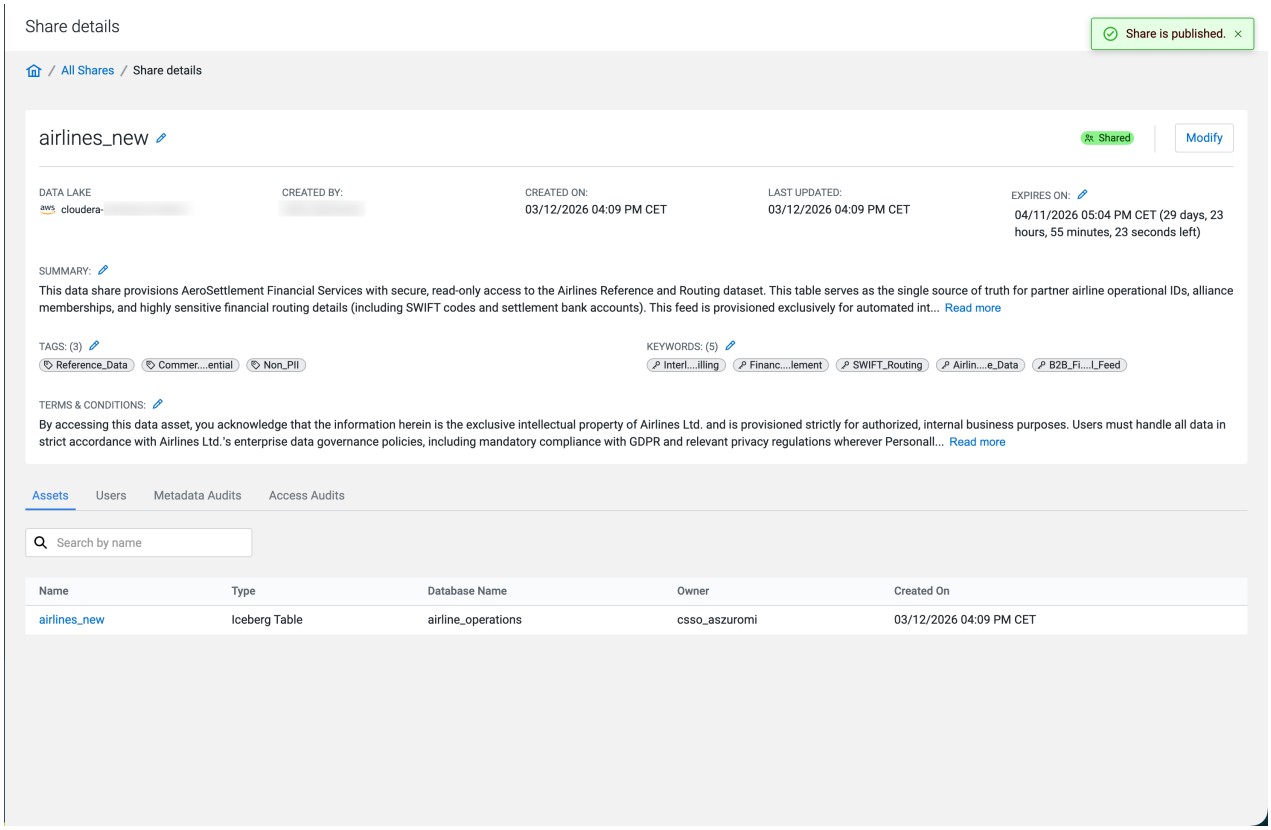
9. Review the summary and click Confirm Sharing to publish the Share.

Note: Once published, you can add new users and change the access time of existing users. However, you cannot remove existing users or change the included assets. To do that, first you need to click Modify to return the share to draft state.

Results

The Data Share is published. The status badge changes to Shared and a success banner confirms publication. The Share button is replaced by a Modify button.

Figure 11: Published Data Share



What to do next

Share the credentials file with the external users so they can configure their Iceberg REST Catalog compatible compute engine. Monitor access activity using the Access Audits tab on the **Share details** page.

Related Information

create-data-share
share-data-share
add-assets-to-data-share
Creating a Data Share with CDP CLI

Modifying Data Shares

Update a Data Share in Cloudera Data Catalog: change some settings while the Share is Shared, or use Modify to edit assets and user membership, then publish again.

About this task

What you can change depends on whether the Data Share is in Shared or Not Shared state (similar to a draft):

Adding or removing assets

You can add or remove Iceberg table assets only while the Share is Not Shared.

Adding or removing users

You can add external users to a Shared Share. Revoking a user's access to the Share (removing them from the Share) is only possible while in Not Shared state after you click Modify.

Per-user access time

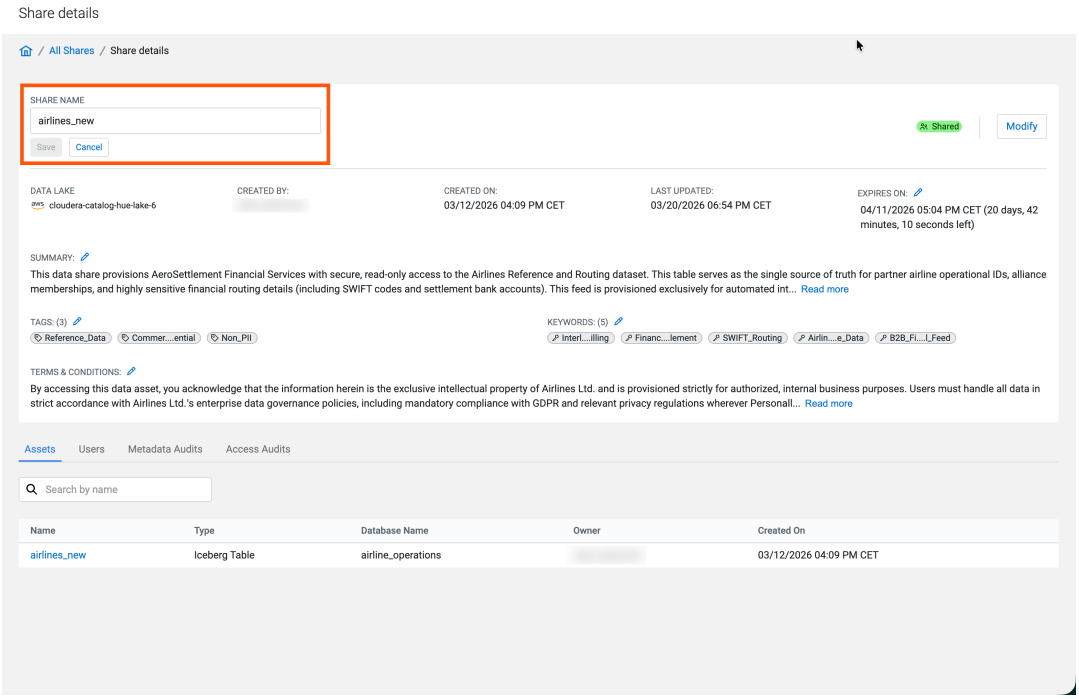
You can change the access date for users who already have access while the Share is Shared.

Share expiration

You can change the overall Data Share expiration time while the Share is Shared.

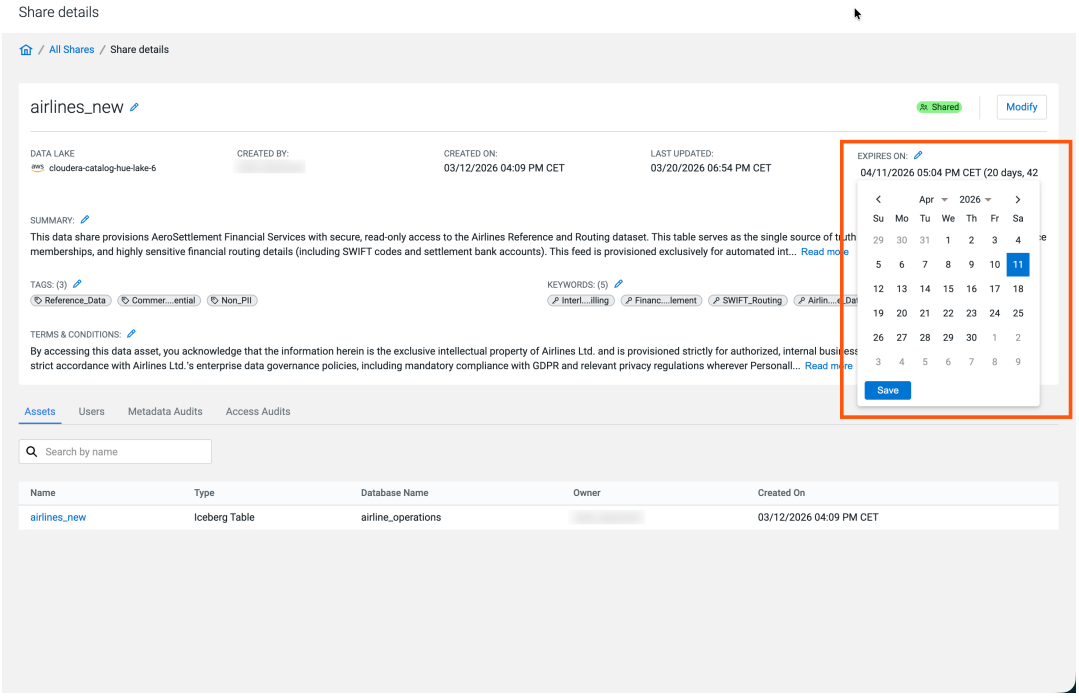
While the Share is Shared, you can edit the following metadata using the icon on the **Share details** page:

Data Share name

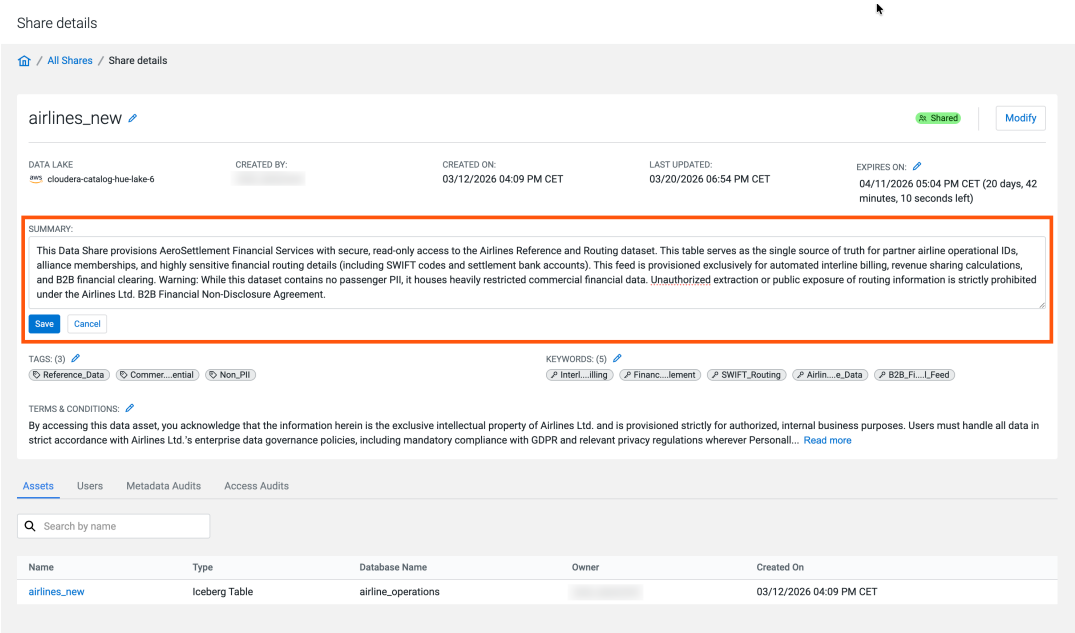


Share expiration

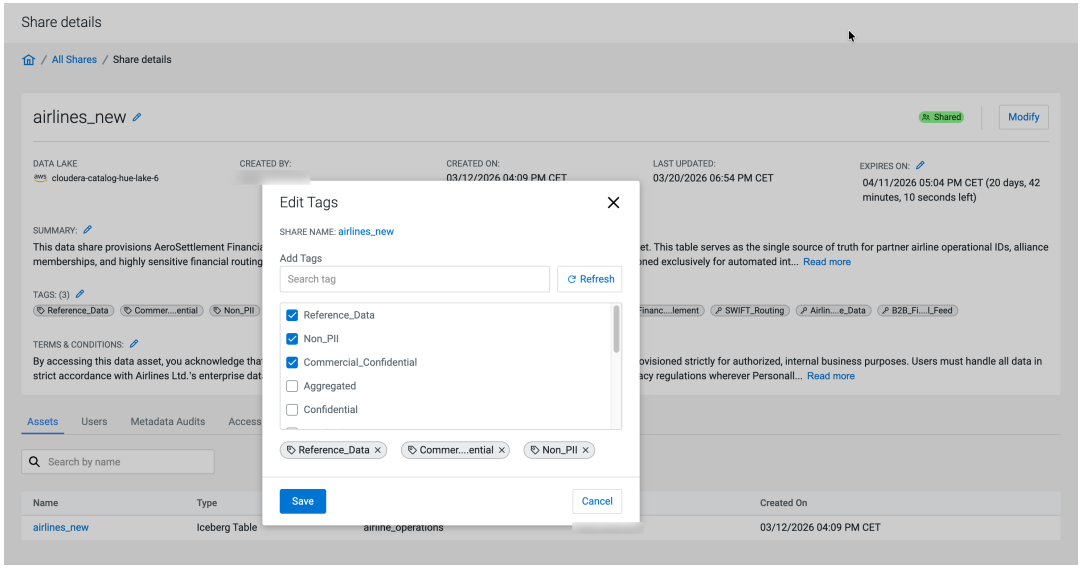
Figure 12: Editing expiration time



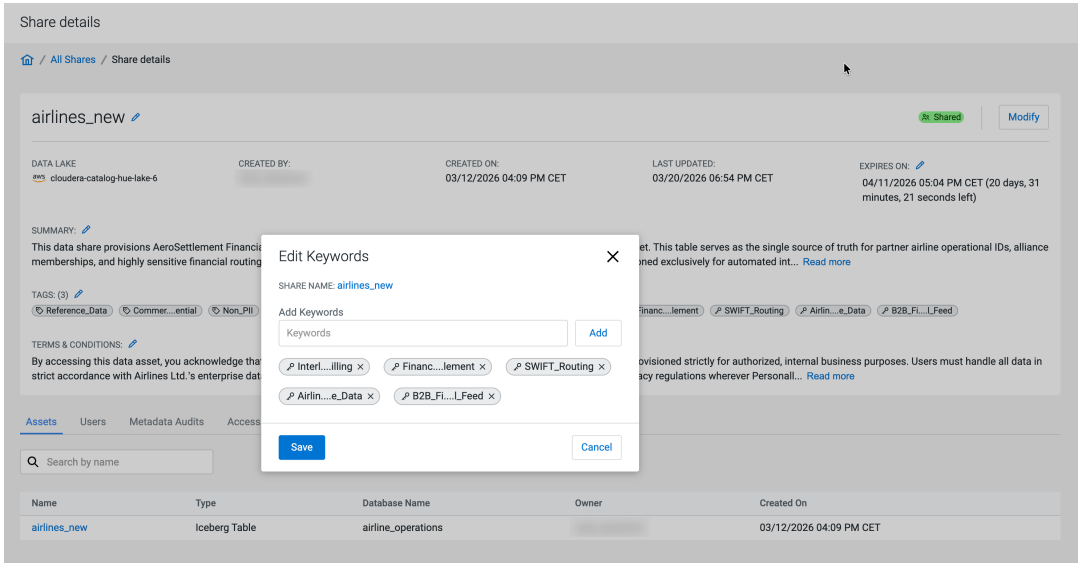
Summary



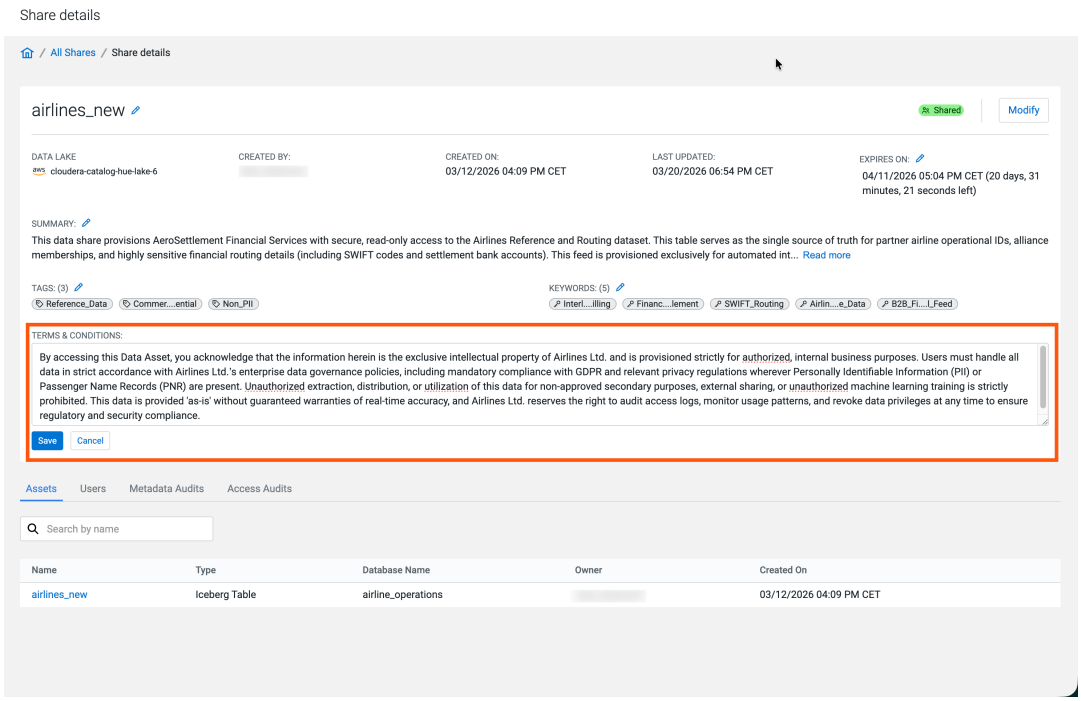
Tags



Keywords



Terms & Conditions



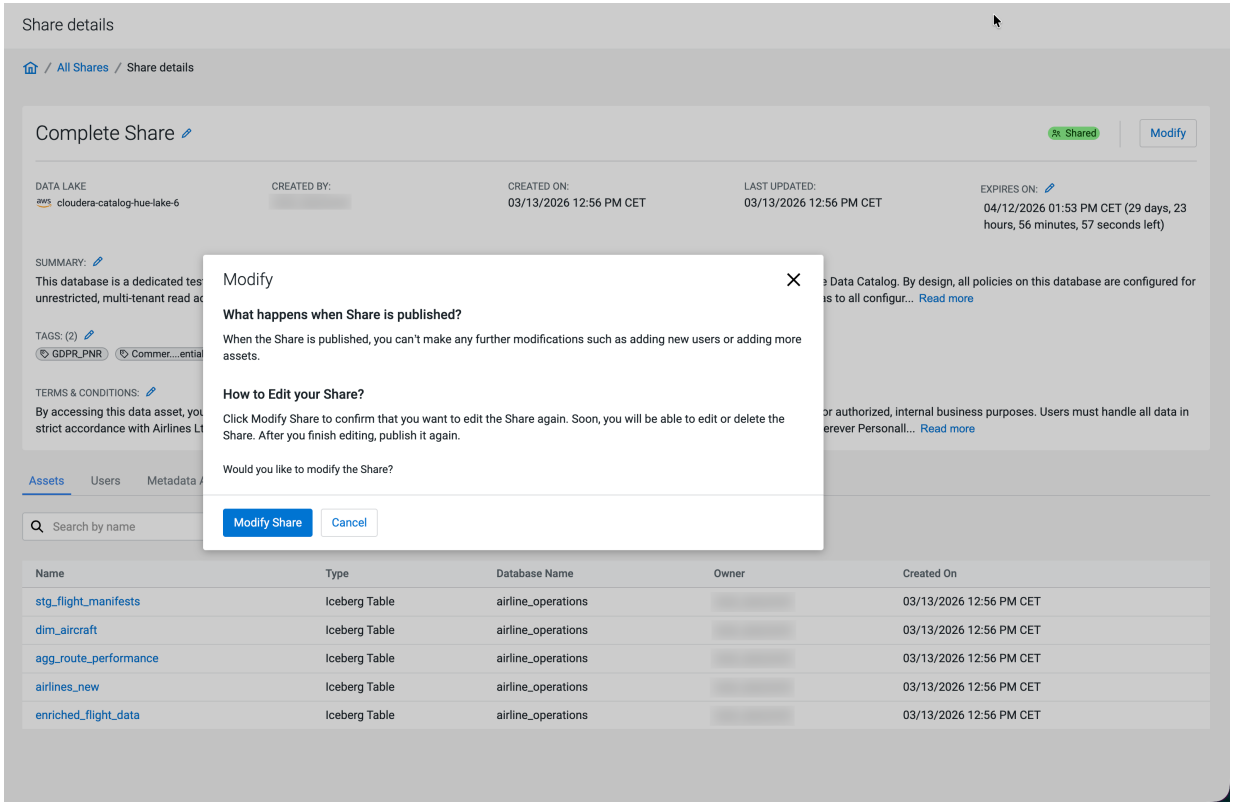
Before you begin

You must have the DataShareAdmin resource role for the relevant environment.

Procedure

1. In Cloudera Data Catalog, go to **Data Sharing** and open the Data Share you want to change.
2. If the Share status is Shared, click Modify.

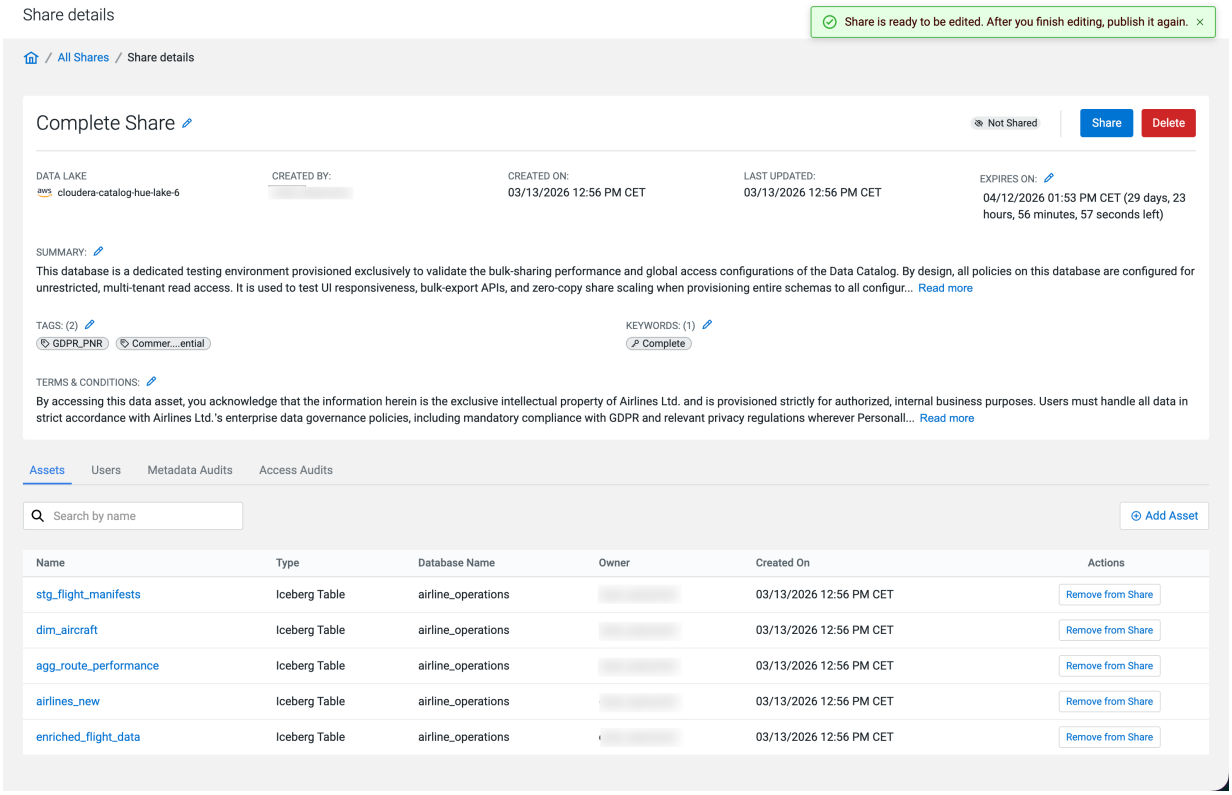
Figure 13: Modify dialog for a published Data Share



3. In the confirmation dialog, click Modify Share.

   The **Share details** page reloads with Not Shared status. You can now edit the Share using the same sections as in the **New Share** wizard flow. For step-level guidance, see https://docs.cloudera.com/data-catalog/cloud/data-sharing-data-catalog/topics/dc-creating-a-new-share.html.

Figure 14: Share details page in Not Shared state

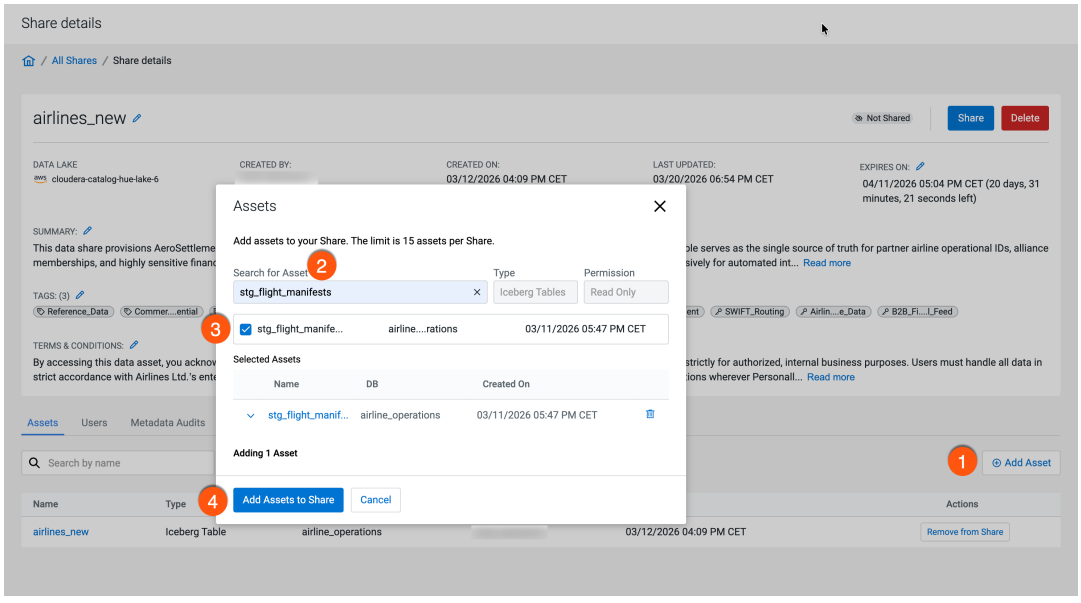


Make your changes in the Data Share:

Adding or removing assets

You can add assets, or remove them individually by clicking the icon on the asset row. To remove multiple assets in bulk, select the checkboxes for the assets and click Remove Asset.

Figure 15: Adding assets



Important:

When you add an Iceberg table that Apache Ranger Fine-Grained Access Control (FGAC) protects with column masking or row-level filtering, Cloudera Data Catalog displays a warning titled Table with fine-grained access control policies. A warning icon also appears on the asset row for tables already included in the Share.

External consumers who access the shared table through the REST Catalog might not receive the same column masking or row-level filtering protections. Sharing the table could expose data more broadly than intended. Review the table access policies before you add or keep the asset. You can still include the table after you acknowledge the warning.

Figure 16: FGAC warning when adding an asset

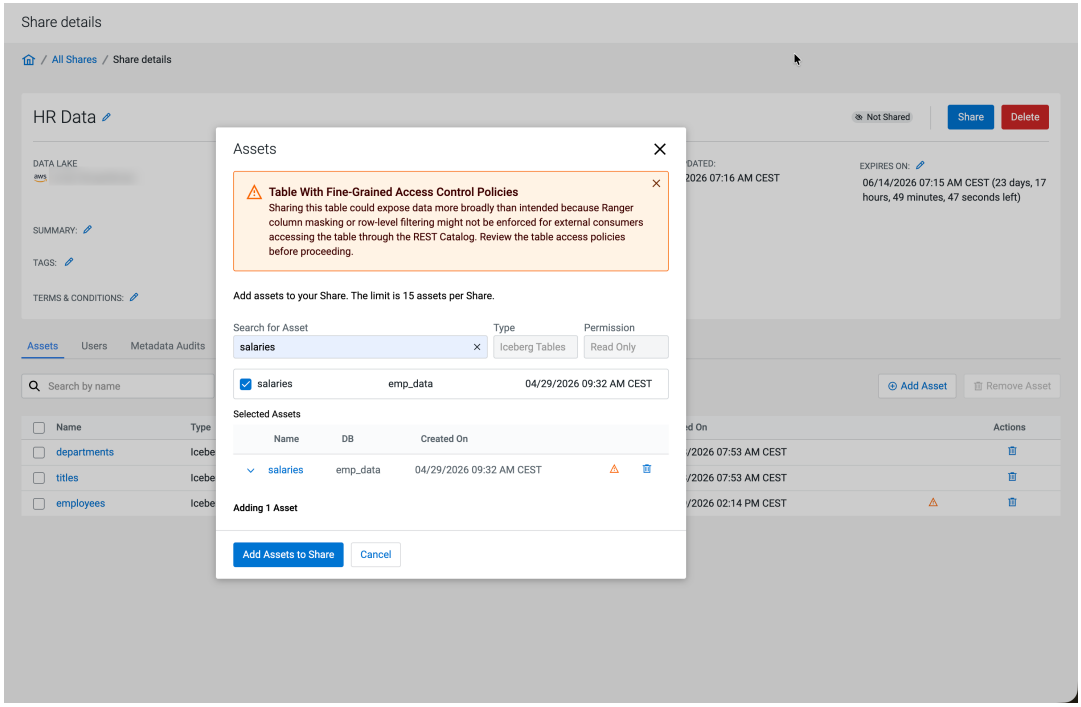


Figure 17: FGAC warning icon on a shared asset

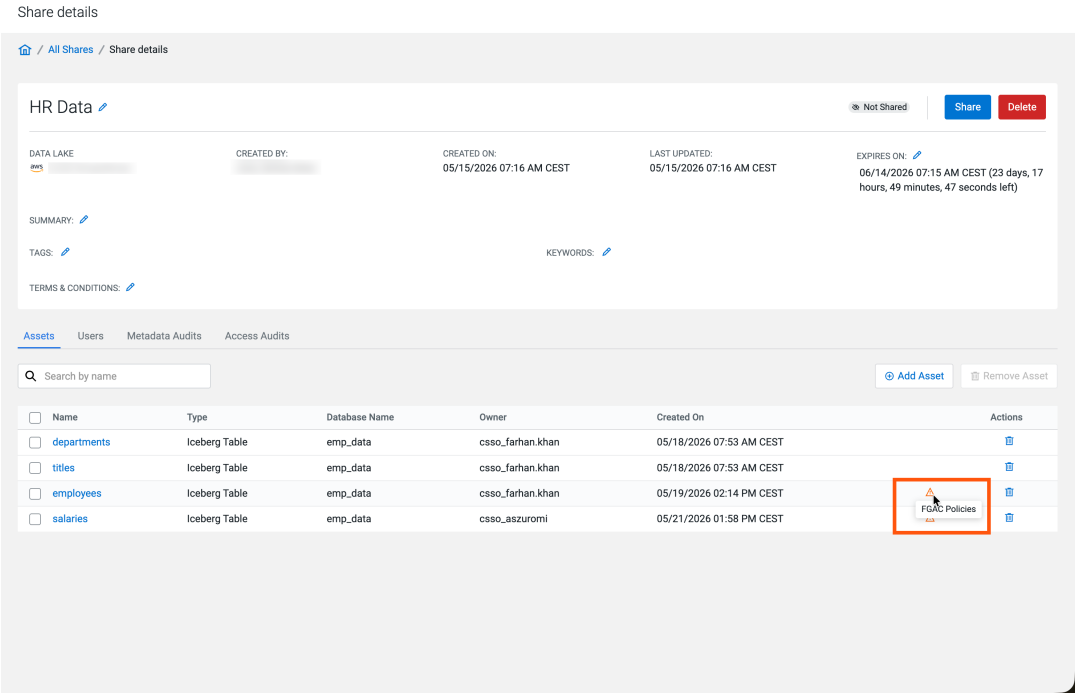
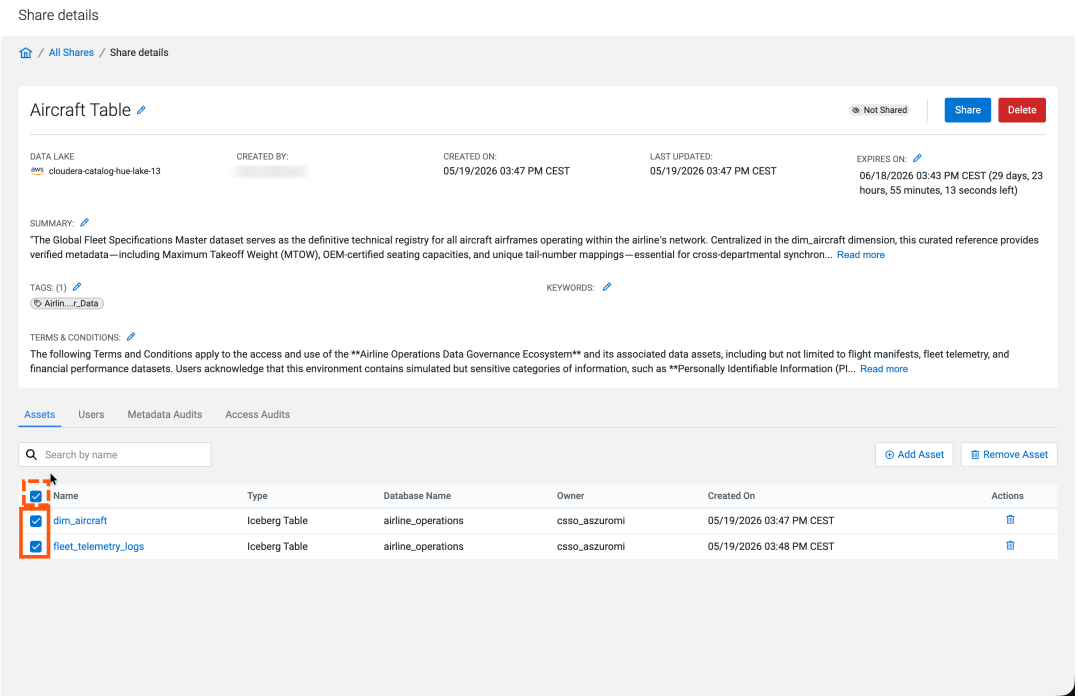


Figure 18: Removing assets

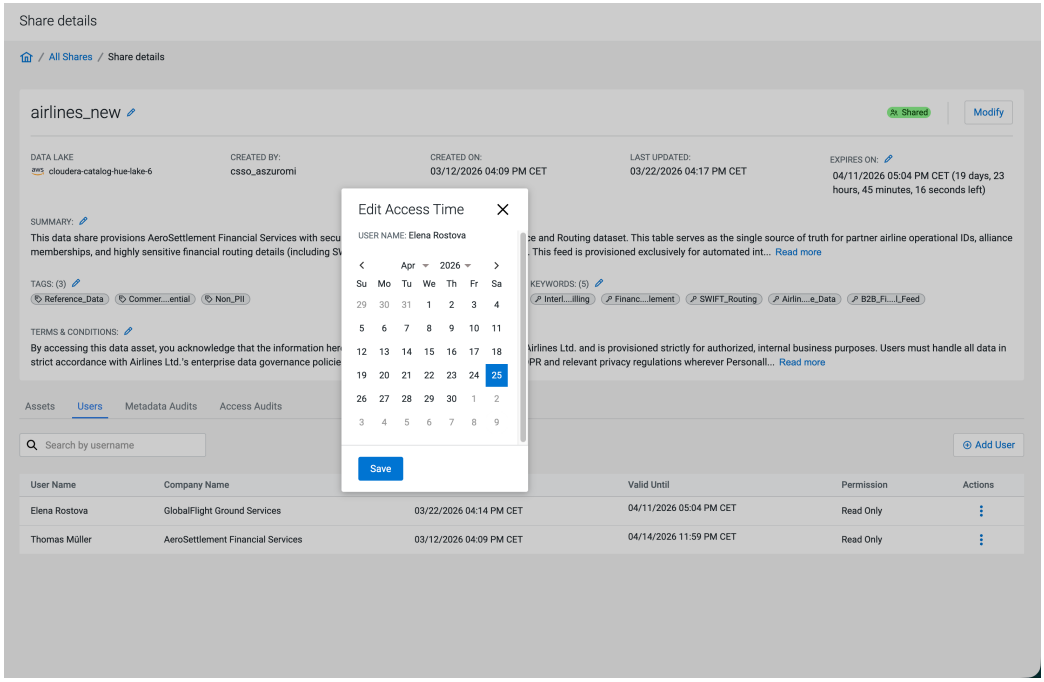


Warning: Removing assets from a Data Share does not have a confirmation prompt.

Editing access time

a. Click the ⋮ icon, then Edit Access Time.

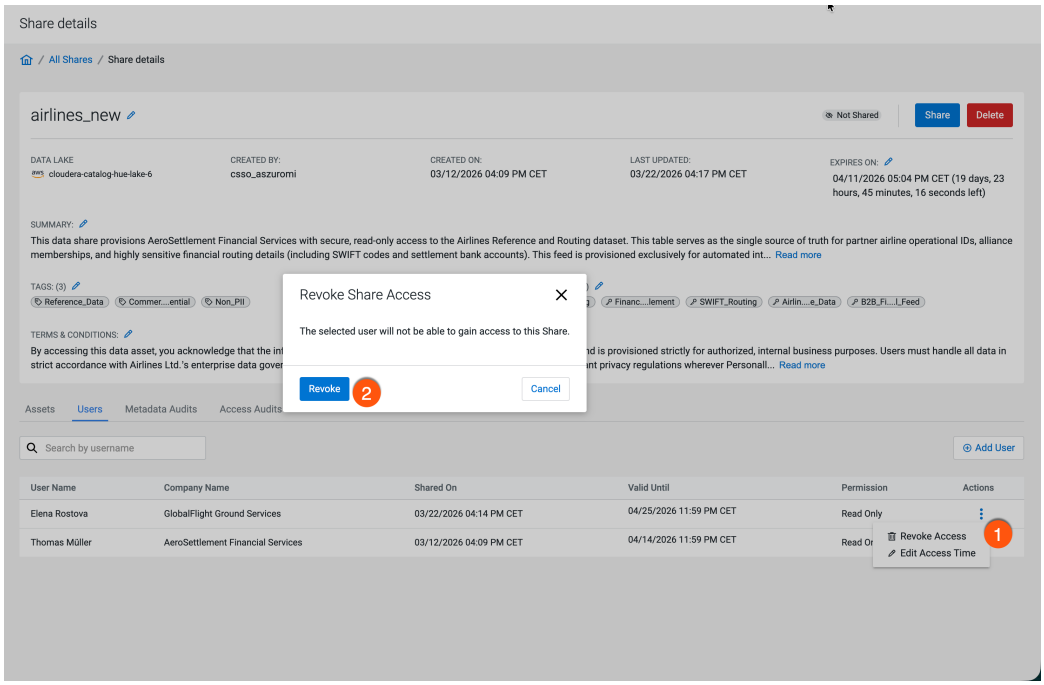
b. Select a new date.



Revoking user access

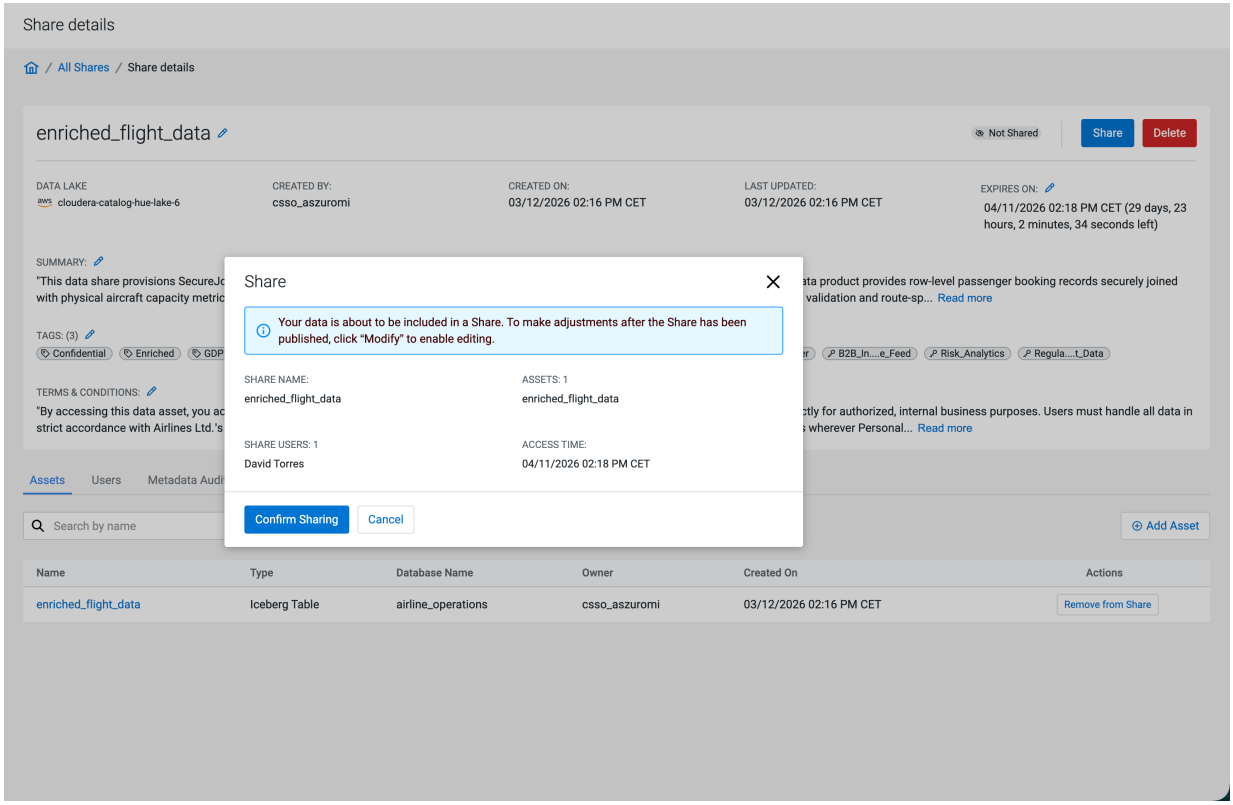
a. Click the ⋮ icon, then Revoke Access.

b. Confirm revoking the user's access.



4. Update the Share as needed (for example, assets, users, terms and conditions, or additional information), then click Share and Confirm Sharing to publish the updated Share.

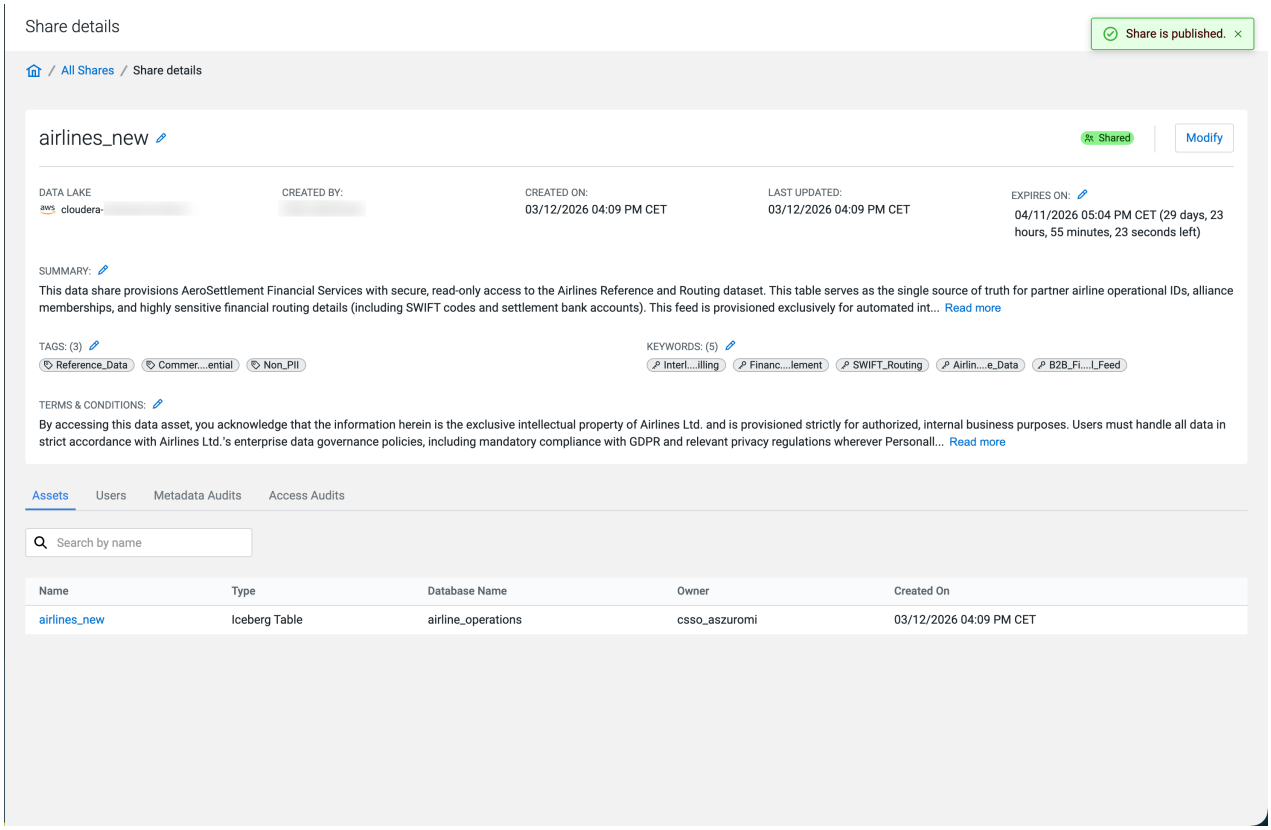
Figure 19: Confirm Sharing dialog



Results

The Data Share is published again with Shared status. External users receive access consistent with the updated configuration and assignments.

Figure 20: Published Data Share



Related Information

unshare-data-share
update-data-share
add-assets-to-data-share
grant-access-to-external-users-on-data-share
update-access-of-external-users-on-data-share
remove-access-of-external-users-on-data-share
share-data-share

Revoking user credentials

Revoke the credentials of an external user to immediately block their access to all Data Shares until new credentials are regenerated.

About this task

Revoking credentials invalidates a user's Client ID and Secret. The user will not be able to access any of their previous Data Shares or new ones until their credential is regenerated. Use this action when a user's credentials are compromised, when access needs to be temporarily suspended, or when a user leaves the organization.

Before you begin

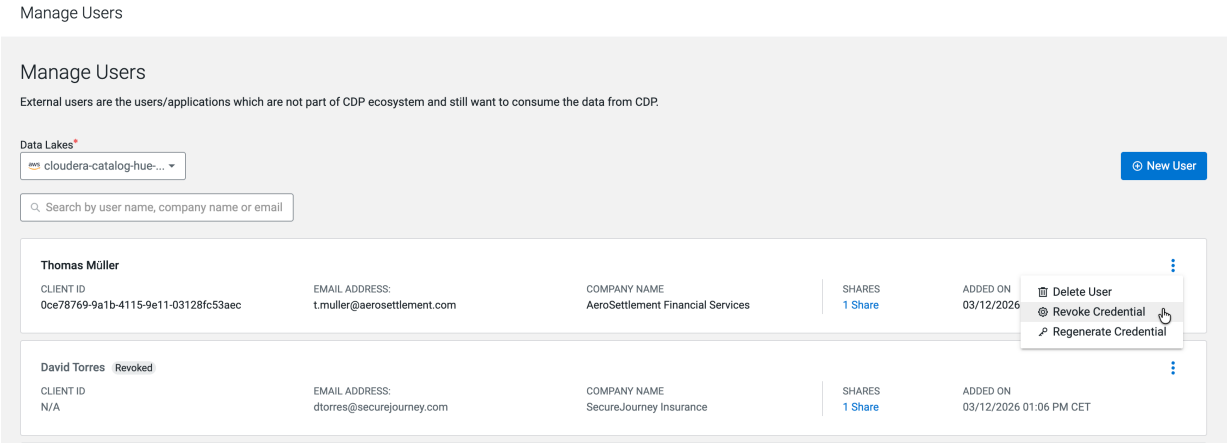
You must have the DataShareAdmin resource role for the relevant environment.

Procedure

1. In Cloudera Data Catalog, navigate to Manage Users and select the target Data Lake from the dropdown.

2. In the user row, click the actions menu (#) and select Revoke Credential.

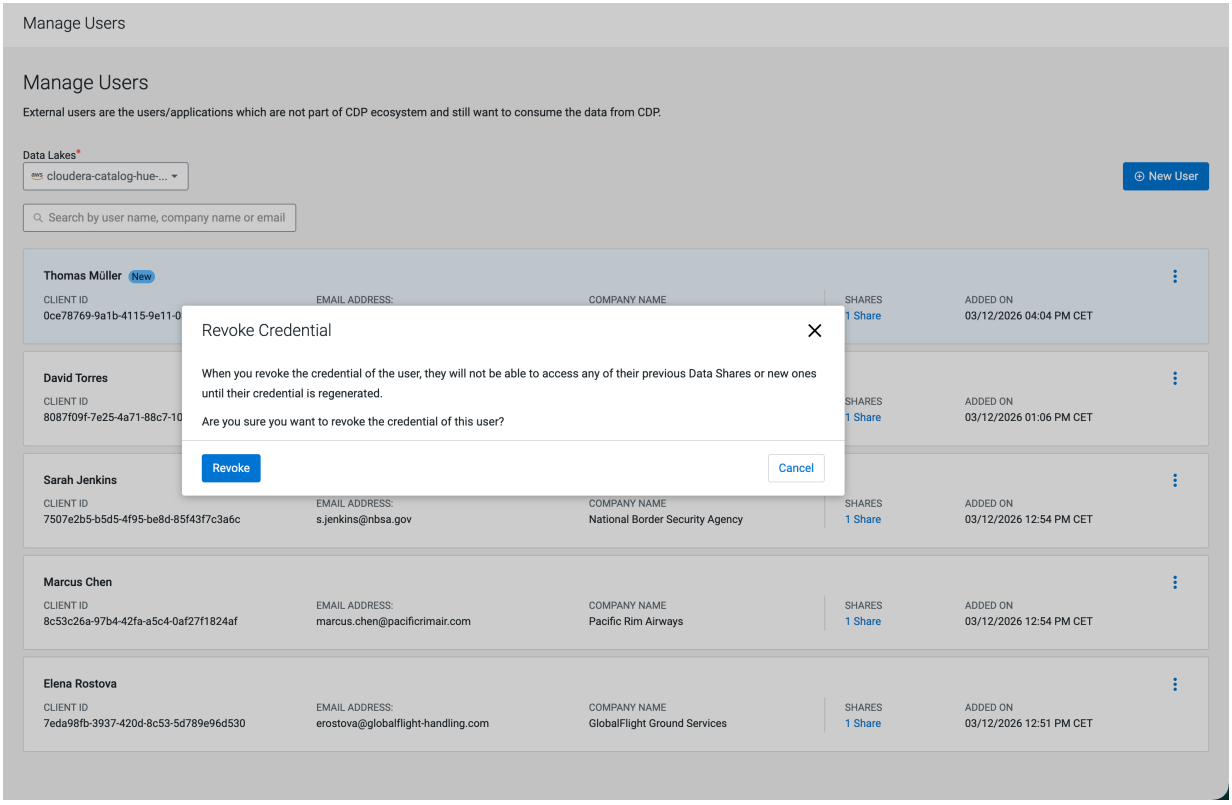
Figure 21: Actions menu with Revoke Credential option



3. In the Revoke Credential confirmation dialog, click Revoke.

Warning: Revoking credentials immediately blocks the user from accessing all their current and future Data Shares until credentials are regenerated.

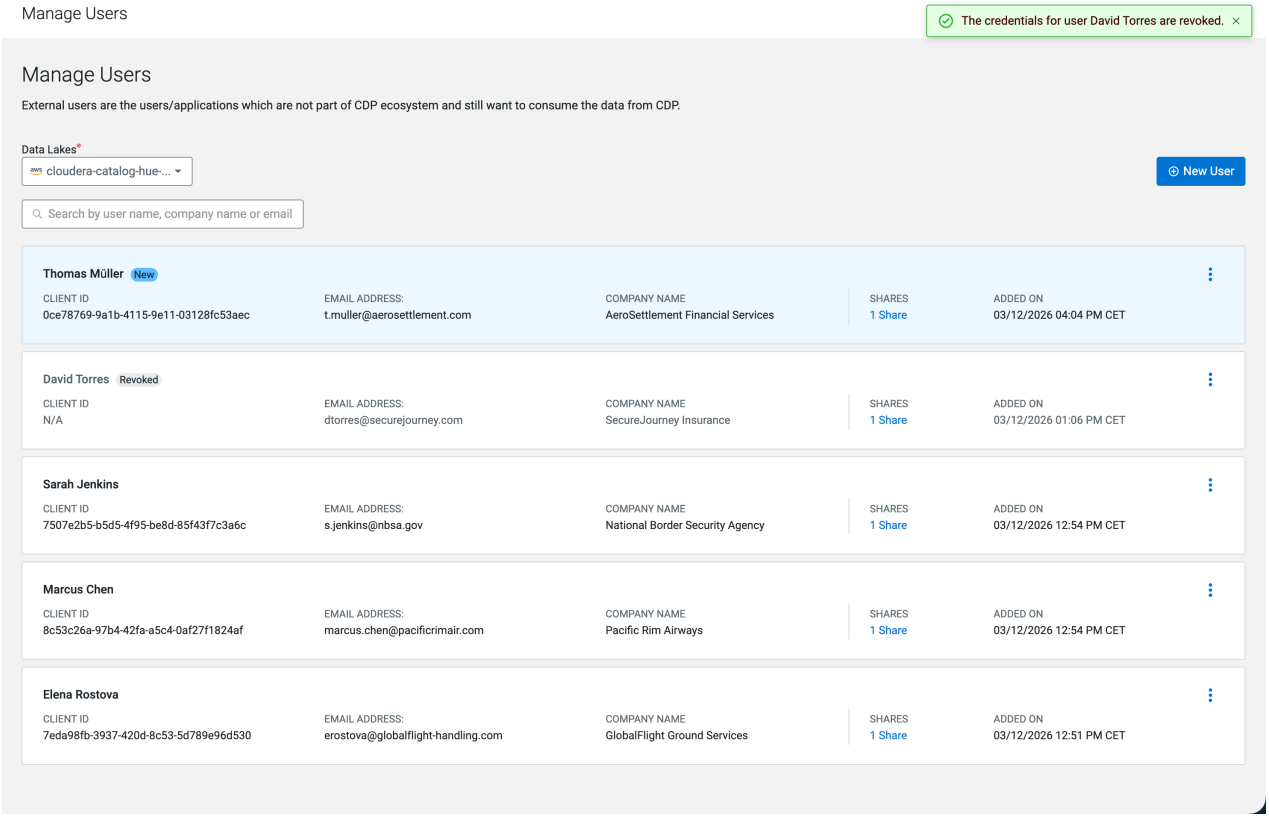
Figure 22: Revoke Credential confirmation dialog



Results

A success banner confirms that the credentials are revoked. The user row displays a Revoked badge and the Client ID field shows N/A.

Figure 23: User row in revoked state



What to do next

To restore the user's access, regenerate their credentials from the same actions menu.

Related Information

revoke-external-user-credentials
remove-access-of-external-users-on-data-share
Regenerating user credentials

Regenerating user credentials

Regenerate the credentials of an external user to issue a new Client ID and Secret, restoring their access to Data Shares.

About this task

Regenerating credentials invalidates the user's existing Client ID and Secret and issues a new pair. Use this action to restore access for a user whose credentials were revoked, or to replace credentials that were lost or potentially compromised.

Before you begin

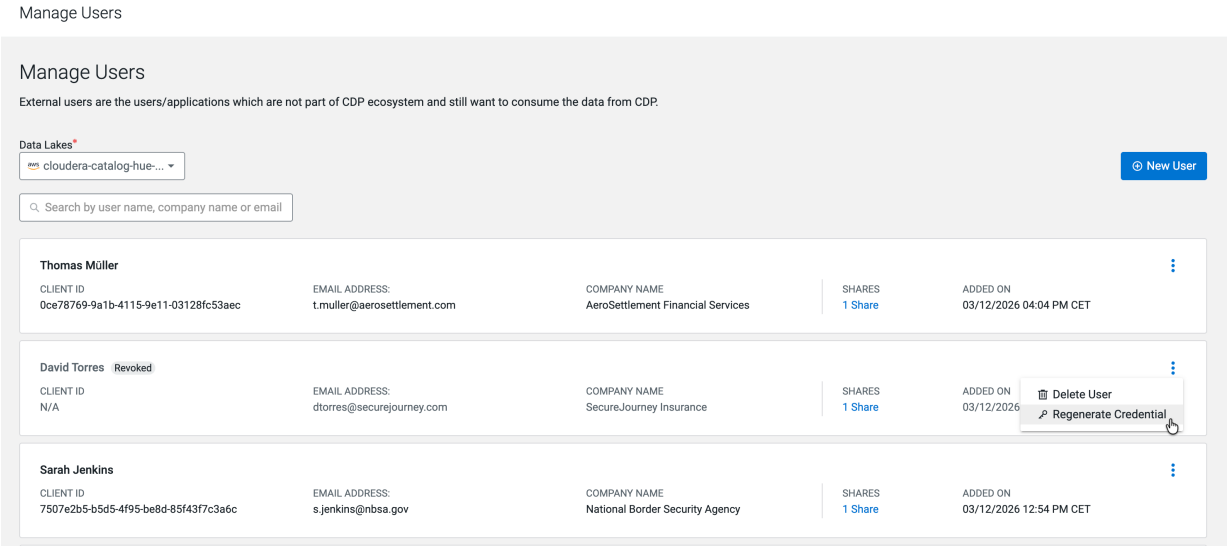
You must have the DataShareAdmin resource role for the relevant environment.

Procedure

1. In Cloudera Data Catalog, go to Manage Users and select the target Data Lake from the dropdown.

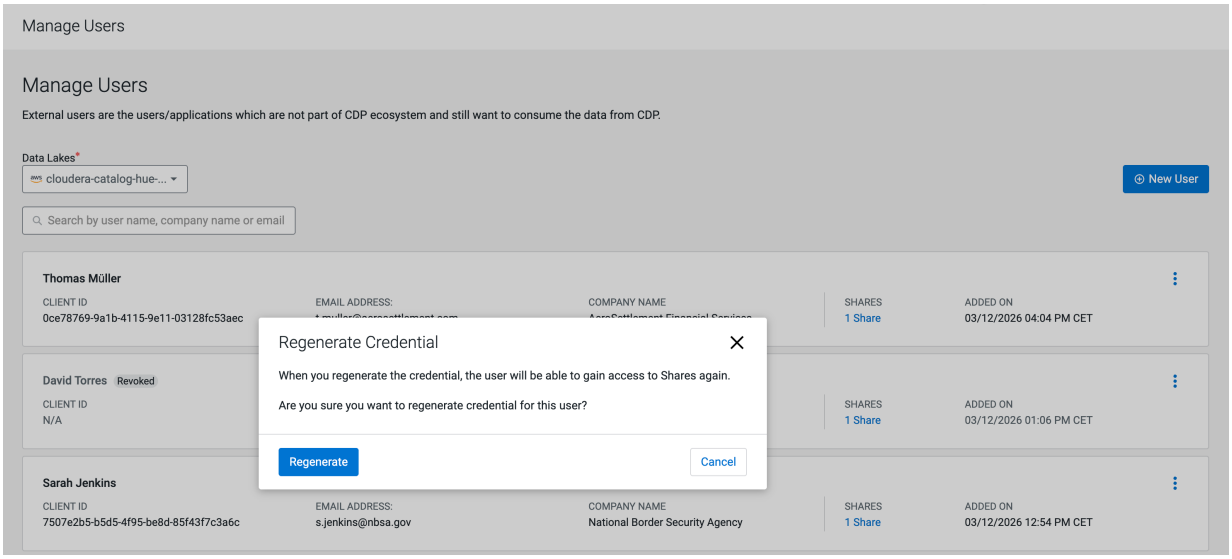
2. In the user row, click the actions menu (#) and select Regenerate Credential.

Figure 24: Actions menu with Regenerate Credential option



3. In the Regenerate Credential confirmation dialog, click Regenerate.

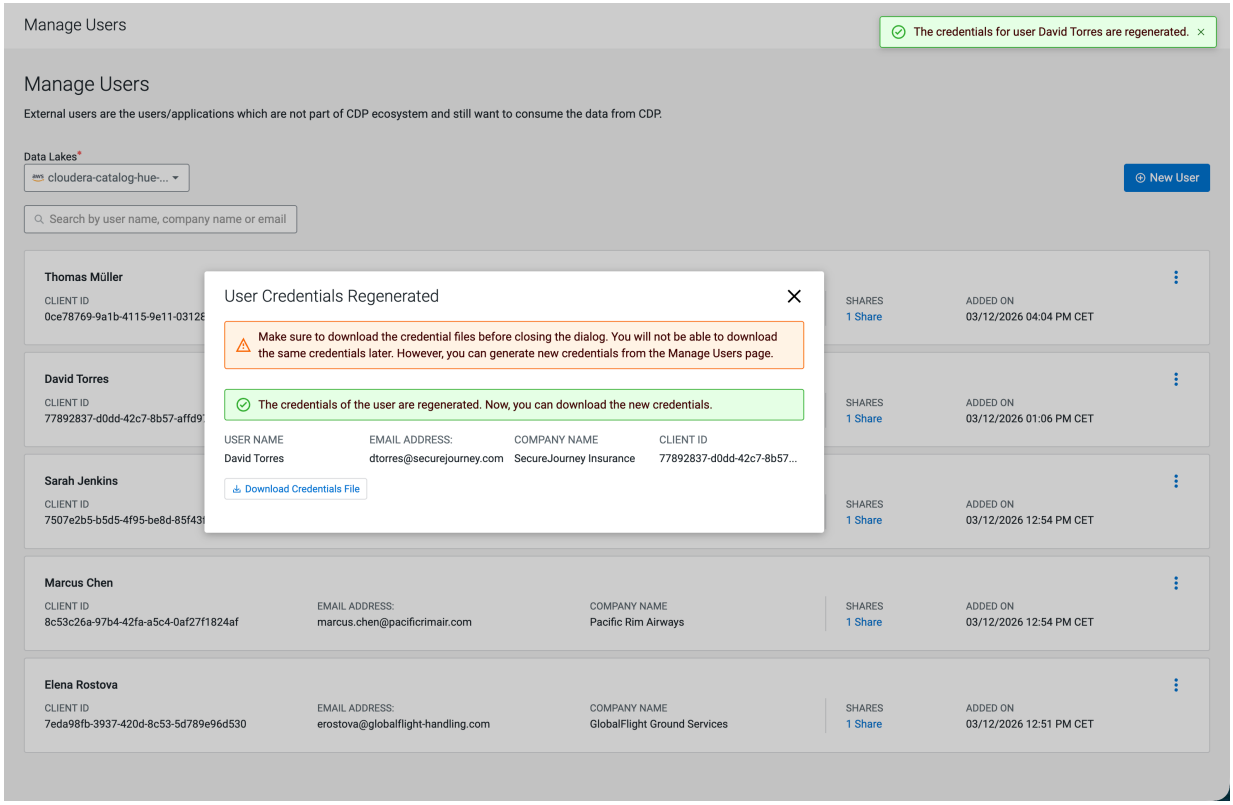
Figure 25: Regenerate Credential confirmation dialog



4. In the User Credentials Regenerated dialog, click Download Credentials File to save the new credentials before closing the dialog.

Important: Download the credentials file before closing this dialog. The same credentials cannot be downloaded again. If needed, you can regenerate new credentials again from the Manage Users page.

Figure 26: User Credentials Regenerated dialog



Results

A success banner confirms that the credentials are regenerated. The user's Client ID is updated in the Manage Users list. The user can now access their Data Shares using the new credentials.

What to do next

Securely share the downloaded credentials file with the external user so they can reconfigure their Iceberg REST Catalog compatible compute engine with the new Client ID and Secret.

Related Information

regenerate-external-user-credentials

Deleting a Data Share

Permanently delete a Data Share in Cloudera Data Catalog to remove it and revoke all associated user access.

About this task

Deleting a Data Share permanently removes it and immediately revokes access for all users who were granted access through it. The underlying Iceberg table assets are not affected. A Data Share can only be deleted when it is in Not Shared status. If the Share is currently published, you must put it into draft state first by clicking Modify.

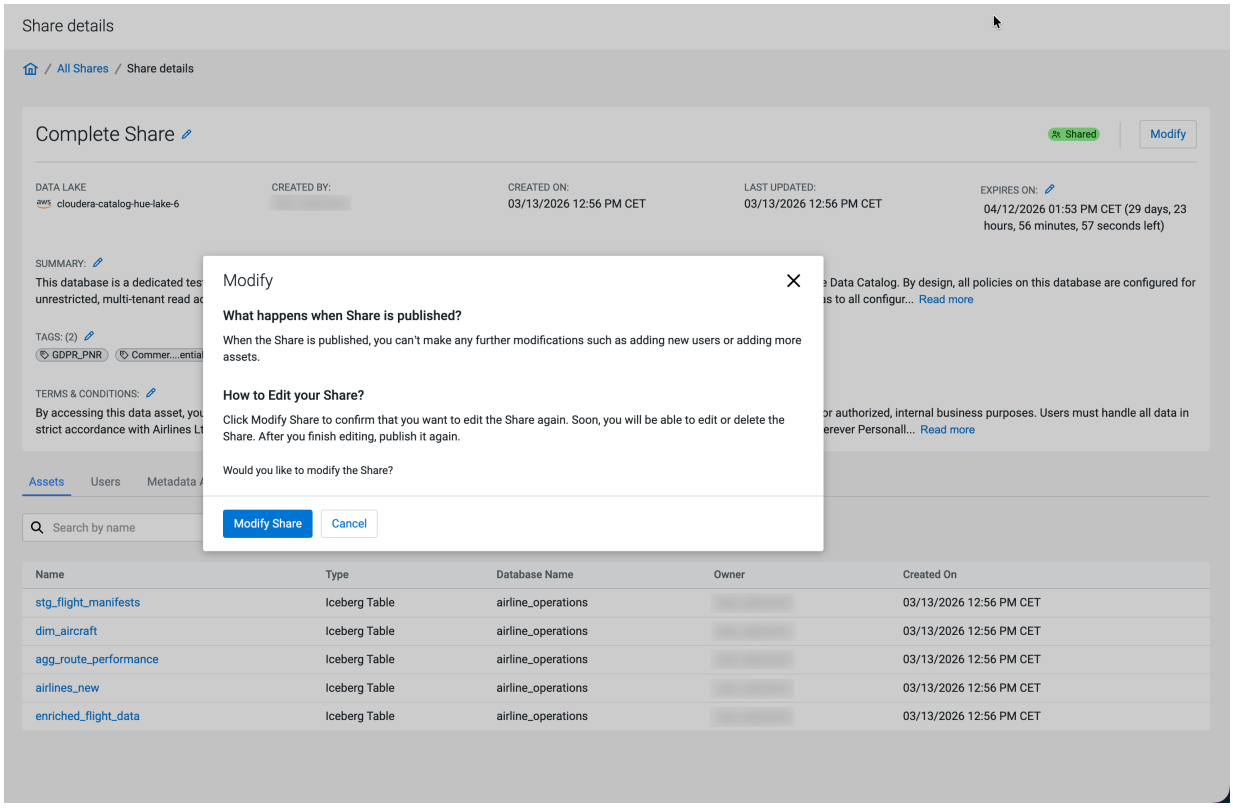
Before you begin

You must have the DataShareAdmin resource role for the relevant environment.

Procedure

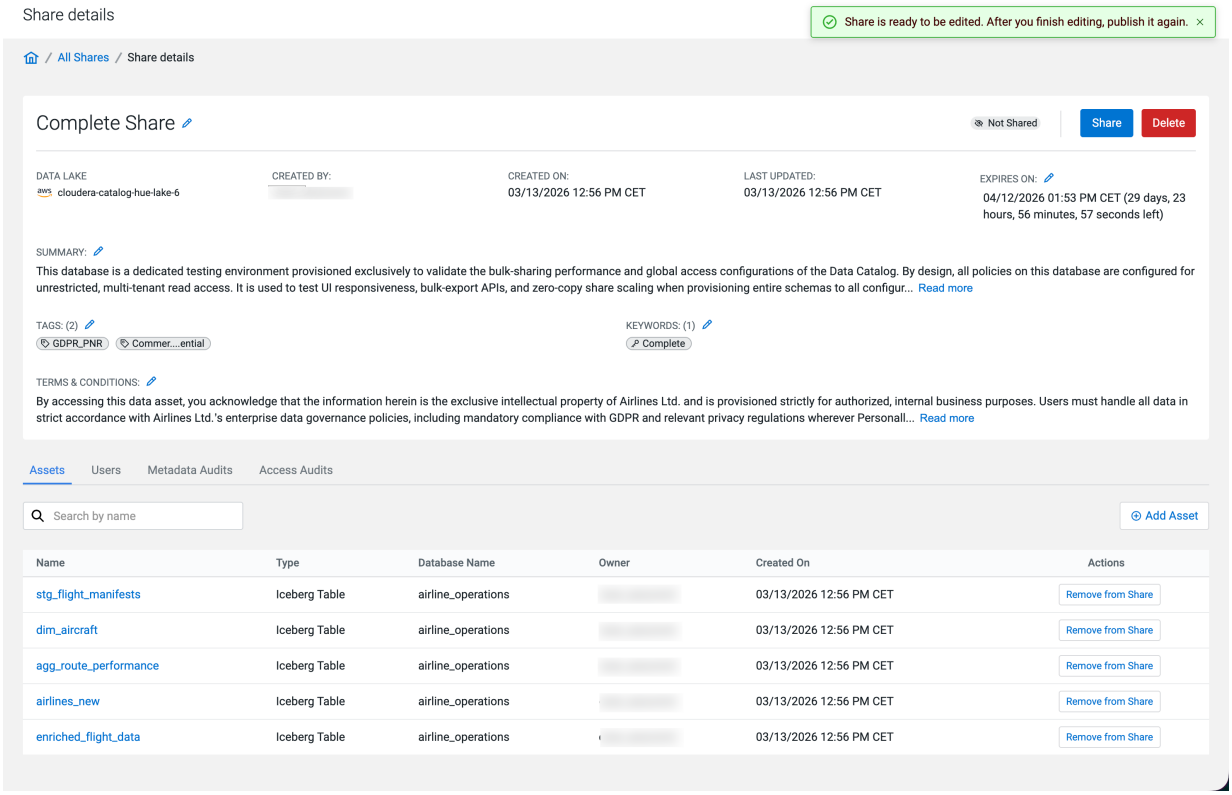
1. In Cloudera Data Catalog, go to **Data Sharing** and open the Data Share you want to delete.
2. If the Data Share status is Shared, click Modify and then click Modify Share in the confirmation dialog to put the Share back into draft state.

Figure 27: Modify dialog for a published Data Share



The **Share details** page reloads with Not Shared status and the Delete button becomes available.

Figure 28: Share details page in Not Shared state

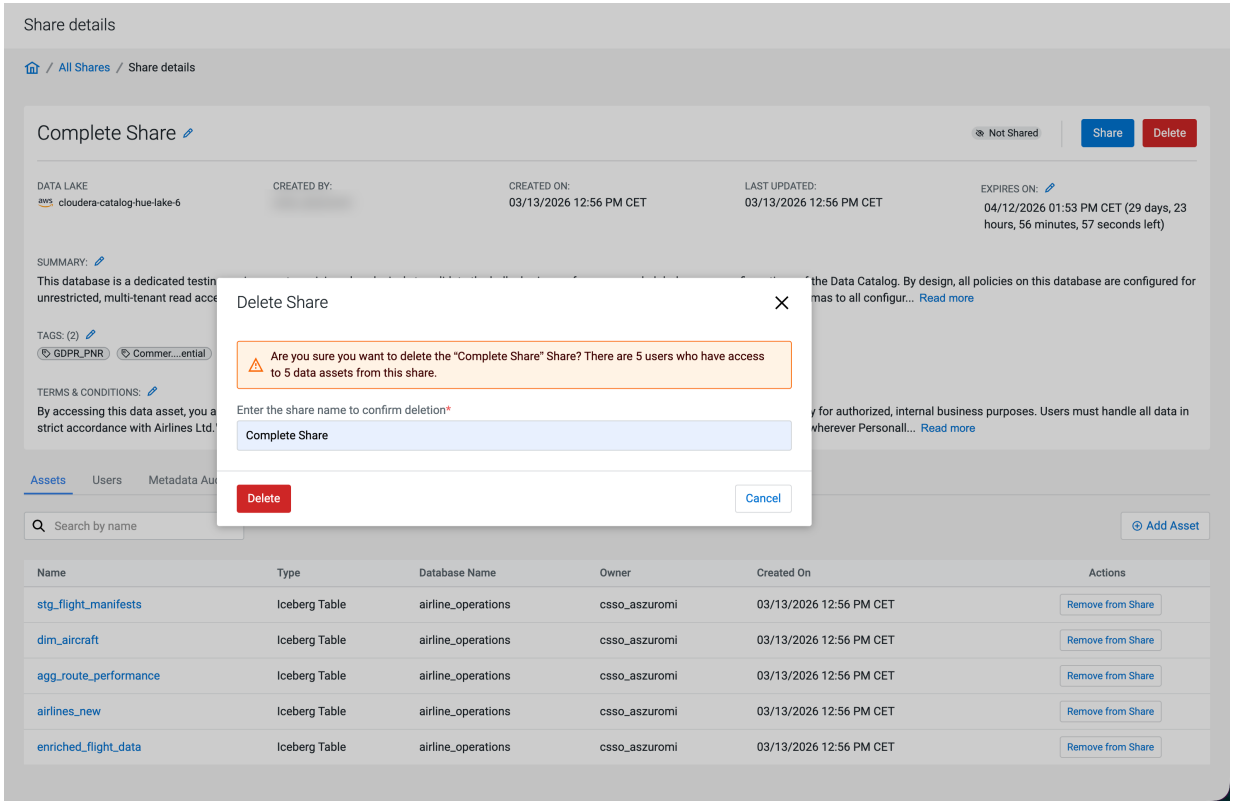


3. Click Delete on the **Share details** page.

   The **Delete Share** dialog opens, warning you of the number of users and assets that will lose access.
4. In the Delete Share dialog, type the exact share name in the confirmation field to enable the Delete button.

   Warning: Deletion is permanent and cannot be undone. All users currently granted access through this Share will immediately lose access.

Figure 29: Delete Share confirmation dialog

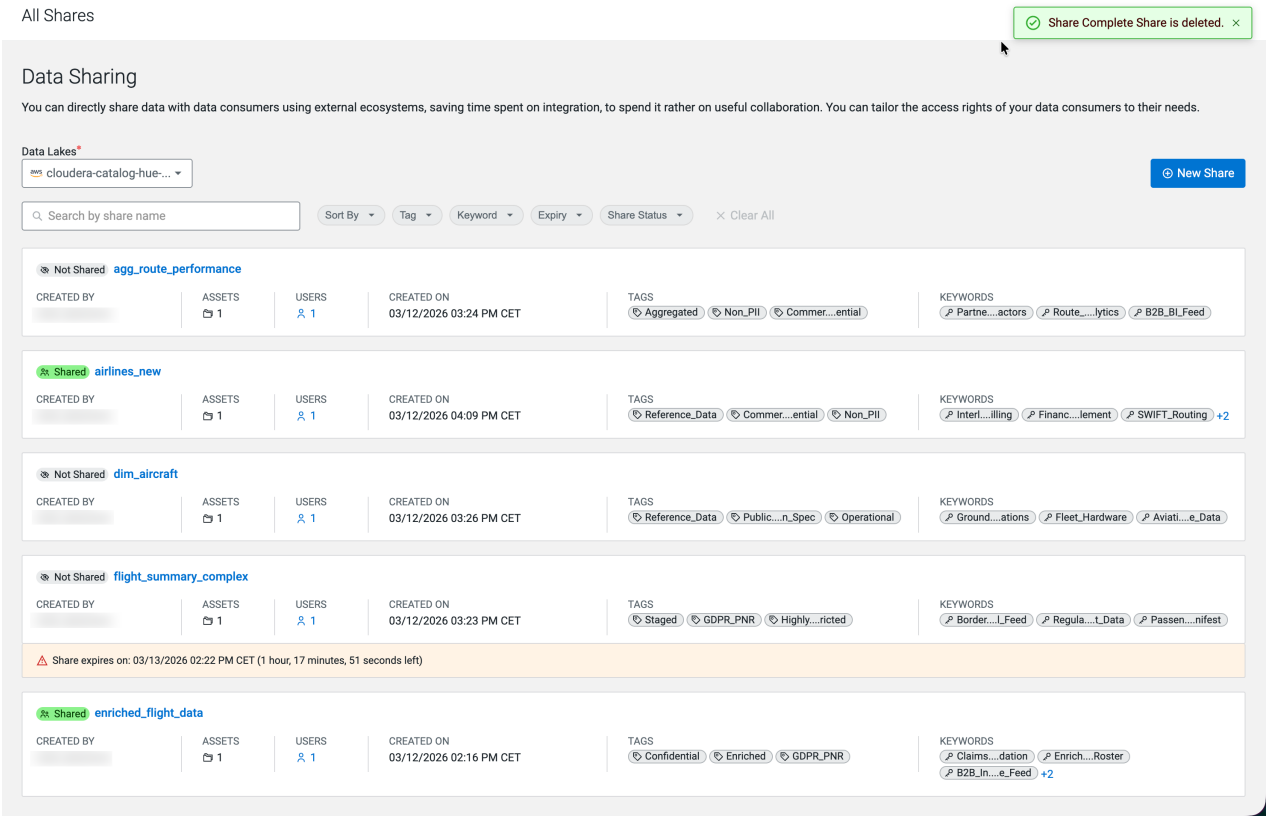


5. Click Delete to confirm.

Results

The Data Share is permanently deleted. You are redirected to the **All Shares** page, where a success banner confirms the deletion.

Figure 30: All Shares page after successful deletion



Related Information

delete-data-share

unshare-data-share

Verifying external access to a Data Share

Confirm that a Data Share is correctly set up and accessible to an external Data Consumer by exchanging the Client ID and Secret for a Knox access token and listing the shared Iceberg tables using the Cloudera Iceberg REST Catalog API.

About this task

After publishing a Data Share and distributing credentials to a Data Consumer, you can verify end-to-end access without a compute engine such as Spark, Snowflake, or Databricks. This procedure uses `curl` to perform two checks: it exchanges the Client ID and Secret for a short-lived Knox JWT access token, then uses that token to call the Cloudera Iceberg REST Catalog API and list the namespaces and tables that are visible to the consumer. Successful completion confirms that Knox credential exchange, Ranger authorization, and Iceberg REST routing are all functioning correctly for the Data Share.

Before you begin

- The Data Share is published (Shared status) and the target Iceberg table has been added as an asset.
- The external user is registered and their credentials CSV file is available. For more information, see https://docs.cloudera.com/data-catalog/cloud/data-sharing-data-catalog/topics/dc-creating-external-users.html.
- `curl` is installed on the machine running the test.

Procedure

1. Set your environment variables from the credentials CSV file and your Data Lake details:

```
CLIENT_ID="[***CLIENT-ID***]"
CLIENT_SECRET="[***CLIENT-SECRET***]"

DATALAKE_NAME="[***YOUR-DATALAKE-NAME***]"
GATEWAY_HOST="[***DATALAKE-MASTER-HOSTNAME***]"
TOPOLOGY="cdp-datashare-access"

BASE_URL="https://${GATEWAY_HOST}/${DATALAKE_NAME}/${TOPOLOGY}"
```

```
CLIENT_ID="e8a4239e-e3e2-4d3f-ba8d-3540ee4a895a"
CLIENT_SECRET="WlRoaE5ES..."

DATALAKE_NAME="cloudera-catalog-hue-lake-6"
GATEWAY_HOST="cloudera-catalog-hue-lake-6-master0.cloudera.svbr-nqvp.int.
cldr.work"
TOPOLOGY="cdp-datashare-access"

BASE_URL="https://cloudera-catalog-hue-lake-6-master0.cloudera.svbr-nqvp
.int.cldr.work/cloudera-catalog-hue-lake-6/cdp-datashare-access"
```

2. Request a Knox JWT access token by exchanging the Client ID and Secret:

```
TOKEN_RESPONSE=$(curl -skS -X POST \
  -H "Content-Type: application/x-www-form-urlencoded" \
  --data-urlencode "grant_type=client_credentials" \
  --data-urlencode "client_id=${CLIENT_ID}" \
  --data-urlencode "client_secret=${CLIENT_SECRET}" \
  "${BASE_URL}/knoxtoken/api/v2/token")

echo $TOKEN_RESPONSE | grep -o '"expires_in":[0-9]*' | cut -d':' -f2 | awk
 '{print int($1/1000)}' | xargs -I{} date -r {} "+Expires: %Y-%m-%d %H:%M:
%S %Z"
```

The expiry line confirms the token was issued and shows how long it is valid. For example: Expires: 2026-03-20 03:58:25 CET.



Important: If the response is an HTTP 401 with Unknown token, the Client ID is not recognized in this Knox topology. Verify that the Data Share is in Shared status and the correct credentials file was used.



Tip: Use the complete CLIENT_SECRET value from the credentials file. A truncated or edited secret often results in an HTTP 400 response from the token endpoint instead of a JWT.



Important: knox.token.ttl controls the lifetime of your Knox access token received in exchange for the Client ID and Secret of your external users. Once it expires, external users need to request a new access token. For more information, see Editing Knox topologies.

```
Expires: 2026-03-20 04:12:38 CET
```

A JSON response containing access_token and other fields confirms the credential exchange was successful.

3. Extract the access token value into a variable:

```
ACCESS_TOKEN=$(echo $TOKEN_RESPONSE | grep -o '"access_token":"[^"]*"' |
 cut -d'"' -f4)
```

Verify the token is a clean JWT string with no surrounding JSON:

```
echo $ACCESS_TOKEN
```



Important: The value must be only the JWT string — three dot-separated base64 segments starting with eyJ. Do not paste the full JSON response. If ACCESS_TOKEN is empty, check that TOKEN_RESPON SE contains a valid JSON response from the previous step.

```
eyJqa3UiOiJodHRwczovL2Nsb3VkZXJhLWNhdGFs...
```

4. Verify namespace access by listing all namespaces visible to the Data Consumer:

```
curl -skS -u "Token:${ACCESS_TOKEN}" \
  "${BASE_URL}/iceberg-rest/v1/namespaces"
```

Confirm that the namespace containing your shared Iceberg table appears in the response.

```
{"namespaces":[["airline_operations"],["default"],["information_schema"],
["sys"]]}
```

5. List the tables in the target namespace to confirm the shared table is accessible:

```
TARGET_NAMESPACE="[***TARGET-NAMESPACE***]"

curl -skS -u "Token:${ACCESS_TOKEN}" \
  "${BASE_URL}/iceberg-rest/v1/namespaces/${TARGET_NAMESPACE}/tables"
```

Verify that the shared Iceberg table name appears in the identifiers list of the response.

```
{"identifiers":[{"namespace":["airline_operations"],"name":"enriched_fli
ght_data"}]}
```

6. Retrieve the table metadata to confirm schema access:

```
TARGET_TABLE="[***TARGET-TABLE***]"

curl -skS -u "Token:${ACCESS_TOKEN}" \
  "${BASE_URL}/iceberg-rest/v1/namespaces/${TARGET_NAMESPACE}/tables/${TAR
GET_TABLE}"
```

The response contains the full Iceberg table metadata. Confirm that the metadata.schemas[0].fields array lists the expected columns, each with a name, type, and required field.

```
{"metadata":{"schemas":[{"fields":[
  {"name":"flight_id","type":"string","required":false},
  {"name":"passenger_name","type":"string","required":false},
  {"name":"seat_assignment","type":"string","required":false},
  {"name":"aircraft_model","type":"string","required":false},
  {"name":"origin_airport","type":"string","required":false},
  {"name":"destination_airport","type":"string","required":false},
  {"name":"passenger_capacity","type":"integer","required":false},
  {"name":"age","type":"integer","required":false},
  {"name":"passport_number","type":"string","required":false}
]}]},...}
```

A JSON response with populated metadata and metadata-location fields confirms that the Data Consumer can read the table schema. The Data Share is correctly set up and externally accessible.

Results

The external consumer's Client ID and Secret are successfully exchanged for a Knox access token, and the Cloudera Iceberg REST Catalog confirms access to the shared namespace, table, and schema. The Data Share end-to-end verification is complete.

What to do next

Once verified, the Data Consumer can configure their Iceberg REST Catalog compatible compute engine using the same Client ID and Secret. For an example using PySpark, see the Related information.

Related Information

Creating a new Data Share
Creating external users
Sample Spark workload to access data
Creating a Data Share with CDP CLI