

Data Catalog Release Notes

Date published: 2019-11-14

Date modified: 2024-06-03



Legal Notice

© Cloudera Inc. 2024. 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

What's new.....	5
June 03, 2024.....	5
May 16, 2024.....	5
April 3, 2024.....	5
Older releases.....	6
Data Catalog releases for 2023.....	6
December 15, 2023.....	6
March 13, 2023.....	6
January 17, 2023.....	6
Data Catalog releases for 2022.....	7
April 1, 2022.....	7
March 2, 2022.....	7
February 16, 2022.....	7
Data Catalog releases for 2021.....	7
December 9, 2021.....	7
October 26, 2021.....	7
August 27, 2021.....	8
October 5, 2021.....	8
July 26, 2021.....	8
May 31, 2021.....	8
April 5, 2021.....	9
March 25, 2021.....	9
March 15, 2021.....	9
Data Catalog releases for 2020.....	9
December 14, 2020.....	9
October 29, 2020.....	9
August 27, 2020.....	10
August 12, 2020.....	10
July 25, 2020.....	10
June 30, 2020.....	10
June 17, 2020.....	10
June 5, 2020.....	11
May 6, 2020.....	11
April 7, 2020.....	11
January 6, 2020.....	11
Data Catalog releases for 2019.....	12
November 14, 2019.....	12
September 23, 2019.....	12
August 22, 2019.....	12

Known issues..... 12

Fixed issues..... 14

What's new

This section lists major features and updates for the Data Catalog service.

Cloudera Data Catalog is now additionally supported in the EU and APAC regional Control Plane. You can now ensure that you manage, secure, collaborate, and govern data assets across multiple clusters and environments within the EU region or APAC region where your organisation operates as per the data protection regulatory requirements.

June 03, 2024

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

This release of the Data Catalog service introduces the following new changes:


This release only contains fixes and updates to prepare Data Catalog for the changes in the upcoming 3.0.0 release.

May 16, 2024

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

This release of the Data Catalog service introduces the following new changes:

Iceberg tables are now supported by the Data Catalog service:

- You are able to filter for them in the Search page.
- Iceberg tables can be viewed in the Asset Details page.
- Iceberg tables can be added to a dataset.
- All subcomponents of Data Catalog support JDK 17.
-  **Important:** Profilers do not support Iceberg tables in this release.

April 3, 2024

This release of the Data Catalog service provides you with a notable behavior change which you must note and act accordingly.

While upgrading your cluster from Cloudera Runtime version **7.2.17** to **7.2.18**, and specifically during the OS upgrade step, the cluster goes into the failure state. The following message is seen:

__NODE_FAILURE:

New node(s) could not be added to the cluster. Reason Please find more details on Cloudera Manager UI. Failed command(s): Start(id=1546339088): Failed to start role profc6cf3856-PROFILER_SCHEDULER_AGENT-484032cb8f17cacf9e684efe50 of service profiler_scheduler in cluster cdp-dc-profilers-258395ef_

Impact on Data Catalog profilers:

If the Data Hub is not created, then the Data Catalog profilers will not be created in Cloudera Runtime 7.2.18 version.

To overcome this scenario, you must use the following process to bring up the Data Catalog profilers in the Cloudera Runtime 7.2.18 version.

First you must delete your existing 7.2.17 clusters. For more information, see [Deleting profiler cluster](#).

Next, after you upgrade to the 7.2.18 Data Lake, then you can launch the Data Catalog profilers. For more information, see [Launch profiler cluster](#).



Note: There is no data loss expected on the users' side or the Profiler analysis. However, the only loss that could be expected is related to the last runtime value of the profiler and the profiler run history. The Profiler Last Runtime history refers to the records of how many runs of the profiler are displayed on the history page. It includes information on whether the runs were completed successfully or resulted in failures.

Older releases

Overview of new features, enhancements, and changed behavior introduced in earlier releases of Data Catalog.

Data Catalog releases for 2023

Release summaries:

December 15, 2023

This release of the Data Catalog service introduces the following new features. Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following features as generally available:

Data Catalog is now integrated with AWS Glue Data Catalog which enables the users to browse and discover data as well as register data into SDX (using metadata translation or copy), so that it can be used with Data Hubs and other relevant experiences.

March 13, 2023

This release of the Data Catalog service introduces the following new features. Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following features as generally available:

- Supporting High Availability for profiler services.
- Navigating to tables and databases in Hue using Data Catalog.
- Using the Download CSV option to save queried results.

January 17, 2023

This release of the Data Catalog service introduces the following new features. Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Cloudera Data Catalog is now additionally supported in the EU and APAC regional Control Plane.
- Associating Customer Sensitivity Profiler tags to Hive assets fails on the 7.2.16 Cloudera Runtime cluster. You must update the configurations. For more information, see [Associating Custom Sensitivity Profiler tags to Hive assets](#).

Data Catalog releases for 2022

Release summaries:

April 1, 2022

This release of the Data Catalog service introduces the following new features. Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

Profiler cluster launch requires specific instance type of the Master node.

March 2, 2022

This release of the Data Catalog service introduces the following new features. Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Support for Cloudera Runtime 7.2.14
- No end-user manual process needed for backing up and restoring profiler data
- Post-upgrade (to Cloudera Runtime 7.2.14 version), deleting profiler cluster does not require saving the modified Custom Sensitivity Profiler rules

February 16, 2022

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Deleting profiler cluster with multiple scenarios
- Bug fixes

Data Catalog releases for 2021

Release summaries:

December 9, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Deleting profiler cluster
- Bug fixes

October 26, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

Available location paths for the Backup and Restore scripts based on the Cloudera Runtime version.

August 27, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

- Bug fixes. Refer to the Fixed Issues section in the Release Notes document.
- Refreshed screenshots for profiler configuration and related pages.
- Profiler cluster creation using the new UI
- The option selected in Depth drop-down and Show Process nodes are now displayed on the upper-left corner of the Lineage module.

October 5, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

- Support for deleting the profiler cluster
- Updated asset details
- Bug fixes and refreshed UI

July 26, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

- Creating classifications on the search page or on the classification widget.
- Bug fixes. Refer to the Fixed Issues section in the Release Notes document.
- Updated in-help documentation for multiple components in the UI.

May 31, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

- Launching Profilers using the Command-line Interface (CLI).
- Bug fixes. Refer to the Fixed Issues section in the Release Notes document.

April 5, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

Support for Business Glossary and usage of "terms" to associate datasets.

March 25, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following addition:

Backing up and Restoring Profiler Database.

March 15, 2021

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Searching for asset types using additional options in Data Catalog and Atlas
- Bug fixes

Data Catalog releases for 2020

Release summaries:

December 14, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Support for AWS and Azure in Cloudera Runtime 7.2.6 version
- Atlas entity audits in the Asset Details page
- Profiler jobs tracking status
- Audit graphs for HBase tables

October 29, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Support for searching assets across multiple data lakes.
- A new tab called Contents is introduced to list the contents of the select entity.
- Support for tagging multiple assets.

August 27, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Multiple entity type selections to search for assets
- Other UI fixes and improvements

August 12, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Multiple filter type selections on the Search page
- Other UI fixes and improvements

July 25, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Support for improved user experience for the search page
- Atlas tags option is now directly accessible from the Data Catalog menu bar

June 30, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following additions:

- Tag management
- On-Demand profiler support for Hive Column Statistics Profiler and Cluster Sensitivity Profiler
- Last Run Check Configuration
- Support for adding Hive asset to one or more datasets from the Asset Details screen

June 17, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following changes and bug fixes:

- UI improvements for search capabilities
- Ability to search Hive and HBase assets using Column tag

June 5, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following changes:

- About accessing Data Lakes
- Support for repopulating the Asset owners
- Support for selecting more entity types for searching assets

May 6, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following changes and bug fixes:

Navigation Support for Hive entity within Lineage.

April 7, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following new features:

- Profiler Tag Rules
- Hive Column Statistical Profiler

January 6, 2020

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments.

Data Catalog introduces the following new features:

- Ability to run Sensitivity Profiler jobs
- Ability to collaborate with other users - you can now rate asseptions, add comments about asset collections and respond to the comments of other users.

Data Catalog releases for 2019

Release summaries:

November 14, 2019

This release of the Data Catalog service introduces the following new features.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments (on-premises, cloud, and IOT).

Data Catalog introduces the following new features:

- Ability to view Profiler jobs
- Ability to edit Profiler configuration
- Ability to toggle a Profiler on and off.

September 23, 2019

This release of the Data Catalog service introduces the following new features:

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments (on-premises, cloud, and IOT).

Data Catalog enables data stewards across the enterprise to work with data assets in the following ways:

- Organize and curate data globally
- Understand where relevant data is located
- Understand how data is interpreted for use
- Understand how data is created and modified
- Understand how data access is secured and protected, and audit use

August 22, 2019

This release of the Data Catalog service introduces the following new features:

This is the first release of the Data Catalog service.

Data Catalog is a service within Cloudera Data Platform that enables you to understand, manage, secure, and govern data assets across enterprise data lakes. Data Catalog helps you understand data across multiple clusters and across multiple environments (on-premises, cloud, and IOT).

Data Catalog enables data stewards across the enterprise to work with data assets in the following ways:

- Organize and curate data globally
- Understand where relevant data is located
- Understand how data is interpreted for use
- Understand how data is created and modified
- Understand how data access is secured and protected, and audit use

Known issues

Learn about the known issues in Data Catalog, the impact or changes to the functionality, and the workaround.

Known Issues

CDPDSS-1953: Location field does not load absolute path for entities in contents tab

Problem: After navigating to the Azure Blob which is created within the Azure directory and later clicking the "Content" tab, the path for Azure blob is displayed as "/" instead of the path that mentions: containers/directory/Blob.

CDPDSS-1956: Issue with count on search page as AWS S3 V2 Object

Problem: The count against the Data Lake displays only the count of AWS S3 V2 Object entity type.

CDPDSS-31173: Atlas API on saving the classifications in Data Catalog is failing while using Cloudera Runtime version 7.2.12

Problem: When the suggested classification by Custom Sensitivity Profiler is saved, Atlas API throws an error response and it is not saved in Data Catalog. This behavior is specifically observed in Cloudera Runtime version 7.2.12.

Workaround: If you are using 7.2.12 version, from Cloudera Manager > Atlas > Configuration > search for 'application.prop' > enter the value as "atlas.tasks.enabled=false". Restart Atlas. Later, log into Data Catalog to use the classification feature.

CDPDSS-2127: On profiler cluster creation step 'external database in creation' the message on search page is not displayed

Problem: The message on Data Catalog search page is not displayed in 'blue' section saying "profilers cluster is getting created".

CDPDSS-2138: ODP fails on an asset imported from Glue to CDP on a medium duty cluster

Problem: The profiler run fails on the table.

CDPDSS-2142: Livy job in Profiler run has failed but profiling status of assets is shown as started

Problem: The livy job fails but each assets shows profiling status as started.

CDPD-32000: Atlas intermittently does not send the classification information to Data Catalog

Problem: On loading of search page filters with 'entity type' filter applied to 'Hive', intermittently the tags/classifications are not loaded in Data Catalog App. Due to this, Data Catalog does not load the filters for 'entity tag' and column tag.

CDPDSS-2145: Applying column tag filter and typing the same search result in search field gives 'no results' found

Problem: On searching the asset which is in the results for a column tag filter, the search results are shown as 'No results found'.

CDPDSS-2236: Issues noticed on data-catalog while performing scale testing

Problem:

- Search page consistently throws 504 gateway time-out error, however on applying global search field or search filters, results are loaded.
- Hive Column profiler goes to "undetermined" state for both ODP or scheduled jobs. To fix this, Profiler must be stopped and restarted. Happens due to Profiler not completing the read events from eventstore in AWS S3 within the time-out limit.(intermittent)
- On Asset Details page, when ODP is triggered, on successful completion of jobs and refresh of page, the ODP widget is not displayed and user needs to refresh a couple of times to load it on UI (reproducible). This ODP trigger job is running in parallel to running scheduled jobs.
- Assets cannot be added to dataset as the API fails with gateway timeout (intermittent)

Workaround: For the first and fourth issues listed, the following configuration change done on Cloudera Manager for Atlas:

atlas.graph.index.search.types.max-query-str-length=1536

Heap size = 4 GB

For the third issue listed, you must refresh the Asset Details page for about three to four times.

Fixed issues

This section lists the issues that have been fixed since the previous version.

CDPDSS-81: Data Catalog Issue

Problem: If an error occurs while fetching results, an error message appears briefly and disappears. Once the message disappears, the results pane is stuck in the Loading state.

CDPDSS-81: Data Catalog Issue

Problem: Despite having no privileges to add, modify, or delete classifications, users are able to edit the tag flow on the Asset Details page. However, users will not be able to save the changes.

CDPDSS-81: Data Catalog Issue

Problem: On the Asset Details page, the Audit tab fails to load if the logged in user does not have Audit read privilege.

CDPDSS-363: The 403 error message seen in Data Catalog > Table / Audit tab

Problem: Error messaging needs to be improved for users accessing Table and Audit tab without sufficient Ranger role.

CDPDSS-437: Unauthorized message for tag management flow does not have column name

Problem: Unauthorized column names and comma visible for tag management flow.

CDPDSS-444: Duplicate tag based policy on Asset Details Policy tab

Problem: When a specific tag is present both on the table as well as on any column of the same table, and a couple of requests to Ranger with the same tag and it displays the policy twice on list.

CDPDSS-527: Data Catalog specific user roles not supported.

Problem: The following Data Catalog user roles are not supported:

- DataCatalogCspRuleManager
- DataCatalogCspRuleViewer

Adding or removing these roles in the User Management will not have any impact on Data Catalog user access.

CDPDSS-81: Data Catalog Issue

Problem: When you navigate to the Data Catalog page from the Environments page or Data Lakes page, the context is not retained and the Data Catalog page loads with the last used context.

Workaround: You must change the context manually by selecting the right data lake.

CDPDSS-302: Profiler Manager server requires a manual restart after installation for DP Tags to reflect in Atlas

Problem: Once the profiler is launched, run the sensitivity profiler, the profiler recognises the sensitive tags but fails to push them to Atlas.

Workaround: You must restart the profiler manager. Note that restarting the profiler manager is a one-time task.

CDPDSS-360: Sensitive and Hive Column Profilers do not support managed Hive tables.

Problem: Profilers cannot access the storage file of the managed table.

CDPDSS-517: Dataset View: Asset list displays loading icon if user is not authorized on Environment on which the dataset is created.

Problem: The dataset view page breaks if user does not have access to the Environment on which the dataset is created.

CDPDSS-570: The Filter type list should be refreshed based on the data lake selection.

Problem: Selecting the filter type from the drop-down does not display the supported entity type and also in the radio button list.

CDPDSS-547: Data lake name missing on the Asset Details page.

Problem: The data lake name was not visible on the Asset Details page .

CDPDSS-594: Database filter and create dataset button gets hidden

Problem: When you clear the query used the clear button, the databaser filter gets hidden and create dataset button on the search page gets disabled.

CDPDSS-636: User able to view assets of a dataset which is available in an inaccessible data lake.

Problem: Users are able to view the assets in a data lake for which they have no permission or access.

CDPDSS-672: On the Asset-Details page, when you make changes and save them, the Scheme page is refreshed, the Schema tab is disabled, and Overview tab becomes operational.

Problem: When you perform some operations on the Asset Details page and save the changes, the Schema tab gets inactive.

CDPDSS-727: CSP Custom Rule Validation is getting stuck in "Validation Running" status

Problem: The rule gets stuck in the Validation Running status.

Workaround: Keep sensitivity profiler enabled when running CSP rule validation.

CDPDSS-955: Data Catalog uses wrong attribute to read createTime of entities and misinterprets lastAccessTime attribute.

Problem: The Data Catalog does not match the Created On date with Atlas.

CDPDSS-949: Asset registration pertaining to GLUE integration is not possible.

Problem: Asset registration failing with GLUE integration.

CDPDSS-921: Table created few days ago and configured recently is getting backlisted.

Problem: When an asset created date is older than the filter applied for black-listing, the asset still gets blacklisted (not picked during the profiler run).

CDPDSS-1007: On-Demand Profiler remains in "in-progress" state if the profiler job fails.

Problem: The On-Demand Profiler remains in the "in-progress" state if the profiler job fails due to some exception (For example: Permission denied exception).

CDPDSS-144: Filter Results Count Mismatch

Problem: When you search and filter the results using Created Before, the count displayed in the filters panel is incorrect. But, the number of search results is accurate.

CDPDSS-144: Search icon within dataset does not work

Problem: When you try to search for an asset within a dataset and click the search icon, it does not yield any results.

Workaround: Press enter to view the search results after entering a string in the search box.

CDPDSS-143: Edit button appears after creating a new dataset

Problem: After creating a new dataset, when you visit the dataset, the Edit button appears instead of Add Assets.

Workaround: Refresh your browser and reopen the dataset. You will be able to see the Add Assets button.

CDPDSS-81: When you select a table tag, owner filters along with the counts disappear

Problem: When you select a table tag, owner filters along with the counts disappear.

Workaround: Make sure you select table tags before selecting the owner filters.

CDPDSS-789: Adding system tag 'dp' prevents further editing of tags on the assets

Problem: Adding system tag 'dp' prevents further editing of tags on the assets

Workaround: Choose not to add the tag 'dp' on any of the assets.

CDPDSS-795: Search bar in schema tab goes empty, but previous results are shown

Problem: Search results displays previous results.

Workaround: Refresh the page.

CDPDSS-797: Schema search tab displays all results for not existing search string

Problem: Search by terms which have no hits and returns all assets.

CDPDSS-1055: Multiple tagging across different types throws 504 Gateway Timeout error

Problem: 504 Bad Gateway error is displayed.

CDPDSS-1075: When one specific asset is tagged to multiple tags from the search page, an incomplete alert is displayed

Problem: Asset name display is inconsistent. Only the first tag is displayed, remaining tag messages do not display the asset name.

CDPDSS-1084: Confusing terminologies.

Problem: Different usage terminologies. Not consistent.

CDPDSS-1373: "Unauthorised" message is displayed by middleware on the Data Catalog UI for every API failing 401

Problem: Data Catalog UI shows a pop-up at the bottom with message: "unauthorised request" and no context.

CDPDSS-1375: Search enhancements by TERM

Problem: In DataCatalog, if the term is included in the asset search query, the count of the assets next to the Data Lakes displays NA. It is because, Atlas search with Glossary term is returns approximate count as -1.

CDPDSS-1410: While navigating to another node from the Lineage tags filter, the tool-tip is retained.

Problem: From the lineage tags filter, the tool-tip is retained.

CDPDSS-1414: The "Cancel" button should exit the "Classifications" pane.

Problem: The "Cancel" button does not go away from the the "Classification" pane.

CDPDSS-1415: The Attribute name field in the "add classification" gets truncated when its very long and it cannot be read completely.

Problem: The attribute name field cannot be read fully.

CDPDSS-1416: Intermittent: Clicking add classifications reloads the lineage.

Problem: Clicking "add classifications" reloads the lineage but takes a very long before completely loading it.

CDPDSS-1474: [Glossary Terms] Difference in the Slider panel used in Asset Details and Search pages.

Problem: Indifferent display of slider panels in Asset Details and search pages.

CDPDSS-1475: Glossary Term | Allow selection on clicking any part of the row.

Problem: While associating a term to column or entity, the user has to click on the + icon for selection and not just tap on any object within the row.

CDPDSS-1484: Atlas and Ranger services UI link for CDP Private Cloud Base clusters is broken.

Problem: The managerUri attribute in Classic Cluster does not contain the Knox gateway URL.

CDPDSS-1498: For AWS Glue assets, users cannot add terms.

Problem: An error is displayed when Glue asset user tries to add terms.

CDPDSS-1501: On applying the term name in the Glossary filter on the search page, the count for data lakes is rendered as NA

Problem: The result count for data lakes displays are NA but the result list loads correctly.

CDPDSS-1664: Inherited attributes are not getting displayed for tags in the Schema tab slider for specific user cases.

Problem: Missing attributes in the Schema tab slider search bar.

CDPDSS-1179: Data Catalog search via Atlas's /v2/search/basic API doesn't retrieve results when the search text mentioned in the entity filter criteria (like searching by DB / table name) has special characters (like + - & | ! () { } [] ^ " ' ~ * ? :)

Workaround: No known workaround via Data Catalog.

CDPDSS-1346: Asset Details view of HiveTable not displaying all columns

Workaround: N/A

CDPDSS-1276: Error is thrown when advanced configuration values are changed to 0 or left blank

Problem: Error is thrown when advanced configuration values are changed to 0 or left blank and saved

CDPDSS-1506: For specific databases, the terms associated to them are not getting deleted

Workaround: N/A

CDPDSS-1595: Data Catalog Profilers are not functioning as expected for Ranger Authorization Service (RAZ) enabled Data Lakes

Problem: When you are using the RAZ enabled clusters, you can launch the Data Catalog profilers but the profilers do not function as expected.

Workaround: To enable RAZ with Azure, you must update the existing Ranger policy and also add a couple of new Ranger policies:

- Existing Policy: Default: Hive warehouse locations
Add dpprofiler user and assign Read + List permission.

Later, add the following policies:

- Policy Name: Profiler locations (Anything that Ranger can accept)

Depending on your storage location, note the following and set them accordingly:

If you have a sample location like:

```
abfs://<Storage account>@<Storage account container>.dfs.core.windows.net/<sub_directory>/dpprofiler
Storage account:<Storage account>
Storage account container : <Storage account Container>
Relative path: /*/dpprofiler
```

Select All permissions to allow conditions to take effect.

- Policy Name :Profiler Ranger Audit Logs (Anything that Ranger can accept)

Depending on your Storage location, note the following and set them accordingly:

If you have a sample location like:

```
abfs://<Storage account>@<Storage account container>.dfs.core.windows.net/
<sub_directory>/ranger/audit
Storage account:<Storage account>
Storage account container : <Storage account Container>
Relative path: /*/ranger/audit
```

Select Read + List permissions for allow conditions.

Restart the Profiler Manager service from DataHub Cloudera Manager instance.

CDPDSS-1748: Release specific features call fails with 404 error on the search page

Problem: Release specific features call is failing with 404 error for data lakes with no profiler cluster.

CDPDSS-1774: In the edit dataset page, on selecting all the assets and deselecting a few does not work

Problem: In the edit dataset page, on selecting all the assets and deselecting a few of the assets, deselects them in the UI but on hitting the 'delete' (trash icon), all the assets (selected and unselected ones) are deleted.

CDPDSS-1801: Options drop-down for a dataset stays open even though drop-down is opened for another dataset

Problem: The drop-down for all the datasets stay open.

CDPDSS-1807: Overlapping b/w icon and inline text in text boxes

Problem: This overlapping of icons and inline text is removed from all the pages.

CDPDSS-1808: UI Issues with new 'add asset' modal to add assets to dataset N/A

Problem: Directly clicking on 'search' field of 'Table Tag' shows no results found.

Typing "_" in the search field erases the entire search string

'Clear' text button is not working for Table tags , however it works for column tags

CDPDSS-1897: Show process nodes option is displayed for entity types which do not have process nodes

Problem: The toggle option to show the "process nodes" is present even for entity types which do not have process nodes.

Workaround: N/A

CDPDSS-1954: User cannot delete profilers when no entities are displayed in search result

Problem: The "Action" button containing the option to delete profilers is also removed when the search results yield no entity type.

CDPDSS-1988: Asset names are not clickable for registered assets in Glue Data Lake

Problem: Once the asset is registered observe that the 'asset name' is in blue but not clickable.

CDPDSS-2125: Unable to create profiler cluster for Cloudera Runtime version 7.2.11 and below

CDPDSS-2130: Unregistered Glue tables consume more time to load Asset Details page.

CDPDSS-2045: Duplicate API calls in Profiler job listing screen

CDPDSS-2047: Azure | Profilers HA | Issue with DB connection

The issue can be solved by upgrading the data lake and the profiler cluster to 7.2.14 or later.

CDPDSS-2032: Profiler Manager Service call is failing for Classic Clusters

CDPDSS-2120: Test run on CSP tag creation fails with 404 error

CDPDSS-2140: System tags status is marked as 'suspended' on a medium duty cluster with Profiler Management Services on SDX

Problem: The system tags are in a suspended state instead of default state which is 'deployed'.

CDPDSS-1965: AWS S3 V2 Directory type nodes could not be clicked in the Asset Details page

The issue was fixed by enabling AWS S3 V2 Directory type nodes in a configuration file.

CDPDSS-2141: Delete profiler option should be available in Data Catalog UI for all the states

Problem: This functionality is present on CloudBreak UI and hence can be made available from Data Catalog App as well.

CDPDSS-2839: Hadoop version upgraded from 7.2.17.200-29 to 7.2.17.301-3

Problem: The Hadoop version is upgraded from 7.2.17.200-29 to 7.2.17.301-3 to remediate CVE-2023-44487 and other CVE fixes are implemented as well.

CDPDSS-2925: The getNextScheduledRun API fails if createtime value is in the Allow List.

The issue was fixed by updating the Data Catalog Profiler configuration.

Workaround: Remove asset filter rules with createTime filter and recreate them.

CDPDSS-2926: Profilers page displayed Start Time and Last Updated time in epoch time format

The Profilers page was updated to use the UTC time format.

CDPDSS-2928: Refreshing the Profiler page deletes configuration details

The issue was fixed by updating the Profiler page.

CDPDSS-2929: The Asset Details page reports a 404 error for impala_process_execution type assets.

The issue was fixed by removing the unsupported lineage widget for impala_process_execution type assets.

CDPDSS-2934: Classification and Glossary filters are not loaded on search page intermittently

The issue was fixed by updating the entity search components.

CDPDSS-2948: 400 Bad Request is generated in the Asset details page for Table Type "VIRTUAL_VIEW"

The issue was fixed by adding a tooltip and disabling the ODP run button for the empty location.