



Collibra Data Intelligence Cloud
Collibra Protect

Collibra Data Intelligence CloudCollibra Data Governance Center - Collibra Protect

Release date: Thursday, November 3, 2022

Revision date: Thu Nov 03, 2022

You can find the most up-to-date technical documentation on our Documentation Center at
https://productresources.collibra.com/docs/collibra/latest/Content/to_collibra-protect.htm

Contents

| | |
|---------------------------------------|-------------|
| Contents | ii |
| About Collibra Protect | i |
| Install Collibra Protect | ii |
| Configure Collibra Protect | iv |
| Essentials for Collibra Protect | ix |
| Overview of Collibra Protect | xvii |
| Data protection standards | xix |
| Data access rules | xxv |
| Data source policies | xxxiv |
| Groups | xxxvi |
| Audit | xxxviii |
| Why rules or standards fail | xl |
| Reference documentation | xlix |
| Collibra Protect | lvii |
| About Collibra Protect | i |
| Install Collibra Protect | ii |
| Configure Collibra Protect | iv |
| Essentials for Collibra Protect | ix |
| Overview of Collibra Protect | xvii |
| Data protection standards | xix |
| Data access rules | xxv |
| Data source policies | xxxiv |
| Groups | xxxvi |

| | |
|-----------------------------------|---------|
| Audit | xxxviii |
| Why rules or standards fail | xl |
| Reference documentation | xlix |

About Collibra Protect

Collibra Protect is a capability of the Data Intelligence Cloud created to protect sensitive data and make it available, or partially available, to specified groups of users.

Collibra Protect solves the problem of protecting sensitive data in an organization. Different groups of people may need varying access levels to the same data set. With Collibra Protect, access rules and data protection standard capabilities allow you to grant access to individuals and protect sensitive information. These rules and standards with different data access levels are managed through the Collibra platform and pushed to the data source. Our aim is to promote a safe data-open culture in organizations.

The goal of Collibra Protect is to centralize and simplify access governance and remove the need of repetitive action and approval. Data access and privacy management promotes an ethical company standard giving permission to view information only to those that need it. Collibra Protect allows you to perform these actions accordingly.

An example use case of Collibra Protect is a data steward giving everyone access to a data set, but only allowing certain access to groups of people based on data categories. This is known as differential access. It is suggested that rules/standards are grouped together, for example by business processes, so you do not have to make a rule or standard for every data set.



Install Collibra Protect

This procedure guides you through a first time installation of Collibra Protect.

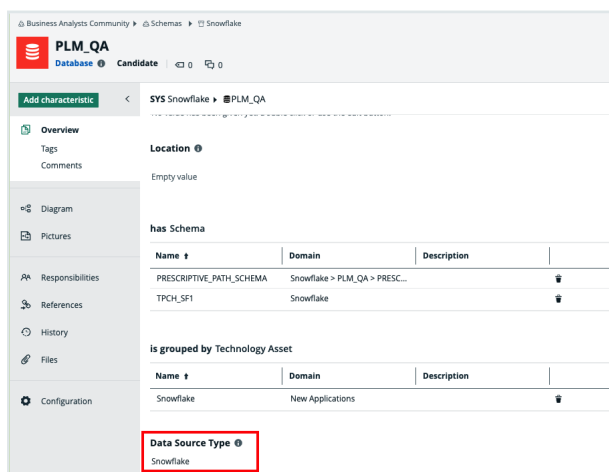
Prerequisites

You must add the [Snowflake capability on Edge](#) as well as perform a catalog ingestion.

Configure the Collibra Protect for Snowflake capability on Edge. Settings → (Edge) Sites → Your site → Capability → Add capability → fill in the needed parameters:

- For "Capability template" choose "Collibra Protect for Snowflake".
- The "Snowflake Connection" can be the same connection used for doing catalog ingestion. Make sure that the Snowflake user/role has enough permissions to create/alter/drop grants, tags, etc.

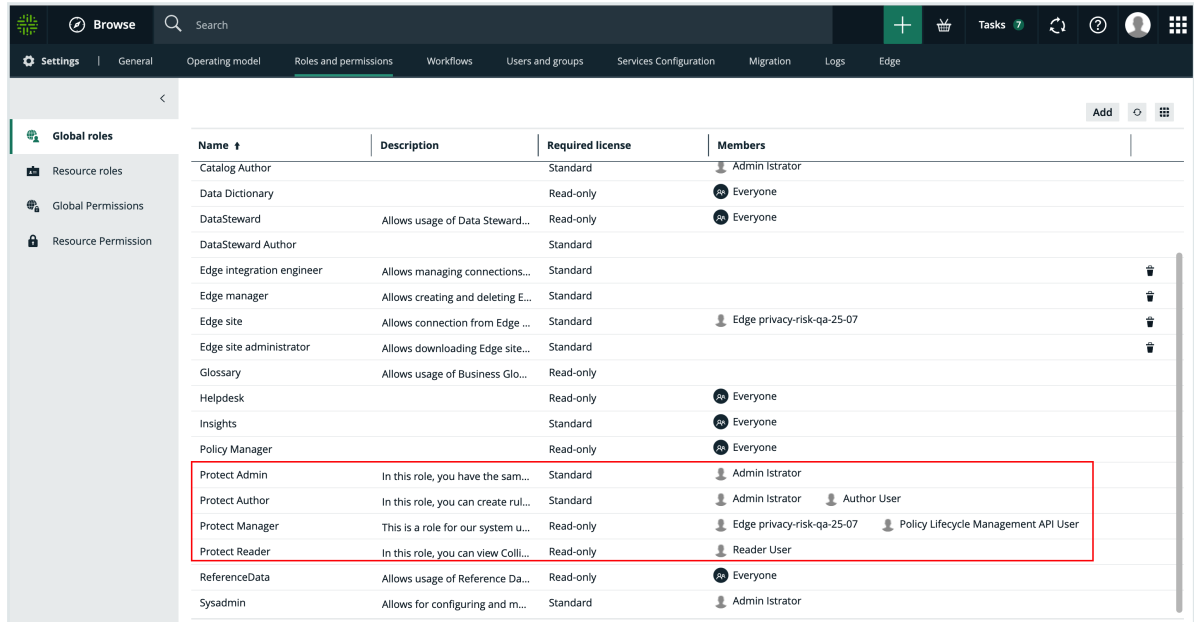
An ingested Snowflake database should look like the example below.




Note The Data Source Type attribute on the database asset should be present. This attribute is automatically added in database assets, after the catalog ingestion process.

Steps

1. Contact a Collibra support or your representative to enable Collibra Protect on your Collibra environment.
2. Ensure [global roles and permissions](#) for Collibra Protect are set correctly.



3. Collibra Protect is installed.
 - » You can now access and start using Collibra Protect via the  menu.

Configure Collibra Protect

Configuring within Collibra Protect is an important part of understanding and using Collibra Protect to its highest ability.

Prerequisites

- You need to have Data Catalog permissions. If not, you cannot see any classification in either standards or rules.
- You need to have a Data Steward role within Collibra. If not, you cannot see the classification page when selecting a classification in Collibra Protect.



Roles in Collibra Protect

It is possible to assign different roles to Collibra users that use Collibra Protect. The roles are provided and have pre-defined permissions that restrict the usage of the application.

| Roles | Description |
|----------------|---|
| Protect Reader | Users in this role can view Collibra Protect with read-only access to the content. This role is assigned to 'Everyone' and grants the users the 'protect' permission. Without this permission, users cannot see 'Protect' as an application in the ☰ menu. They also cannot navigate to protect related URLs or access protect endpoints. |
| Protect Author | Users in this role can create rules and standards , view imported policies and groups , and generate audits as an individual contributor. This role grants the product right permission 'protect' and the 'protect_edit' permission. Authors can only modify rules and standards they own. This role is not assigned to anyone automatically. |
| Protect Admin | Users in this role have the same permissions as the Protect Author role as well as the ability to edit other user's rules and standards. This role grants the product right permission 'protect', 'protect_edit', and an extra 'protect_administration' permission. This role is not assigned to anyone automatically. |

| Roles | Description |
|-----------------|--|
| Protect Manager | This role is restricted to our system user to manage background processes and setup configurations for Collibra Protect and it should not be assigned to other Collibra users. |

Configure groups

Before you start working in Collibra Protect, you need to configure your groups. Protect groups are the basis of all the actions performed in Collibra Protect.

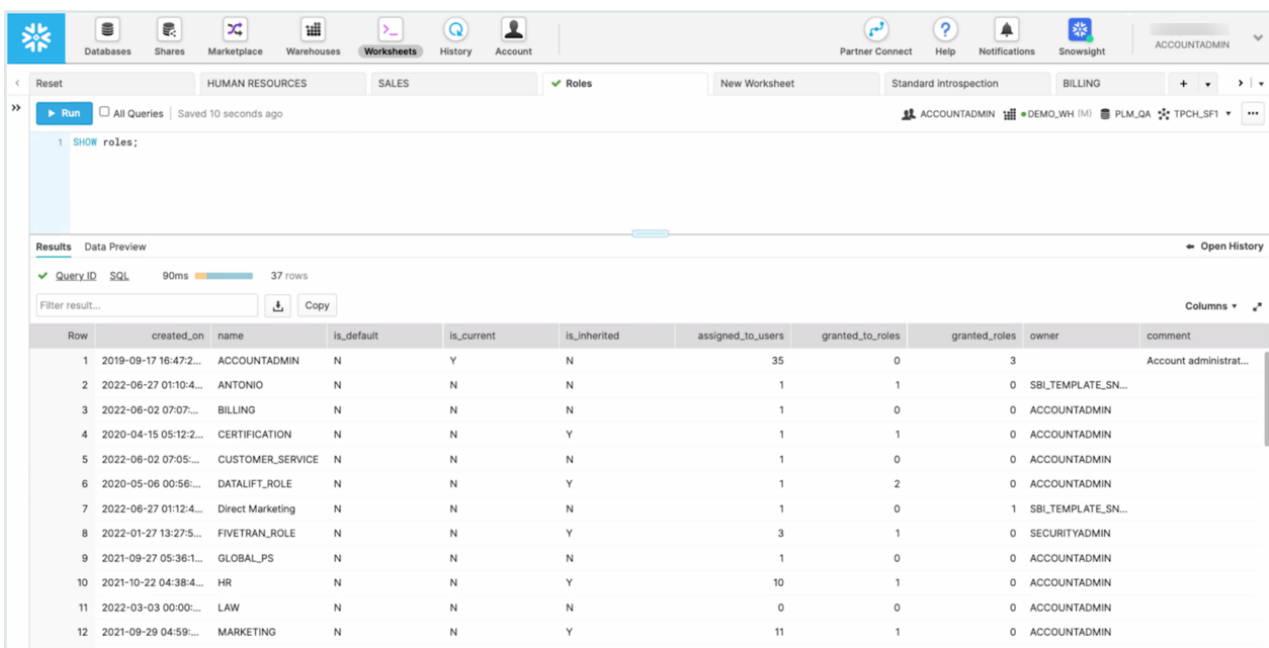
Associate a Protect group with Snowflake

Each Snowflake user is assigned to one or more Snowflake roles. Permissions are based on these roles. View the example below of the roles page in Snowflake. Any/all roles can be correlated to a Collibra Protect group.

| Role | Creation Time | Owner | Comment |
|------------------|------------------------|--------------------|--|
| ACCOUNTADMIN | 9/18/2019, 1:47:25 ... | | Account administrator can manage all aspects of the account. |
| ANTONIO | 6/27/2022, 10:10:4... | SBL_TEMPLATE_SN... | |
| BILLING | 6/2/2022, 4:07:43 ... | ACCOUNTADMIN | |
| CERTIFICATION | 4/15/2020, 2:12:24 ... | ACCOUNTADMIN | |
| CUSTOMER_SERVICE | 6/2/2022, 4:05:29 ... | ACCOUNTADMIN | |
| DATALIFT_ROLE | 5/6/2020, 9:56:54 ... | ACCOUNTADMIN | |
| Direct Marketing | 6/27/2022, 10:12:4... | SBL_TEMPLATE_SN... | |
| FIVETRAN_ROLE | 1/27/2022, 10:27:58... | SECURITYADMIN | |
| GLOBAL_PS | 9/27/2021, 2:36:19 ... | ACCOUNTADMIN | |
| HR | 10/22/2021, 1:38:44... | ACCOUNTADMIN | |
| LAW | 3/3/2022, 9:00:27 ... | ACCOUNTADMIN | |
| MARKETING | 9/29/2021, 1:59:26 ... | ACCOUNTADMIN | |
| MARKETING2 | 9/29/2021, 2:36:17 ... | ACCOUNTADMIN | |
| MARKETING3 | 9/30/2021, 3:56:47 ... | ACCOUNTADMIN | |
| PC_DBT_ROLE | 5/6/2022, 9:08:33 ... | ACCOUNTADMIN | System created role for partner elt integration. |
| PLM | 10/22/2021, 1:30:58... | ACCOUNTADMIN | |
| PLM_QA_HR | 2/24/2022, 3:38:20... | ACCOUNTADMIN | PLM QA HR Read Only Role |

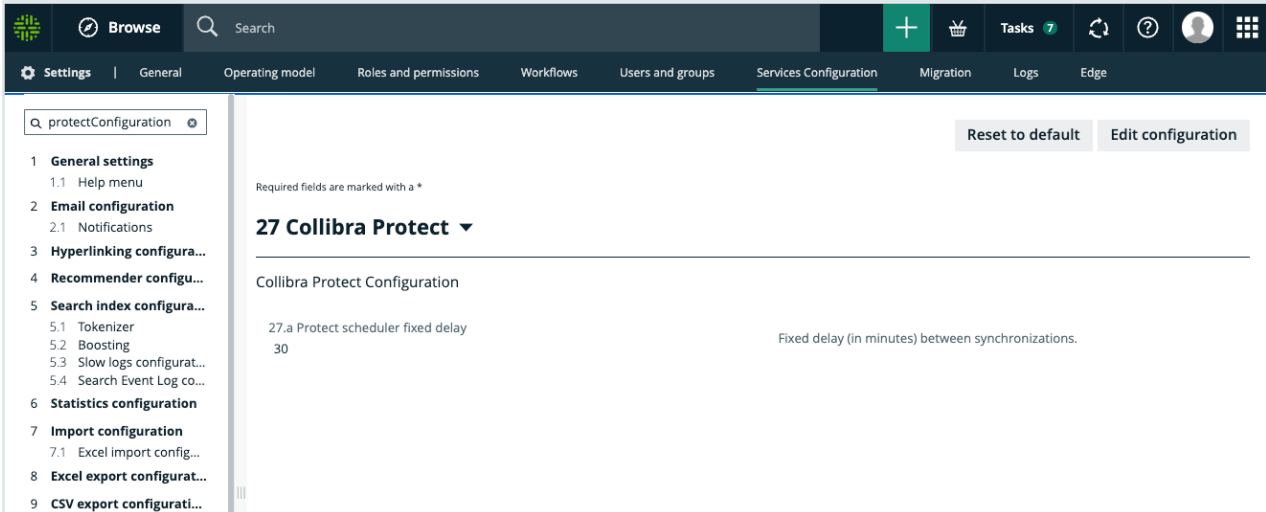
How to create Collibra Protect groups?

When you initially go to the **Groups** tab in Collibra Protect, there are no groups created. There is a link at the top of the page to the Groups API that creates new groups in Collibra Protect. Use this API link to create new groups and associate it with a specific role in Snowflake.



General configuration

Collibra Protect synchronizes standards and rules with the source database(s) at regular intervals. This synchronization runs in the background on a configured frequency. By default, the frequency is every 60 minutes, but this is configurable through Settings → Services Configuration → 27 Collibra Protect.



Important If you do not have access to the **Service Configuration** tab, create a support ticket requesting the JVM Parameter be added to your Collibra Infrastructure Configuration: `-DPROTECT_SYNC_SCHEDULER_DELAY=PT60M`. After the parameter is added, restart Collibra so these changes take effect and the policies are now synchronized with the cloud provider.

Synchronization includes:

1. Aggregate all standards and rules computing:
 - which columns need to be masked for which groups.
 - which tables need to have a row filter.
 - which tables and columns need to be granted access.
2. On the source database(s) such as Snowflake:
 - create and apply maskings.
 - create and apply row filters.
 - grant access to groups on tables and/or columns (depending on the underlying database).

Essentials for Collibra Protect

To use Collibra Protect to the best of its ability, you need to know the following things:

- [How to protect your data](#)
- [Technical background](#)
- [Data protection standards vs. data access rules](#)
- [Prescriptive paths](#)



How to protect your data

1. Access management

The most basic line of protection is to make sure only the right people/groups have access to the data. Data here is referring to the tables and columns in your database. In Collibra Protect, you can grant specific groups access to parts of your data based on Collibra assets.

For example, it is easy to grant the HR team access to the US customers' data set. But, what if some parts of the US customers' data set need to be hidden from the HR team, because it contains restricted information, such as personally identifiable information (PII)? In that case, you can further protect your data by applying column-based protection or row-based protection.

Note Collibra Protect only grants access. It cannot revoke access from people/groups.

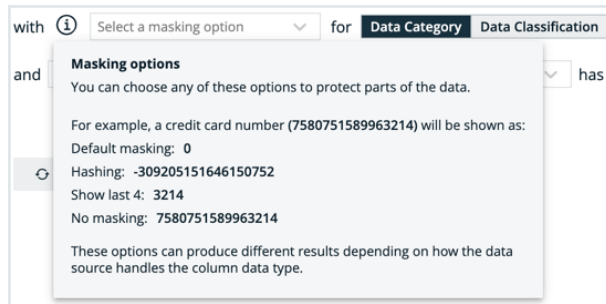
2. Column-based protection

Column based protection allows you to target specific columns and mask their content. By masking the column's data, the group cannot see the content as it is. They will see a masked version of it instead.

For example, you can mask a column of credit card numbers, so the individual group cannot see the full credit card numbers.

We currently support four masking options. They include:

- **Default masking:** Shows the value as 0.
- **Hashing:** Converts the value into a variety of different letters, numbers, and symbols.
- **Show last:** Displays the last letters, numbers, and symbols in the value. You can choose to show the last 1 through 20 of the value. The most common choice is Show last 4.
- **No masking:** Displays the data value as it is originally written.



Collibra Protect allows you to choose to mask columns that are part of a **data category** or a **data classification**. While granting access to a certain asset, you can choose to apply this masking on only a subset of that asset if it is also part of a data category or data classification.

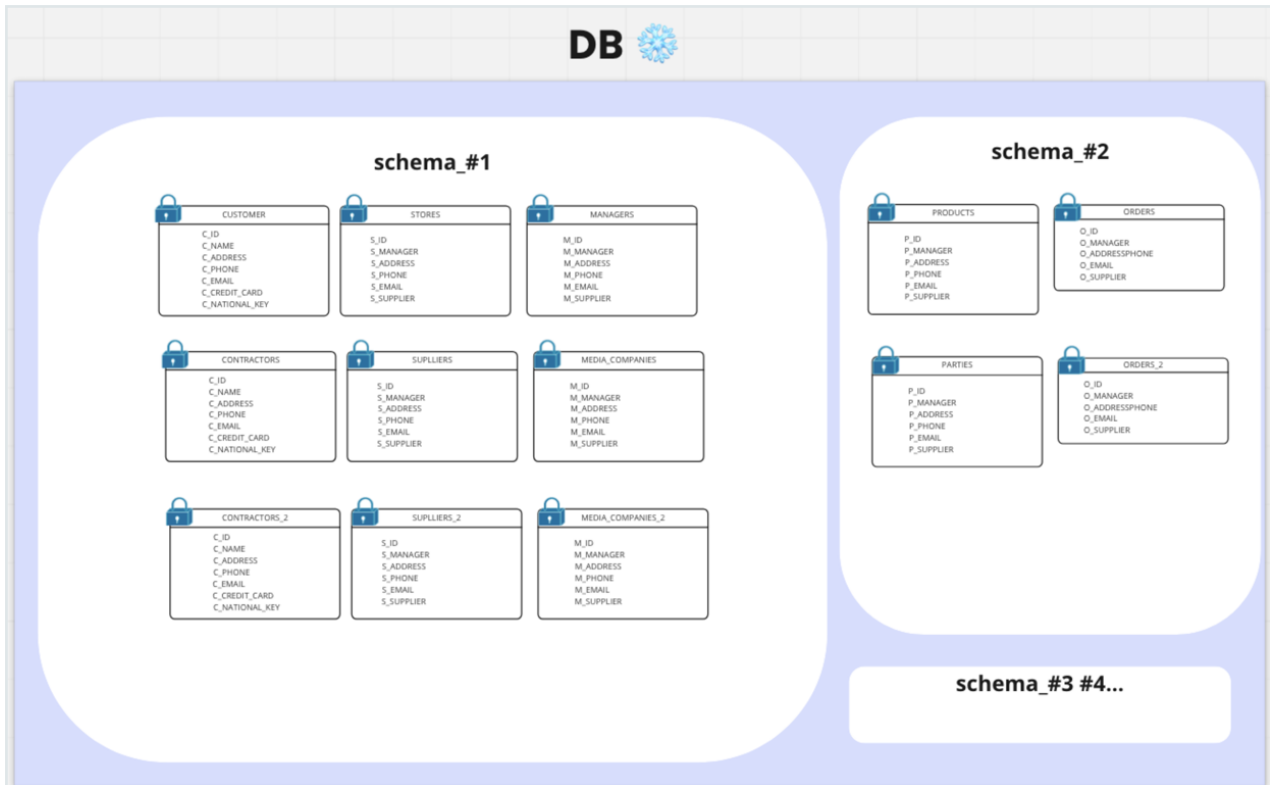
3. Row-based protection

Another way to protect your data is to filter rows of a specific table. If you do not want to expose all of the existing items in a table because one of the columns is part of a certain data classification, you can easily leverage the Collibra operating model to do so. When creating a rule that impacts certain tables in the source database, filter rows on tables by using the row filtering option for tables where one of their columns is part of a data classification. The filtering is based on what value is stored in the cell of that particular column. For instance, in a table that has a column that is classified as **country-code**, you can hide or show all items that have the value of **US**.

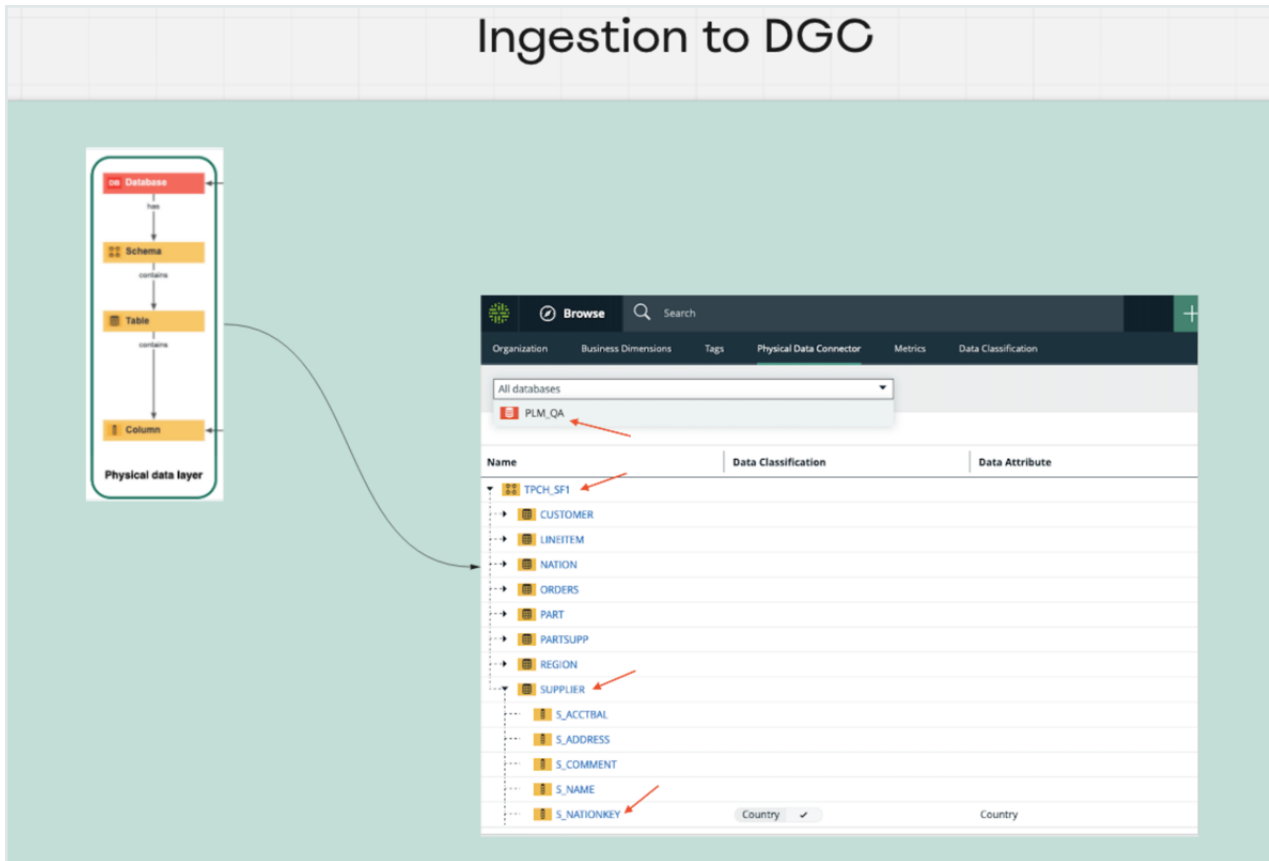
Technical background

The technical background of Collibra Protect explains the connection of the data as it is in the database (DB) with the physical layer (equivalent assets in Collibra Data Intelligence Cloud) and the logical layer (the out of the box model).

Imagine you have this database:



When ingesting this DB to Collibra Data Intelligence Cloud, the physical layer is created as well as an asset for each of the schemas, tables, and columns.



Once there is a physical layer established in our Collibra environment, start creating the logical layer on top of it.

- In this phase, take any column and classify it as any data classification available, or let the platform classify it for you.
- Also, assign a column to a data attribute.

From here, create additional assets or use existing assets of different types (data set, data category, or business process) to establish a relation to these columns.

Data protection standards vs. data access rules

Collibra Protect has both standards and rules to govern your data with ease and clarity.

| | |
|-------------------------|---|
| <p>Standards</p> | <p>Data protection standards create a layer of protection for similar types of data by masking them wherever they are.</p> <p>For example, if columns with first and last names are a part of the PII data category, regardless what tables, schemas, and databases they are part of, create a standard that targets all of these columns by choosing the PII data category and masking it.</p> |
| <p>Rules</p> | <p>After establishing this primary layer (blanket) of protection to your most sensitive data, use data access rules to manage access and enhance protection for specific usages.</p> <p>For example, create a rule that grants access to a specific group, for a specific data set, while knowing that all PII within this data set will be masked by the standard we created before.</p> |

FAQs

1. What if I want to grant access to a group without having the PII masked?
 - » When creating a rule for an asset that contains data masked by a standard, choose to override it by unmasking it or changing its masking type.
2. What If I want to grant access to a group, but the protection from the standard is not enough because there might also be other sensitive data within this supported asset?
 - » When creating a rule, add additional layers of protection over the ones that were set by any existing standard. Further protect the data by applying additional masking on or by filtering the data.

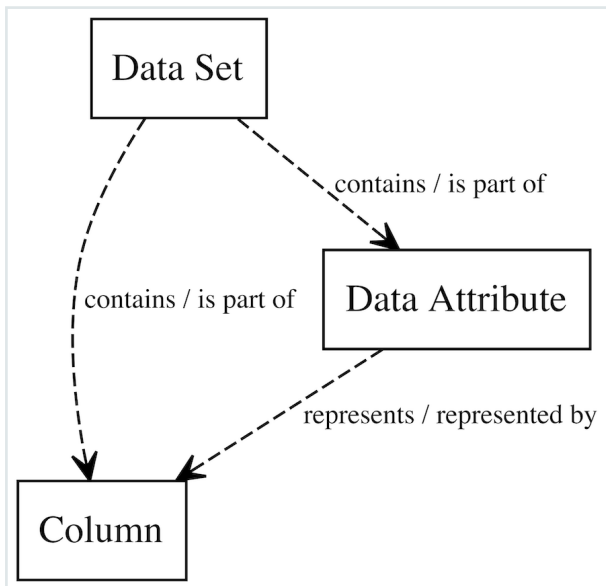
Prescriptive paths

When creating a standard or rule, you select which asset(s) you want to protect and/or grant access to. By default, you can grant access to a data set, a data category, and a business process. Colibra Protect searches the knowledge graph, through relationships and/or

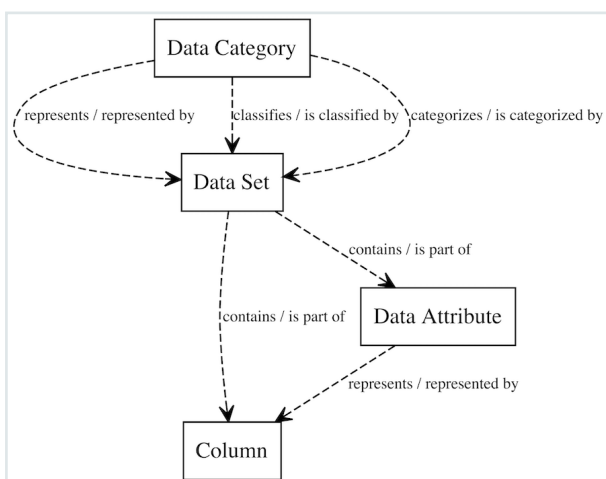
intermediate assets, to find which set of physical data layer assets, such as columns and tables, this resolves.

The traversal of the knowledge graph is done through a set of prescriptive paths. For each type of asset, there is a set of prescriptive paths to traverse to the column assets. See the images below for more details.

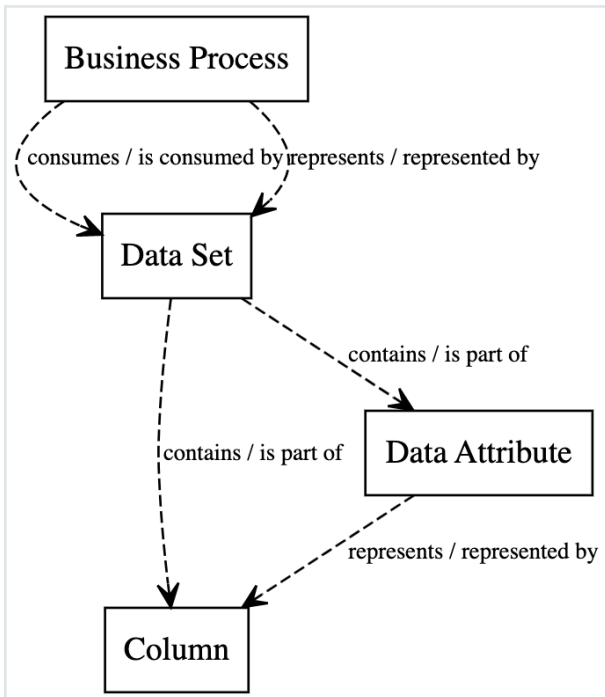
Prescriptive path for data set



Prescriptive path for data category



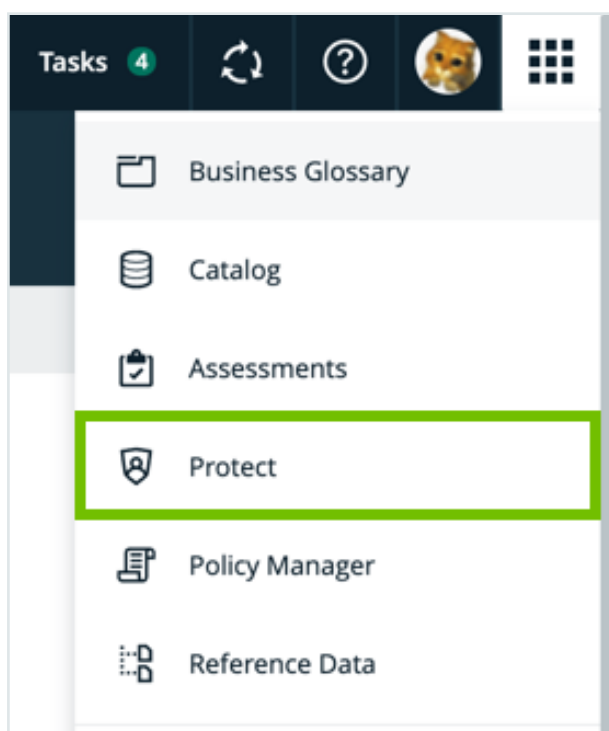
Prescriptive path for business process



Overview of Collibra Protect

To work with Collibra Protect, ensure that you have a global role that has the Protect global permission and that it is [enabled](#) in your environment.

You will find, Collibra Protect, in the main menu . Click **Protect**.



If Collibra Protect is not shown on the menu, the feature is not enabled.

The landing page displays five tabs at the top of the page: **Data Protection Standards**, **Data Access Rules**, **Data Source Policies**, **Groups**, and **Audit**.



Data Protection Standards

Data Access Rules

Data Source Policies

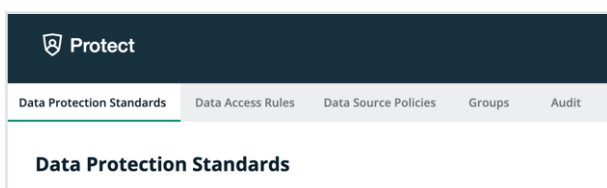
Groups

Audit

| Tab | Description |
|---------------------------|---|
| Data Protection Standards | <p>Define default data source access to data types based on data categories, data attributes, or classes/classifications through data protection standards</p> <p>Note Data access rules for particular groups can override created standards.</p> |
| Data Access Rules | Use data access rules to grant groups different access to the same data in data sets, in business processes, or identified by data categories. |
| Data Source Policies | View a list of policies that are currently active in the source data tables. You can also import policies from your source database using the Colibra Protect Data Source Policies API. |
| Groups | <p>Add groups through custom code via the Data Access API link and view existing current data access groups.</p> <p>Note You must add at least one group before you can create a standard or a rule.</p> |
| Audit | Generate an audit log for a preview of the last hour of ingested data from the data source. |

Data protection standards

The Data Protection Standards page contains an overview of the available standards in your environment.




| Page Section | Description |
|-----------------------------|--|
| Standards summary | Under the heading, there is a summary about data protection standards. Click the Create a Data Protection Standard button to create a standard and get started in Collibra Protect. |
| Recently Modified Standards | This section shows the five most recently modified standards. |
| Standards table | This table displays a detailed view of the created data protection standards. |

In the **Synchronization status** column of the standards table, there are five status options that can appear. To view the status of the standard in the data source, go to the source database.

| Synchronization Status | Description |
|------------------------|---|
| Active | This standard is currently active in Collibra Protect and in the data source. |



| Synchronization Status | Description |
|------------------------|---|
| Pending | This standard has been created or edited, and is pending synchronization. |
| Failed | The synchronization of this standard has failed. Click the  icon next to the failed status to view additional information about the error. |
| Delete Pending | This standard will be deleted from the data source in the next synchronization. |
| Not Deleted | The deletion of this standard has failed. |

Note Collibra Protect periodically synchronizes with the data source and statuses will be updated along with the synchronization. To learn more, go to the [general configuration](#) page.

Create a data protection standard

Data protection standards create a layer of protection by masking data wherever they appear. Create a data protection standard to get started using Collibra Protect.

Create a Data Protection Standard
✕

Data protection standards apply default data source access to types of data based on data categories or data classifications. Data Access Rules for particular groups will override these defaults.

Standard Name*

Description

for the group* + -

protect* Data Category Data Classification

with* ⓘ


Summary
 For the Group Human Resources
 protect Personal Information
 with Hashing

Cancel
Save Standard

Steps

1. In Collibra Protect, go to the **Data Protection Standards** tab.
2. Click the green **Create a Data Protection Standard** button.
 - » The **Create Data Protection Standard** dialog box appears.
3. Enter the required information. It is important to note that when selecting assets, user permissions are defined in Collibra. If an asset is not visible for you, it will not appear as an option in the drop down menus.

| Field | Description |
|---------------|-------------------------------------|
| Standard name | Name of the standard being created. |

| Field | Description |
|-------------------------------------|--|
| Description (optional) | Description of the standard. |
| Group | Group(s) for which the standard is created. |
| Data Category / Data Classification | A data category or data classification to apply the protection on. |
| Masking | Masking option for the standard. <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;"> <p>Note Click  to learn more about the masking options for standards.</p> </div> |

Note Click the plus sign to add more to each field where applicable. For example, after selecting a group, click + to add another group into the standard, and click – to delete a selected group. When entering the required information, you can view the selections you made in the **Summary** section.

4. Click the green **Save Standard** button.
 - » The saved data protection standard appears in the standards table.

Modify a data protection standard


You can edit or delete a data protection standard after it has been created.

Edit a standard

Editing a data protection standard might be necessary in certain situations. For example, change the masking method from default masking to hashing.

Important You will only be able to edit standard assets if you have view asset permissions. If one of the assets in the standard is unauthorized, you will not be able to edit the standard until the view access permission is granted.

Steps

1. In the standards table, click the standard name, and then click the **Edit** button or click  in the appropriate row
 - » The **Edit a Data Protection Standard** dialog box appears.
2. Edit the [required information](#).
3. Click the green **Save Standard** button.
 - » The updated data protection standard appears in the standards table.

Edit a Data Protection Standard ✕

Data protection standards apply default data source access to types of data based on data categories or data classifications. Data Access Rules for particular groups will override these defaults.


Standard Name *

Description

for the group * + -

and the group + -

protect * **Data Category** **Data Classification**

with * 


Summary
 For the Group Human Resources and Marketing
 protect [GDPR data related to criminal convictions and offences](#)
 with Default masking

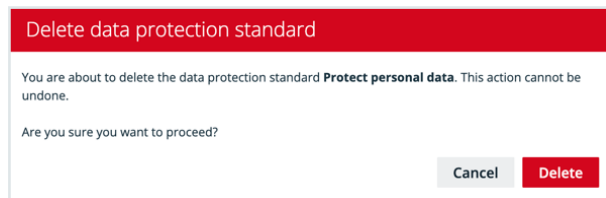
Cancel
Save Standard

Delete a standard

If you have an [author/admin role](#), delete a data protection standard that is no longer necessary.

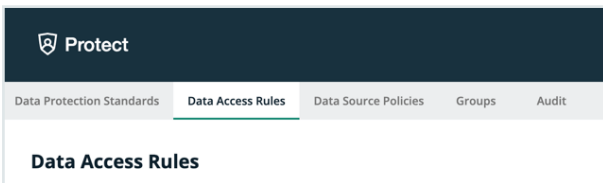
Steps

1. In the standards table, click the  icon in the appropriate row
 - » The **Delete data protection standard** dialog box appears.
2. Click the red **Delete** button.



Data access rules

The Data Access Rules page contains an overview of the available rules in your environment.




| Page Section | Description |
|-------------------------|---|
| Rules summary | Under the heading, there is a summary about data access rules. Click the Create a Data Access Rule button to create a standard . |
| Recently Modified Rules | This section shows the five most recently modified rules. |
| Rules table | This table displays a detailed view of the created data access rules. |

In the **Synchronization status** column, there are five status options that can appear. To view the status of the rule in the data source, go to the source database.

| Synchronization Status | Description |
|------------------------|---|
| Active | This rule is currently active in Collibra Protect and in the data source. |
| Pending | This rule has been created or edited, and is pending synchronization. |



| Synchronization Status | Description |
|------------------------|--|
| Failed | The synchronization of this rule has failed. Click the  icon next to the failed status to view additional information about the error. |
| Delete Pending | This rule will be deleted from the data source in the next synchronization. |
| Not Deleted | The deletion of this rule has failed. |

Note Collibra Protect periodically synchronizes with the data source and statuses will be updated along with the synchronization. To learn more, go to the [general configuration page](#).

Create a data access rule

After establishing a primary layer (blanket) of protection to your most sensitive data using standards, create data access rules to manage access to the data sources and enhance protection for specific usages.

Create a Data Access Rule
✕

Use data access rules to grant groups different access to the same data in data sets, business processes, or identified by data categories. You can mask or hide columns by their data category and you can also conditionally filter rows based on code set values.

Rule Name *
Marketing GI Rule

Description
Set rule for the marketing group for the geographic information asset
Apply default masking for genetic data

Set rule for

group * Marketing + -

asset * Geographic Information + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Default masking + -

for Data Category Data Classification Genetic data + -

and Select an action + -

rows where Unauthorized + -

has Select a code set + -

Select a code value + -

Summary
Grant access to Marketing
for [Geographic Information](#)
with Default masking for [Genetic data](#)

Generate Preview

Cancel
Save Rule

Steps

1. In Collibra Protect, go to the **Data Access Rules** tab.
2. Click the green **Create a Data Access Rule** button.
 - » The **Create a Data Access Rule** dialog box appears.
3. Enter the required information. It is important to note that when selecting assets, user permissions are defined in Collibra. If an asset is not visible for you, it will not appear as an option in the drop down menus.

| Field | Description |
|------------------------|--|
| Rule name | Name of the rule being created |
| Description (optional) | Description of the rule. |
| Group | Group for which the rule is being created. |
| Asset Name | Data asset that the rule is protecting. Protect enables you to protect the following asset types: Business process, data set, and data category. Learn more in technical background and prescriptive paths . |

| Field | Description |
|---|--|
| <p>Masking (optional)</p> <ul style="list-style-type: none">◦ Data Category / Data Classification | <p>Masking option for the rule. Click the i to learn more about masking options.</p> <ul style="list-style-type: none">◦ Select a data category or a data classification to apply masking to. |

| Field | Description |
|---|--|
| Action (optional) <ul style="list-style-type: none"> ◦ Data Classification ◦ Code Set ◦ Code Value | Filter the data by selecting hide or show. <ul style="list-style-type: none"> ◦ Select data classification that is either hidden or shown ◦ Code set to set up row filtering in the tables. A code set must be selected to filter by a code value. ◦ Code value of the code set selected. |

Important The grant access checkbox is selected by default. By leaving this checkbox selected, you are granting access to the tables in the database with columns linked to the selected assets to the selected group(s). If you do not want to grant this kind of access to these groups, clear the grant access checkbox.

Note Click the plus sign to add more to each field where applicable. For example, after selecting a group, click + to add another group into the standard, and click – to delete a selected group. When entering the required information, you can view the selections you made in the **Summary** section.

- Click **Generate Preview** to see a preview of the new rule.

Summary
Grant access to Marketing
for [Geographic Information](#)
with Default masking for [Genetic data](#)

Geographic Information ▾

| Column ↑ | Access | Masking Agent | Masking | Code Value |
|--|--------|---------------|---------|------------|
| C_ADDRESS_sdfxgxcfhcjnhjvbkjbjhgxdfzs... | Masked | Genetic data | 0 | |
| C_NAME | Masked | Genetic data | 0 | |
| DS_TBL0001_COL0001 | Masked | Genetic data | 0 | |

Tip Use the preview to verify the data access rule is set up correctly. The preview only shows the first 1,000 affected columns. The drop-down below the **Generate Preview** button is used to switch between the different selected assets in the rule. Each asset has its own preview table.

- Click the green **Save Rule** button.
 - » The saved data access rule appears in the rules table.

Modify a data access rule


You can edit or delete a data access rule after it has been created.

Edit a rule

Editing a data access rule might be necessary in certain situations. For example, change the code set value from BE to US.

Important You will only be able to edit rule assets if you have view asset permissions. If one of the assets in the rule is unauthorized, you will not be able to edit the rule until the view access permission is granted.

Steps

1. In the rules table, click the rule name, and then click the **Edit** button or click  in the appropriate row
 - » The **Edit a Data Access Rule** dialog box appears.
2. Edit the [required information](#).
3. Click the green **Save Rule** button.
 - » The updated data access rule appears in the rules table

Edit a Data Access Rule
✕

Use data access rules to grant groups different access to the same data in data sets, business processes, or identified by data categories. You can mask or hide columns by their data category and you can also conditionally filter rows based on code set values.

Rule Name*

Description

Set rule for

group* + -

asset* + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ for **Data Category** + -


and rows where has

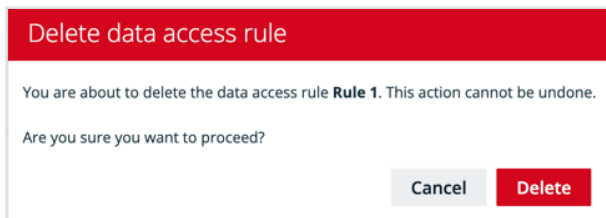
Summary
 Grant access to Marketing
 for [Customer Data](#)
 with Hashing for [Personal Information](#)

Delete a rule

If you have an [author/admin role](#), delete a data access rule that is no longer necessary.

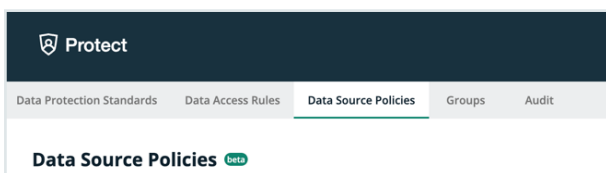
Steps

1. In the rules table, click the  icon in the appropriate row
 - » The **Delete data access rule** dialog box appears.
2. Click the red **Delete** button.



Data source policies

The Data Source Policies page contains an overview of the available policies in your environment.



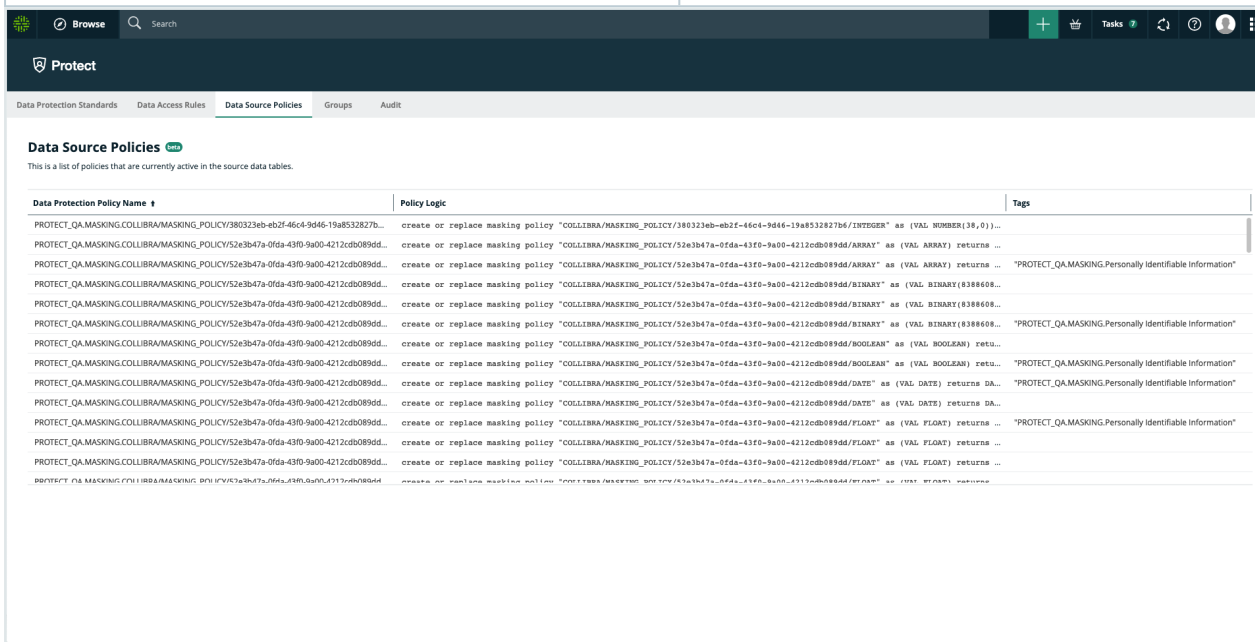
The data protection policy table displays a list of policies that are currently active in the source data tables. This includes policies that were created via Collibra Protect as well as policies that were created in the data source manually.

Note Collibra Protect currently only supports the Snowflake data source.

The table columns include:

| Column name | Description |
|-----------------------------|---|
| Data Protection Policy Name | Policies that originated in Protect have this structure: [DB name].[SCHEMA name].[policy type*].[asset id]. *Policy type can also be masking/row-filtering |
| Policy Logic | This column contains the SQL command that is executed in Snowflake whenever the user tries to access the protected object and will determine how to display the data to the user. |

| Column name | Description |
|-------------|--|
| Tags | For policies that originated in a standard, this column lists the name of the attached tag. The convention is that each tag has the name of the asset that is included in that standard. |



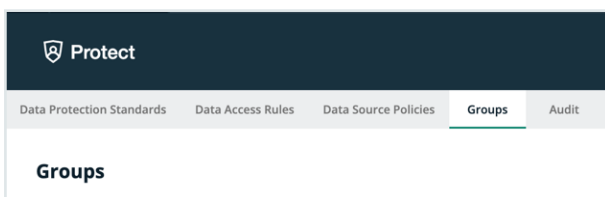
Types of policies on Snowflake

There are three types of policies on Snowflake: Column-based policies, row access policies, and tag-based policies. Each type can be created in Protect or on Snowflake.

For rules, policies are created directly on the column level. Row access policies are created when row filters are specified. For standards, the policy is created, attached to a Snowflake tag, and attached to the tab on any affected column.

Groups

The Groups page contains an overview of the created Collibra Protect groups in your environment.



The groups table displays a list of groups that are currently active in the data source.

The screenshot shows the main content area of the 'Groups' page. It features a table with the following data:

| Group Name | System Reference | Created By | Created date |
|-----------------|--------------------------|----------------|------------------------|
| CID | "Snowflake": "string" | Admin Istrator | Jun 16, 2022, 8:52 AM |
| Human Resources | "Snowflake": "HR" | Admin Istrator | May 11, 2022, 11:39 AM |
| Marketing | "Snowflake": "MARKETING" | Admin Istrator | May 11, 2022, 11:39 AM |

Below the table, there is a section titled 'Adding Groups' with a note: 'To add a group, you have to use the [Collibra Protect Group API](#). Currently, only Snowflake data sources are supported.'

Note Collibra Protect currently only supports the Snowflake data source.

The table columns include:

| Column name | Description |
|------------------|------------------------------------|
| Group Name | Name of the Protect group |
| System Reference | |
| Created By | User who created the Protect group |

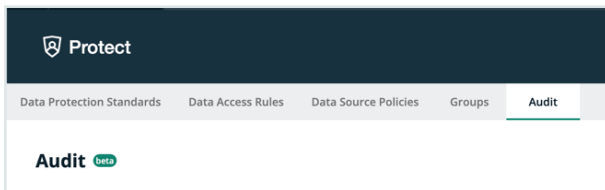
| Column name | Description |
|--------------|----------------------------|
| Created Date | Date the group was created |

Adding groups in Collibra Protect

To add a group, use the [Collibra Protect Group API link](#). This action must be done before any data protection standards or data access rules can be created.

Audit

The Audit page allows you to generate an audit log of access records from the data source .



Generate an audit log

Why would I need to generate an audit log?

Note The actions you take in Collibra Protect only appear in an audit log after one hour.

Steps

1. In Collibra Protect, go to the **Audit** tab.

Click one of the predefined time buttons (Today, Yesterday, A week ago, or 30 days ago) or use the date picker to specify a start date for the audit log.

2. Click the green **Generate Log** button.

» The audit log displays the first 1,000 records after the selected start time.

Note | Important Generating an audit log can take some time. Do not navigate away from the page or your request is canceled. For full details, contact your data source administrator.

Audit 🔍

Today Yesterday A week ago 30 days ago

Start Date: 09/29/2022

Generate Log

For audit log generation, data sources may have latency to summarize access records. Logs generated here for today may not contain information for the most recent access.

| Query ID | Query Start Time | Source User Name | Direct Objects Accessed | Base Objects Accessed |
|--------------------------------------|-----------------------|------------------|---|-----------------------------|
| 01a74800-0501-ec9a-0001-000306f6b19e | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.MAIN_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ec9a-0001-000306f6b1a2 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.MAIN_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ea4f-0001-000306f69dd2 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.DEPENDS_ON_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ec9a-0001-000306f6b1a6 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.NODES_DEPENDS_ON_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74801-0501-ea4f-0001-000306f69dda | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ec9a-0001-000306f6b1b6 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1ba | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ea4f-0001-000306f69de2 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1be | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ec9a-0001-000306f6b1c2 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1c6 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |

1-50 51-100 101-150 ... 701-712

The audit log table columns include:

| Column name | Description |
|------------------------|---|
| Query ID | The ID of the query in the source DB. |
| Query Start Time | Date and time of the query in the source DB. |
| Source User Name | Name of the user in the source DB that conducted the query (accessed the data). |
| Direct Object Accessed | The DB object that was used to access and view the data (a table or a view). |
| Base Object Accessed | The DB object that was accessed and viewed. |

Why rules or standards fail

Certain rules or standards may fail due to logical errors. This section describes some of the common scenarios that cause them to fail.



Different types of masking affecting the same column

Note In this topic, the term *agent* refers to a data category or a data classification.

Masking within a rule

Scenario

A rule that is set for a group masks multiple agents using different types of masking, and the agents share the same column. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group. The rule masks the **Personal Information** data category by hashing and masks the **Personal and family details** data category by showing only the last two digits. Suppose that both these data categories share the same column. Then, the rule will fail because the same column cannot be masked using two different masking types for a given group.

Rule Name*
Masking within a rule

Description

Set rule for

group* Marketing + -

asset* Customer Data + -

and the asset Audit & Internal Controls + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Hashing + - for **Data Category** **Data Classification** Personal Information + -

with ⓘ Show last + - 2 + - for **Data Category** **Data Classification** Personal and family details + -

and Select an action + - rows where Select a data classification + - has Select a code set + - Select a code value + -

Summary
Grant access to Marketing
for Customer Data and Audit & Internal Controls
with Hashing for Personal Information and
with Show last 2 characters for Personal and family details

Masking between rules

This scenario is similar to the [previous scenario](#) except that this scenario considers two rules, instead of one, that are set for the same group. The masking types for the agents in the two rules are different, and both the agents share the same column. Then, a conflict occurs because the same column cannot be masked using two different masking types for a given group.

When two rules conflict with each other, if the synchronization status of only one of them is **Active**, then the other rule fails. If, however, the synchronization status of both the rules is **Active** or **Pending**, then both of them fail.

This scenario is applicable regardless of whether the agents are the same or different, and regardless of whether the rule applies to a single asset or multiple assets.

Rule Name*
Masking between rules - 1

Description

Set rule for

group* Marketing

asset* Customer Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with Hashing for Data Category Data Classification Personal Information

and Select an action rows where Select a data classification has Select a code set Select a code value

Summary
Grant access to Marketing for Customer Data with Hashing for Personal Information

Rule Name*
Masking between rules - 2

Description

Set rule for

group* Marketing

asset* Audit & Internal Controls

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with Show last 2 for Data Category Data Classification Personal and family details

and Select an action rows where Select a data classification has Select a code set Select a code value

Summary
Grant access to Marketing for Audit & Internal Controls with Show last 2 characters for Personal and family details

Masking between standards

Scenario

Two standards mask different agents, and the agents share the same column. This scenario is applicable regardless of whether the groups and the masking types are the same or different.

Example

Consider two standards. The first standard masks the **Personal Information** data category, and the second standard masks the **Name** data classification. Suppose that both the agents share the same column. Then, a conflict occurs because more than one standard cannot be applied to the same column via different agents.

Note This is a limitation on how Collibra Protect implements standards on Snowflake.

When two standards conflict with each other, if the synchronization status of only one of them is **Active**, then the other standard fails. If, however, the synchronization status of both the standards is **Active** or **Pending**, then both of them fail.

Standard Name *

Description

for the group * + -

protect * Data Category Data Classification

with * ⓘ

Summary
 For the Group Marketing
 protect [Personal Information](#)
 with Hashing

Standard Name *
 Masking between standards - 2

Description

for the group * Human Resources

protect * Data Category **Data Classification** Name

with * ⓘ Show last 2

Summary
 For the Group Human Resources
 protect Name
 with Show last 2

Conflicting filters affecting the same column

Filtering within a rule for the same data classification

Scenario

A rule that is set for a group contains conflicting filters for the same data classification. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group and the **Customer Data** asset. The rule contains two filters for the **Country** data classification.

Rule Name *
Filtering within a rule for the same data classification

Description

Set rule for

group * Marketing

asset * Customer Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for Data Category Data Classification Select a data category

and Show rows where Country has Country code BE

and Hide rows where Country has Country code PL

Summary
Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE
and Hide rows where Country has Country code: PL

If any of the tables in the asset contain a column that is classified as **Country**:

- The first filter shows the rows that contain **BE** in that column.
- The second filter hides the rows that contain **PL** in that column.

Then, this rule will fail because two conflicting filters affect the same column.

When applying a filter for a specific data classification, you must select only one type of action. That is, you can choose to either show rows based on one or more values or hide rows based on one or more values. You must not use the show and hide filter actions together for the same data classification.

Filtering within a rule for different data classifications

Scenario

A rule that is set for a group contains conflicting filters for different data classifications that share the same column. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group and the **Customer Data** asset. The rule contains two filters: one for the **Country** data classification, and another for the **State** data classification.

Rule Name *
Filtering within a rule for different data classifications

Description

Set rule for

group * Marketing + -

asset * Customer Data + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option Select a data category

for **Data Category** **Data Classification**

and Show + -
rows where Country Country code BE + -

and Hide + -
rows where State Country code PL + -

Summary

Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE
and Hide rows where State has Country code: PL

If any of the tables in the asset contain columns that are classified as **Country**, the first filter shows only the rows that contain **BE** in those columns.

If any of the tables in the asset contain columns that are classified as **State**, the second filter hides only the rows that contain **PL** in those columns.

Suppose that a column is classified as both **Country** and **State**. That is, data classifications **Country** and **State** share the same column. Then, this rule will fail because two conflicting filters affect the same column.

Filtering between rules for same or different data classifications

This scenario is similar to the [previous scenarios](#) except that this scenario considers two rules, instead of one, that are set for the same group. The filter in one rule is different from the filter in

the other rule, and both the filters affect the same column. Then, a conflict occurs because two conflicting filters affect the same column.

When two rules conflict with each other, if the synchronization status of only one of them is **Active**, then the other rule fails. If, however, the synchronization status of both the rules is **Active** or **Pending**, then both of them fail.

Rule Name*
Filtering between rules for same or different data classifications - 1

Description

Set rule for

group* Marketing + -

asset* Customer Data + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for **Data Category** **Data Classification** Select a data category

and Show rows where Country has Country code BE + -

Summary
Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE

Rule Name*
Filtering between rules for same or different data classifications - 2

Description

Set rule for

group* Marketing + -

asset* Personal Information + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for **Data Category** **Data Classification** Select a data category

and Hide rows where Country has Country code PL + -

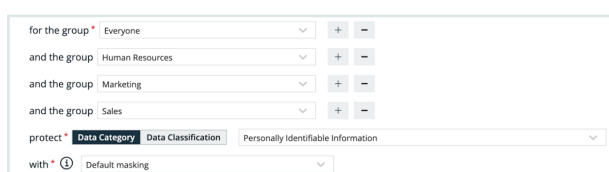
Summary
Grant access to Marketing
for Personal Information
and Hide rows where Country has Country code: PL

Reference documentation

As described in the [DB synchronization](#) section, Collibra Protect periodically does an aggregation of all data protection standards and data access rules available. These standards and rules prepare a representation containing all databases, schemas, tables, and columns involved as well as their protections and accesses. The synchronization process then triggers Edge capabilities, like Collibra Protect for Snowflake, that are responsible for translating the representation to actions toward the data source provider using their technology. This process might involve JDBC and REST calls to perform low-level operations to guarantee that the protections and accesses are applied.

Collibra Protect for Snowflake

Data protection standards rely on [tag-based masking policies](#) available in Snowflake. The name of the data category or data classification specified in the standard becomes a tag, which is applied to all affected columns to enforce data protection. For example, let's say a standard is created on the Personally Identifiable Information data category to restrict access for different groups with the organization.



for the group * Everyone + -
and the group Human Resources + -
and the group Marketing + -
and the group Sales + -
protect * Data Category Data Classification Personally Identifiable Information
with * Default masking

When synchronized and active, the standard resolves to 14 masking policies, which is one for each Snowflake data type. The masking policies are created at the schema level and use the following naming convention: COLLIBRA/MASKING_POLICY/<asset ID>/<snowflake type>.

Chapter 12

| Row | created_on | name ↑ | database_name | schema_name | kind | owner |
|-----|------------------------|--|---------------|-------------|----------------|--------------|
| 1 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/ARRAY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 2 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/BINARY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 3 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/BOOLEAN | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 4 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/DATE | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 5 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/FLOAT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 6 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/GEOGRAPHY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 7 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/NUMBER | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 8 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/OBJECT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 9 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/STRING | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 10 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIME | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 11 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 12 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP_LTZ | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 13 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP_TZ | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 14 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/VARIANT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |

The example below shows a masking policy created for the type STRING.

Note The data the consumers see depends on the masking option selected. Go to the Masking and Data Types page to learn more.

```
1 CASE
2   WHEN CURRENT_ROLE() = 'PUBLIC' THEN '*'
3   WHEN CURRENT_ROLE() = 'HR' THEN '*'
4   WHEN CURRENT_ROLE() = 'MARKETING' THEN '*'
5   WHEN CURRENT_ROLE() = 'SALES' THEN '*'
6   ELSE val
7 END
```

Done

All masking policies are then associated with the Personally Identifiable Information tab, which is created at the schema level and assigned to all columns where the protection needs to be applied. At runtime, Snowflake fetches the right masking policy based on column data type.

| Row | created_on | name | database_name | schema_name | owner | comment |
|-----|------------------------|-------------------------------------|---------------|-------------|--------------|---|
| 1 | 2022-09-06 03:41:13... | Personally Identifiable Information | PROTECT_QA | DEMO | ACCOUNTADMIN | Generated by Collibra: 28d226cc-0ab0-4d23-b912-985312fb36b1 |

Data Access Rules are translated as a combination of [grant instructions](#), [dynamic masking](#), and [row access policies](#) when specified in the rule. For example, a data set named Employee Data has sensitive columns categorized as Personally Identifiable Information.

Chapter 12

| # | Name | is part of |
|----|---------------|------------|
| 1 | EMPLOYEE_NAME | EMPLOYEES |
| 2 | EMP_ID | EMPLOYEES |
| 7 | DEPT_ID | EMPLOYEES |
| 10 | SALARY | EMPLOYEES |

In Collibra Protect, a rule is created to grant access of that data set to Human Resources. Since the Grant Access checkbox is enabled, each database, schema, and table in that data set received a grant for the Snowflake role specified and each column that has protection received a column masking policy.

Set rule for

group: Human Resources

asset: Employee Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to the tables in the database with columns linked to the selected assets. If this box is unchecked, no access will be given to these columns.

with: No masking for Data Category: Data Classification, Personally identifiable information

Let's look closer at one of the columns, such as EMPLOYEE_NAME. It belongs to the EMPLOYEES table within the DEMO schema within the PROTECT_QA database.

EMPLOYEE_NAME

Column Candidate

Summary: Description from source system: Empty value

Data Profiling: Technical Data Type: VARCHAR

In Snowflake, each column has a masking policy assigned to it. The masking policies created at the schema level follow the naming convention: COLLIBRA/MASKING_POLICY/<asset ID>.

| Row | created_on | name | database_name | schema_name | kind | owner |
|-----|-----------------------|--|---------------|-------------|----------------|--------------|
| 18 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY78620268-9754-47af-8044-c684-c9487961 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 17 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY74647679-2307-468f-826f-c6c0793237 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 16 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY93869506-6987-425e-8f52-96469099984a | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 15 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY18327836-0651-4884-0344-23888609104a | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 14 | 2022-09-08 03:48:09.8 | COLLIBRAMASKING_POLICY12822280-0467-4423-9912-96531291981581688187 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |

The content of the masking policy created for the column EMPLOYEE_NAME is shown below.

```
Details
1 CASE
2     WHEN CURRENT_ROLE() = 'HR' THEN va1
3     WHEN CURRENT_ROLE() = 'PUBLIC' THEN '*'
4     WHEN CURRENT_ROLE() = 'MARKETING' THEN '*'
5     WHEN CURRENT_ROLE() = 'SALES' THEN '*'
6     ELSE va1
7 END
```

Done

The Human Resources group has access to the plain data without any masking while the other groups have masked access as created in the data protection standard.

Important In this example, the column EMPLOYEE_NAME has the policy tag and a column masking policy assigned to it. In Snowflake, when both are present, the column masking policy takes precedence and the policy tag is not executed. To mitigate this behavior and ensure that the protection defined in the standard is applied, we prepare the column masking policy with the conditions defined in the policy tag.

Masking and data types

Snowflake provides several functions to transform the data. In Collibra Protect, we support four masking options.

1. Default Masking is not supported by Snowflake. This implementation was added in our Protect capability, so protection can be applied to a wide range of data types. Each column received a default value according to the column data type. Below is a list of Snowflake data types and their default values.

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|-----------------------|
| 1 | NUMBER | NUMBER | 0 |
| 2 | DECIMAL | NUMBER | 0 |
| 3 | NUMERIC | NUMBER | 0 |
| 4 | INT | NUMBER | 0 |
| 5 | INTEGER | NUMBER | 0 |
| 6 | BIGINT | NUMBER | 0 |
| 7 | SMALLINT | NUMBER | 0 |
| 8 | TINYINT | NUMBER | 0 |
| 9 | BYTEINT | FLOAT | 0 |
| 10 | FLOAT | FLOAT | 0 |
| 11 | FLOAT4 | FLOAT | 0 |
| 12 | FLOAT8 | FLOAT | 0 |

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|----------------------------|
| 13 | DOUBLE | FLOAT | 0 |
| 14 | DOUBLE PRECISION | FLOAT | 0 |
| 15 | REAL | FLOAT | 0 |
| 16 | VARCHAR | VARCHAR | * |
| 17 | CHAR | VARCHAR | * |
| 18 | CHARACTER | VARCHAR | * |
| 19 | STRING | VARCHAR | * |
| 20 | TEXT | VARCHAR | * |
| 21 | BINARY | BINARY | 00 |
| 22 | VARBINARY | BINARY | 00 |
| 23 | BOOLEAN | BOOLEAN | false |
| 24 | DATE | DATE | 1970-01-01 |
| 25 | DATETIME | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |
| 26 | TIME | TIME | 00:00:00 |
| 27 | TIMESTAMP | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|--|
| 28 | TIMESTAMP_LTZ | TIMESTAMP_LTZ | 1969-12-31 16:00:00.000-0800 <i>Might change based on user TZ</i> |
| 29 | TIMESTAMP_NTZ | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |
| 30 | TIMESTAMP_TZ | TIMESTAMP_TZ | 1969-12-31 16:00:00.000-0800 <i>Might change based on user TZ</i> |
| 31 | VARIANT | VARIANT | 0 |
| 32 | OBJECT | OBJECT | {} |
| 33 | ARRAY | ARRAY | [] |
| 34 | GEOGRAPHY | GEOGRAPHY | {"coordinates": [0,0], "type": "Point"} (aka point(0, 0) and visualization can change based on user preferences) |

In Collibra Protect, we also support the hashing and show last masking options. These can only be applied to Snowflake data types STRING, NUMBER, and FLOAT.

2. Hashing allows us to use Snowflake's SHA2 value function for strings, and the HASH value for numbers
3. Show Last allows us to use the substr(to_varchar(value), length(value) - n, n) expression for strings, and mod(value, power(10,n)) for numbers. Value is the content and n is the number of characters to be shown.
4. No Masking is when the raw content is returned.

Note Whenever a masking option cannot be applied, like hashing on the DATE type, default masking is applied, so protection is guaranteed.

Collibra Protect

About Collibra Protect

Collibra Protect is a capability of the Data Intelligence Cloud created to protect sensitive data and make it available, or partially available, to specified groups of users.

Collibra Protect solves the problem of protecting sensitive data in an organization. Different groups of people may need varying access levels to the same data set. With Collibra Protect, access rules and data protection standard capabilities allow you to grant access to individuals and protect sensitive information. These rules and standards with different data access levels are managed through the Collibra platform and pushed to the data source. Our aim is to promote a safe data-open culture in organizations.

The goal of Collibra Protect is to centralize and simplify access governance and remove the need of repetitive action and approval. Data access and privacy management promotes an ethical company standard giving permission to view information only to those that need it. Collibra Protect allows you to perform these actions accordingly.

An example use case of Collibra Protect is a data steward giving everyone access to a data set, but only allowing certain access to groups of people based on data categories. This is known as differential access. It is suggested that rules/standards are grouped together, for example by business processes, so you do not have to make a rule or standard for every data set.



Install Collibra Protect

This procedure guides you through a first time installation of Collibra Protect.

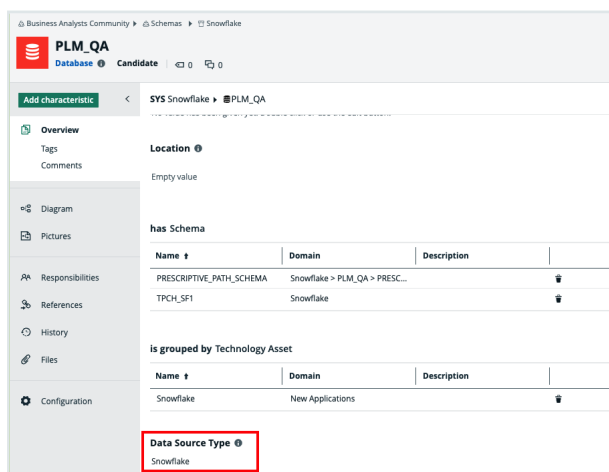
Prerequisites

You must add the [Snowflake capability on Edge](#) as well as perform a catalog ingestion.

Configure the Collibra Protect for Snowflake capability on Edge. Settings → (Edge) Sites → Your site → Capability → Add capability → fill in the needed parameters:

- For "Capability template" choose "Collibra Protect for Snowflake".
- The "Snowflake Connection" can be the same connection used for doing catalog ingestion. Make sure that the Snowflake user/role has enough permissions to create/alter/drop grants, tags, etc.

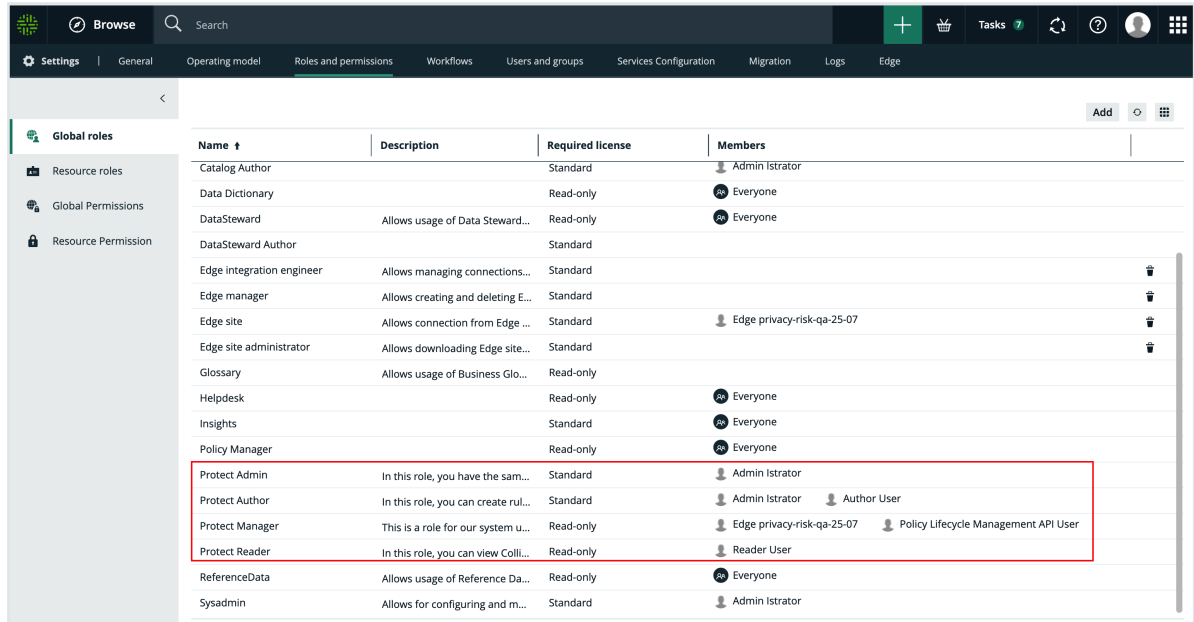
An ingested Snowflake database should look like the example below.




Note The Data Source Type attribute on the database asset should be present. This attribute is automatically added in database assets, after the catalog ingestion process.

Steps

1. Contact a Collibra support or your representative to enable Collibra Protect on your Collibra environment.
2. Ensure [global roles and permissions](#) for Collibra Protect are set correctly.



3. Collibra Protect is installed.
 - » You can now access and start using Collibra Protect via the  menu.

Configure Collibra Protect

Configuring within Collibra Protect is an important part of understanding and using Collibra Protect to its highest ability.

Prerequisites

- You need to have Data Catalog permissions. If not, you cannot see any classification in either standards or rules.
- You need to have a Data Steward role within Collibra. If not, you cannot see the classification page when selecting a classification in Collibra Protect.



Roles in Collibra Protect

It is possible to assign different roles to Collibra users that use Collibra Protect. The roles are provided and have pre-defined permissions that restrict the usage of the application.

| Roles | Description |
|----------------|---|
| Protect Reader | Users in this role can view Collibra Protect with read-only access to the content. This role is assigned to 'Everyone' and grants the users the 'protect' permission. Without this permission, users cannot see 'Protect' as an application in the ☰ menu. They also cannot navigate to protect related URLs or access protect endpoints. |
| Protect Author | Users in this role can create rules and standards , view imported policies and groups , and generate audits as an individual contributor. This role grants the product right permission 'protect' and the 'protect_edit' permission. Authors can only modify rules and standards they own. This role is not assigned to anyone automatically. |
| Protect Admin | Users in this role have the same permissions as the Protect Author role as well as the ability to edit other user's rules and standards. This role grants the product right permission 'protect', 'protect_edit', and an extra 'protect_administration' permission. This role is not assigned to anyone automatically. |

| Roles | Description |
|-----------------|--|
| Protect Manager | This role is restricted to our system user to manage background processes and setup configurations for Collibra Protect and it should not be assigned to other Collibra users. |

Configure groups

Before you start working in Collibra Protect, you need to configure your groups. Protect groups are the basis of all the actions performed in Collibra Protect.

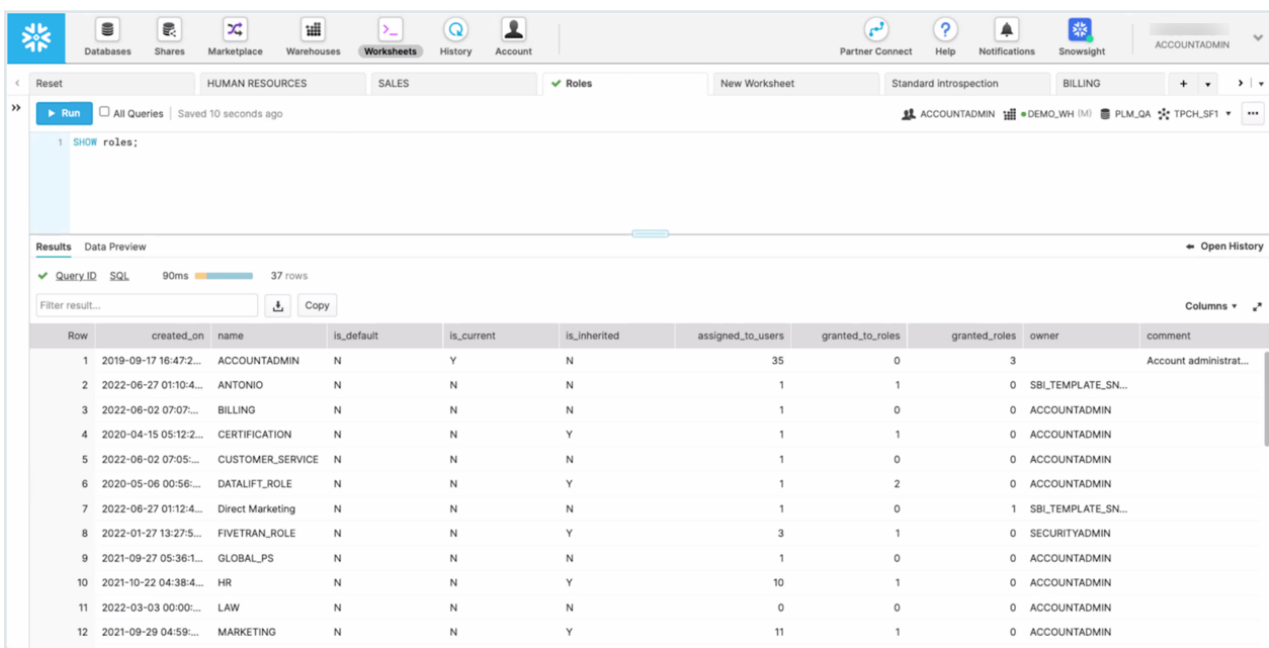
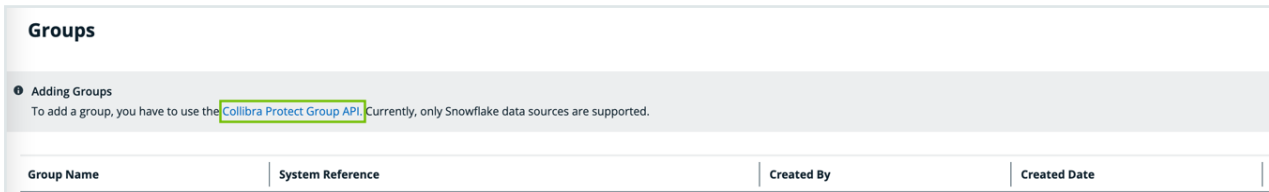
Associate a Protect group with Snowflake

Each Snowflake user is assigned to one or more Snowflake roles. Permissions are based on these roles. View the example below of the roles page in Snowflake. Any/all roles can be correlated to a Collibra Protect group.

| Role | Creation Time | Owner | Comment |
|------------------|------------------------|--------------------|--|
| ACCOUNTADMIN | 9/18/2019, 1:47:25 ... | | Account administrator can manage all aspects of the account. |
| ANTONIO | 6/27/2022, 10:10:4... | SBL_TEMPLATE_SN... | |
| BILLING | 6/2/2022, 4:07:43 ... | ACCOUNTADMIN | |
| CERTIFICATION | 4/15/2020, 2:12:24 ... | ACCOUNTADMIN | |
| CUSTOMER_SERVICE | 6/2/2022, 4:05:29 ... | ACCOUNTADMIN | |
| DATALIFT_ROLE | 5/6/2020, 9:56:54 ... | ACCOUNTADMIN | |
| Direct Marketing | 6/27/2022, 10:12:4... | SBL_TEMPLATE_SN... | |
| FIVETRAN_ROLE | 1/27/2022, 10:27:58... | SECURITYADMIN | |
| GLOBAL_PS | 9/27/2021, 2:36:19 ... | ACCOUNTADMIN | |
| HR | 10/22/2021, 1:38:44... | ACCOUNTADMIN | |
| LAW | 3/3/2022, 9:00:27 ... | ACCOUNTADMIN | |
| MARKETING | 9/29/2021, 1:59:26 ... | ACCOUNTADMIN | |
| MARKETING2 | 9/29/2021, 2:36:17 ... | ACCOUNTADMIN | |
| MARKETING3 | 9/30/2021, 3:56:47 ... | ACCOUNTADMIN | |
| PC_DBT_ROLE | 5/6/2022, 9:08:33 ... | ACCOUNTADMIN | System created role for partner elt integration. |
| PLM | 10/22/2021, 1:30:58... | ACCOUNTADMIN | |
| PLM_QA_HR | 2/24/2022, 3:38:20... | ACCOUNTADMIN | PLM QA HR Read Only Role |

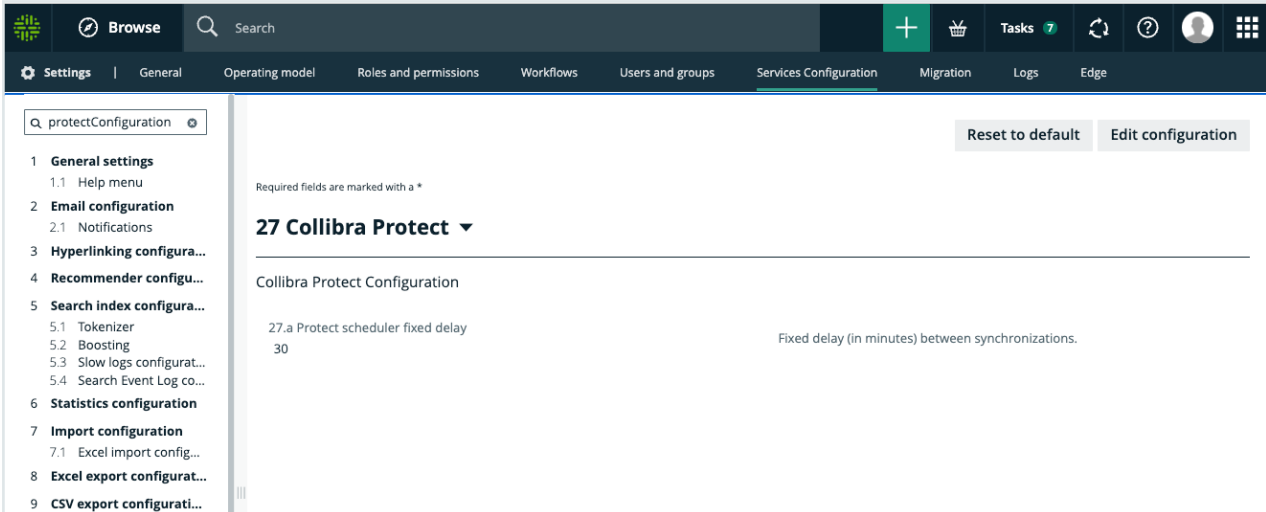
How to create Collibra Protect groups?

When you initially go to the **Groups** tab in Collibra Protect, there are no groups created. There is a link at the top of the page to the Groups API that creates new groups in Collibra Protect. Use this API link to create new groups and associate it with a specific role in Snowflake.



General configuration

Collibra Protect synchronizes standards and rules with the source database(s) at regular intervals. This synchronization runs in the background on a configured frequency. By default, the frequency is every 60 minutes, but this is configurable through Settings → Services Configuration → 27 Collibra Protect.



Important If you do not have access to the **Service Configuration** tab, create a support ticket requesting the JVM Parameter be added to your Collibra Infrastructure Configuration: `-DPROTECT_SYNC_SCHEDULER_DELAY=PT60M`. After the parameter is added, restart Collibra so these changes take effect and the policies are now synchronized with the cloud provider.

Synchronization includes:

1. Aggregate all standards and rules computing:
 - which columns need to be masked for which groups.
 - which tables need to have a row filter.
 - which tables and columns need to be granted access.
2. On the source database(s) such as Snowflake:
 - create and apply maskings.
 - create and apply row filters.
 - grant access to groups on tables and/or columns (depending on the underlying database).

Essentials for Collibra Protect

To use Collibra Protect to the best of its ability, you need to know the following things:

- [How to protect your data](#)
- [Technical background](#)
- [Data protection standards vs. data access rules](#)
- [Prescriptive paths](#)



How to protect your data

1. Access management

The most basic line of protection is to make sure only the right people/groups have access to the data. Data here is referring to the tables and columns in your database. In Collibra Protect, you can grant specific groups access to parts of your data based on Collibra assets.

For example, it is easy to grant the HR team access to the US customers' data set. But, what if some parts of the US customers' data set need to be hidden from the HR team, because it contains restricted information, such as personally identifiable information (PII)? In that case, you can further protect your data by applying column-based protection or row-based protection.

Note Collibra Protect only grants access. It cannot revoke access from people/groups.

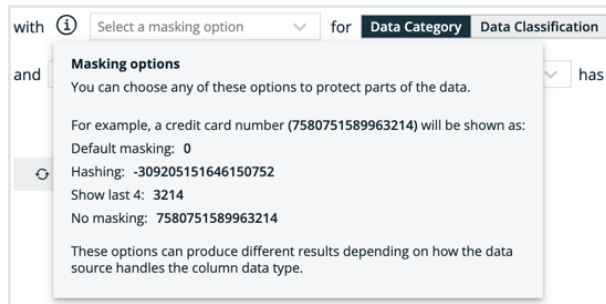
2. Column-based protection

Column based protection allows you to target specific columns and mask their content. By masking the column's data, the group cannot see the content as it is. They will see a masked version of it instead.

For example, you can mask a column of credit card numbers, so the individual group cannot see the full credit card numbers.

We currently support four masking options. They include:

- **Default masking:** Shows the value as 0.
- **Hashing:** Converts the value into a variety of different letters, numbers, and symbols.
- **Show last:** Displays the last letters, numbers, and symbols in the value. You can choose to show the last 1 through 20 of the value. The most common choice is Show last 4.
- **No masking:** Displays the data value as it is originally written.



Collibra Protect allows you to choose to mask columns that are part of a **data category** or a **data classification**. While granting access to a certain asset, you can choose to apply this masking on only a subset of that asset if it is also part of a data category or data classification.

3. Row-based protection

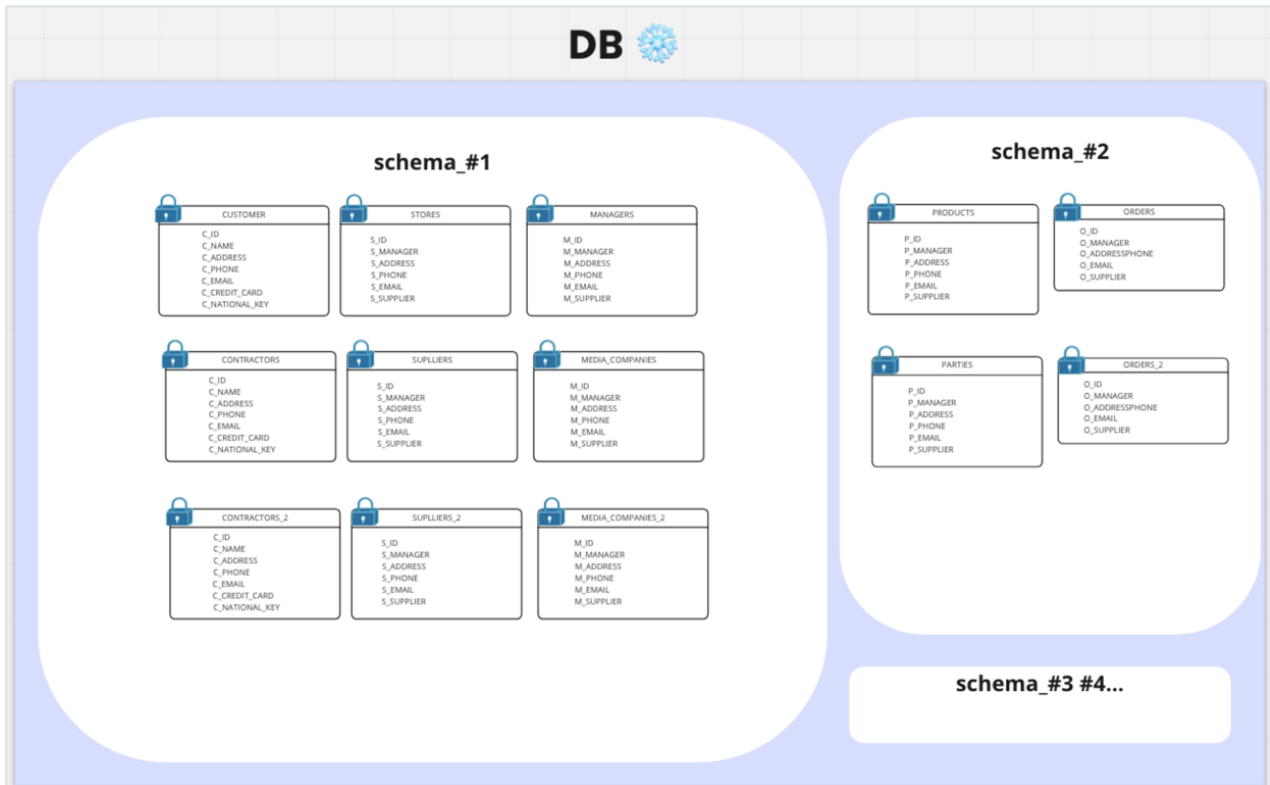
Another way to protect your data is to filter rows of a specific table. If you do not want to expose all of the existing items in a table because one of the columns is part of a certain data classification, you can easily leverage the Collibra operating model to do so.

When creating a rule that impacts certain tables in the source database, filter rows on tables by using the row filtering option for tables where one of their columns is part of a data classification. The filtering is based on what value is stored in the cell of that particular column. For instance, in a table that has a column that is classified as **country-code**, you can hide or show all items that have the value of **US**.

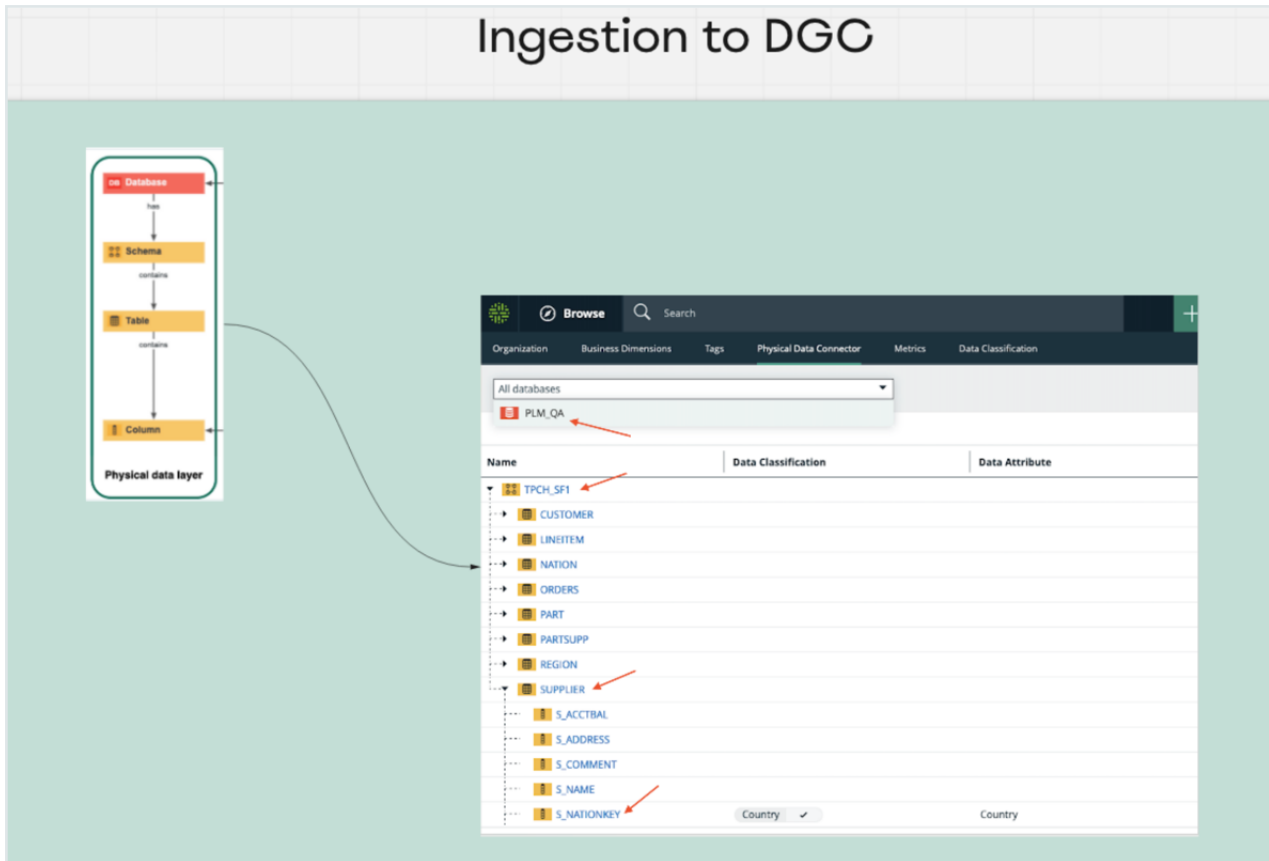
Technical background

The technical background of Collibra Protect explains the connection of the data as it is in the database (DB) with the physical layer (equivalent assets in Collibra Data Intelligence Cloud) and the logical layer (the out of the box model).

Imagine you have this database:



When ingesting this DB to Collibra Data Intelligence Cloud, the physical layer is created as well as an asset for each of the schemas, tables, and columns.



Once there is a physical layer established in our Collibra environment, start creating the logical layer on top of it.

- In this phase, take any column and classify it as any data classification available, or let the platform classify it for you.
- Also, assign a column to a data attribute.

From here, create additional assets or use existing assets of different types (data set, data category, or business process) to establish a relation to these columns.

Data protection standards vs. data access rules

Collibra Protect has both standards and rules to govern your data with ease and clarity.

| | |
|-------------------------|---|
| <p>Standards</p> | <p>Data protection standards create a layer of protection for similar types of data by masking them wherever they are.</p> <p>For example, if columns with first and last names are a part of the PII data category, regardless what tables, schemas, and databases they are part of, create a standard that targets all of these columns by choosing the PII data category and masking it.</p> |
| <p>Rules</p> | <p>After establishing this primary layer (blanket) of protection to your most sensitive data, use data access rules to manage access and enhance protection for specific usages.</p> <p>For example, create a rule that grants access to a specific group, for a specific data set, while knowing that all PII within this data set will be masked by the standard we created before.</p> |

FAQs

1. What if I want to grant access to a group without having the PII masked?
 - » When creating a rule for an asset that contains data masked by a standard, choose to override it by unmasking it or changing its masking type.
2. What If I want to grant access to a group, but the protection from the standard is not enough because there might also be other sensitive data within this supported asset?
 - » When creating a rule, add additional layers of protection over the ones that were set by any existing standard. Further protect the data by applying additional masking on or by filtering the data.

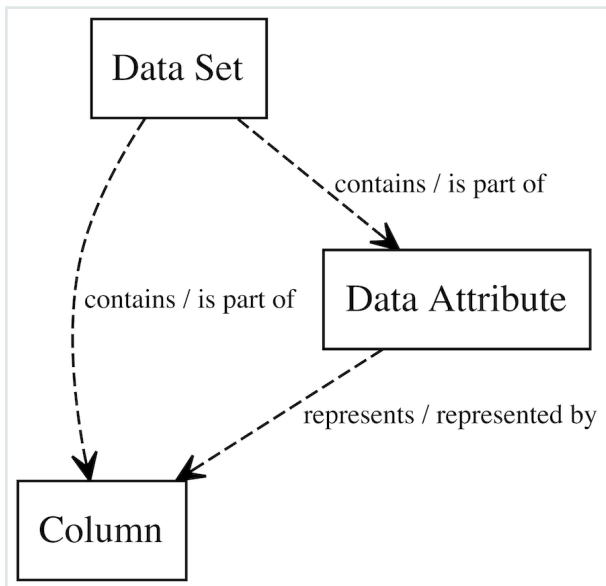
Prescriptive paths

When creating a standard or rule, you select which asset(s) you want to protect and/or grant access to. By default, you can grant access to a data set, a data category, and a business process. Colibra Protect searches the knowledge graph, through relationships and/or

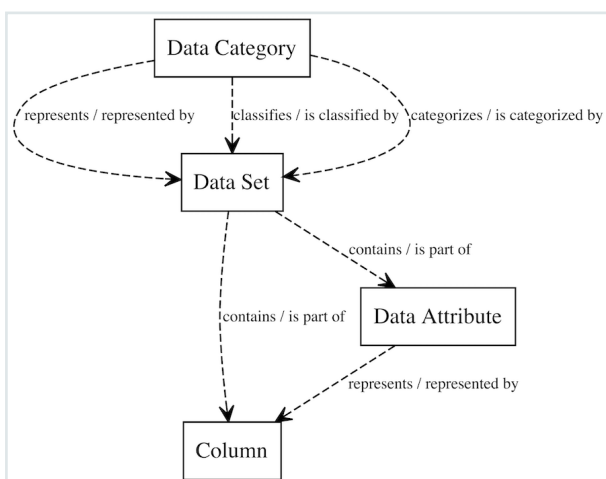
intermediate assets, to find which set of physical data layer assets, such as columns and tables, this resolves.

The traversal of the knowledge graph is done through a set of prescriptive paths. For each type of asset, there is a set of prescriptive paths to traverse to the column assets. See the images below for more details.

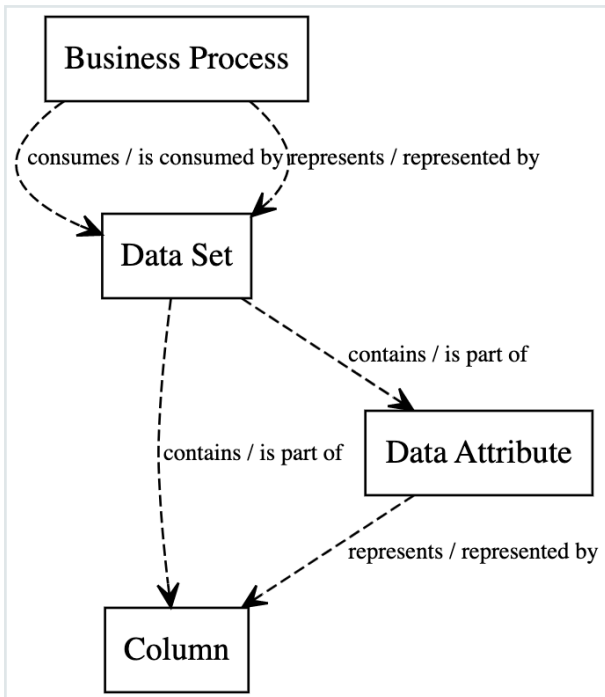
Prescriptive path for data set



Prescriptive path for data category



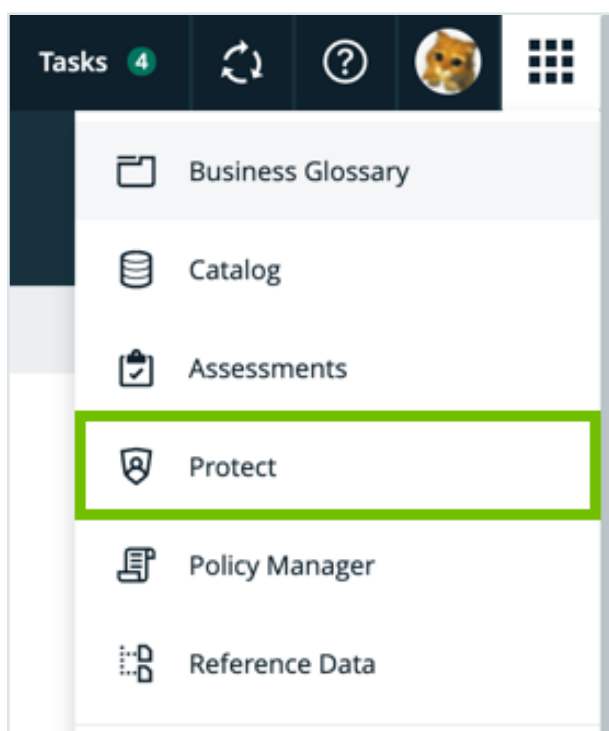
Prescriptive path for business process



Overview of Collibra Protect

To work with Collibra Protect, ensure that you have a global role that has the Protect global permission and that it is [enabled](#) in your environment.

You will find, Collibra Protect, in the main menu . Click **Protect**.



If Collibra Protect is not shown on the menu, the feature is not enabled.

The landing page displays five tabs at the top of the page: **Data Protection Standards**, **Data Access Rules**, **Data Source Policies**, **Groups**, and **Audit**.



Data Protection Standards

Data Access Rules

Data Source Policies

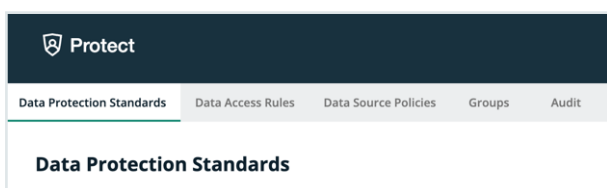
Groups

Audit

| Tab | Description |
|---------------------------|---|
| Data Protection Standards | <p>Define default data source access to data types based on data categories, data attributes, or classes/classifications through data protection standards</p> <p>Note Data access rules for particular groups can override created standards.</p> |
| Data Access Rules | Use data access rules to grant groups different access to the same data in data sets, in business processes, or identified by data categories. |
| Data Source Policies | View a list of policies that are currently active in the source data tables. You can also import policies from your source database using the Colibra Protect Data Source Policies API. |
| Groups | <p>Add groups through custom code via the Data Access API link and view existing current data access groups.</p> <p>Note You must add at least one group before you can create a standard or a rule.</p> |
| Audit | Generate an audit log for a preview of the last hour of ingested data from the data source. |

Data protection standards

The Data Protection Standards page contains an overview of the available standards in your environment.




| Page Section | Description |
|-----------------------------|--|
| Standards summary | Under the heading, there is a summary about data protection standards. Click the Create a Data Protection Standard button to create a standard and get started in Collibra Protect. |
| Recently Modified Standards | This section shows the five most recently modified standards. |
| Standards table | This table displays a detailed view of the created data protection standards. |

In the **Synchronization status** column of the standards table, there are five status options that can appear. To view the status of the standard in the data source, go to the source database.

| Synchronization Status | Description |
|------------------------|---|
| Active | This standard is currently active in Collibra Protect and in the data source. |



| Synchronization Status | Description |
|------------------------|---|
| Pending | This standard has been created or edited, and is pending synchronization. |
| Failed | The synchronization of this standard has failed. Click the  icon next to the failed status to view additional information about the error. |
| Delete Pending | This standard will be deleted from the data source in the next synchronization. |
| Not Deleted | The deletion of this standard has failed. |

Note Collibra Protect periodically synchronizes with the data source and statuses will be updated along with the synchronization. To learn more, go to the [general configuration](#) page.

Create a data protection standard

Data protection standards create a layer of protection by masking data wherever they appear. Create a data protection standard to get started using Collibra Protect.

Create a Data Protection Standard
✕

Data protection standards apply default data source access to types of data based on data categories or data classifications. Data Access Rules for particular groups will override these defaults.

Standard Name*

Description

for the group* + -

protect* Data Category Data Classification

with* ⓘ


Summary
 For the Group Human Resources
 protect Personal Information
 with Hashing

Cancel
Save Standard

Steps

1. In Collibra Protect, go to the **Data Protection Standards** tab.
2. Click the green **Create a Data Protection Standard** button.
 - » The **Create Data Protection Standard** dialog box appears.
3. Enter the required information. It is important to note that when selecting assets, user permissions are defined in Collibra. If an asset is not visible for you, it will not appear as an option in the drop down menus.

| Field | Description |
|---------------|-------------------------------------|
| Standard name | Name of the standard being created. |

| Field | Description |
|-------------------------------------|--|
| Description (optional) | Description of the standard. |
| Group | Group(s) for which the standard is created. |
| Data Category / Data Classification | A data category or data classification to apply the protection on. |
| Masking | Masking option for the standard. <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;"> <p>Note Click  to learn more about the masking options for standards.</p> </div> |

Note Click the plus sign to add more to each field where applicable. For example, after selecting a group, click + to add another group into the standard, and click – to delete a selected group. When entering the required information, you can view the selections you made in the **Summary** section.

4. Click the green **Save Standard** button.
 - » The saved data protection standard appears in the standards table.

Modify a data protection standard


You can edit or delete a data protection standard after it has been created.

Edit a standard

Editing a data protection standard might be necessary in certain situations. For example, change the masking method from default masking to hashing.

Important You will only be able to edit standard assets if you have view asset permissions. If one of the assets in the standard is unauthorized, you will not be able to edit the standard until the view access permission is granted.

Steps

1. In the standards table, click the standard name, and then click the **Edit** button or click  in the appropriate row
 - » The **Edit a Data Protection Standard** dialog box appears.
2. Edit the [required information](#).
3. Click the green **Save Standard** button.
 - » The updated data protection standard appears in the standards table.

Edit a Data Protection Standard
✕

Data protection standards apply default data source access to types of data based on data categories or data classifications. Data Access Rules for particular groups will override these defaults.


Standard Name *

Description

for the group *

and the group

protect * Data Category Data Classification

with * 


Summary

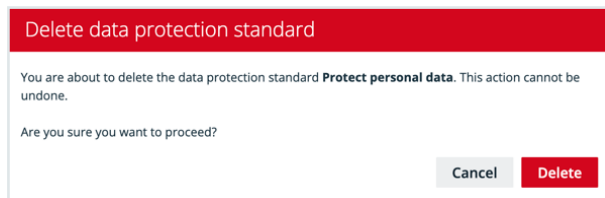
For the Group Human Resources and Marketing
protect [GDPR data related to criminal convictions and offences](#)
with Default masking

Delete a standard

If you have an [author/admin role](#), delete a data protection standard that is no longer necessary.

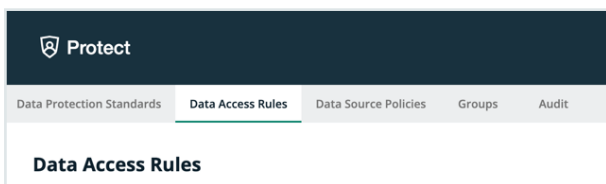
Steps

1. In the standards table, click the  icon in the appropriate row
 - » The **Delete data protection standard** dialog box appears.
2. Click the red **Delete** button.



Data access rules


The Data Access Rules page contains an overview of the available rules in your environment.



| Page Section | Description |
|-------------------------|---|
| Rules summary | Under the heading, there is a summary about data access rules. Click the Create a Data Access Rule button to create a standard . |
| Recently Modified Rules | This section shows the five most recently modified rules. |
| Rules table | This table displays a detailed view of the created data access rules. |

In the **Synchronization status** column, there are five status options that can appear. To view the status of the rule in the data source, go to the source database.

| Synchronization Status | Description |
|------------------------|---|
| Active | This rule is currently active in Collibra Protect and in the data source. |
| Pending | This rule has been created or edited, and is pending synchronization. |

| Synchronization Status | Description |
|------------------------|--|
| Failed | The synchronization of this rule has failed. Click the  icon next to the failed status to view additional information about the error. |
| Delete Pending | This rule will be deleted from the data source in the next synchronization. |
| Not Deleted | The deletion of this rule has failed. |

Note Collibra Protect periodically synchronizes with the data source and statuses will be updated along with the synchronization. To learn more, go to the [general configuration page](#).

Create a data access rule

After establishing a primary layer (blanket) of protection to your most sensitive data using standards, create data access rules to manage access to the data sources and enhance protection for specific usages.

Create a Data Access Rule
✕

Use data access rules to grant groups different access to the same data in data sets, business processes, or identified by data categories. You can mask or hide columns by their data category and you can also conditionally filter rows based on code set values.

Rule Name *
Marketing GI Rule

Description
Set rule for the marketing group for the geographic information asset
Apply default masking for genetic data

Set rule for

group * Marketing + -

asset * Geographic Information + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Default masking + -

for Data Category Data Classification Genetic data + -

and Select an action + -

rows where Unauthorized + -

has Select a code set + -

Select a code value + -

Summary
Grant access to Marketing
for [Geographic Information](#)
with Default masking for [Genetic data](#)

Generate Preview

Cancel
Save Rule

Steps

1. In Collibra Protect, go to the **Data Access Rules** tab.
2. Click the green **Create a Data Access Rule** button.
 - » The **Create a Data Access Rule** dialog box appears.
3. Enter the required information. It is important to note that when selecting assets, user permissions are defined in Collibra. If an asset is not visible for you, it will not appear as an option in the drop down menus.

| Field | Description |
|------------------------|--|
| Rule name | Name of the rule being created |
| Description (optional) | Description of the rule. |
| Group | Group for which the rule is being created. |
| Asset Name | Data asset that the rule is protecting. Protect enables you to protect the following asset types: Business process, data set, and data category. Learn more in technical background and prescriptive paths . |

| Field | Description |
|---|--|
| <p>Masking (optional)</p> <ul style="list-style-type: none">◦ Data Category / Data Classification | <p>Masking option for the rule. Click the i to learn more about masking options.</p> <ul style="list-style-type: none">◦ Select a data category or a data classification to apply masking to. |

| Field | Description |
|---|--|
| Action (optional) <ul style="list-style-type: none"> ◦ Data Classification ◦ Code Set ◦ Code Value | Filter the data by selecting hide or show. <ul style="list-style-type: none"> ◦ Select data classification that is either hidden or shown ◦ Code set to set up row filtering in the tables. A code set must be selected to filter by a code value. ◦ Code value of the code set selected. |

Important The grant access checkbox is selected by default. By leaving this checkbox selected, you are granting access to the tables in the database with columns linked to the selected assets to the selected group(s). If you do not want to grant this kind of access to these groups, clear the grant access checkbox.

Note Click the plus sign to add more to each field where applicable. For example, after selecting a group, click + to add another group into the standard, and click – to delete a selected group. When entering the required information, you can view the selections you made in the **Summary** section.

- Click **Generate Preview** to see a preview of the new rule.

Summary
Grant access to Marketing
for [Geographic Information](#)
with Default masking for [Genetic data](#)

Geographic Information ▾

| Column ↑ | Access | Masking Agent | Masking | Code Value |
|---|--------|---------------|---------|------------|
| C_ADDRESS_sdfxgxcfhjvhjvbkjbjhgxdfzs... | Masked | Genetic data | 0 | |
| C_NAME | Masked | Genetic data | 0 | |
| DS_TBL0001_COL0001 | Masked | Genetic data | 0 | |

Tip Use the preview to verify the data access rule is set up correctly. The preview only shows the first 1,000 affected columns. The drop-down below the **Generate Preview** button is used to switch between the different selected assets in the rule. Each asset has its own preview table.

- Click the green **Save Rule** button.
 - » The saved data access rule appears in the rules table.

Modify a data access rule


You can edit or delete a data access rule after it has been created.

Edit a rule

Editing a data access rule might be necessary in certain situations. For example, change the code set value from BE to US.

Important You will only be able to edit rule assets if you have view asset permissions. If one of the assets in the rule is unauthorized, you will not be able to edit the rule until the view access permission is granted.

Steps

1. In the rules table, click the rule name, and then click the **Edit** button or click  in the appropriate row
 - » The **Edit a Data Access Rule** dialog box appears.
2. Edit the [required information](#).
3. Click the green **Save Rule** button.
 - » The updated data access rule appears in the rules table

Edit a Data Access Rule
✕

Use data access rules to grant groups different access to the same data in data sets, business processes, or identified by data categories. You can mask or hide columns by their data category and you can also conditionally filter rows based on code set values.

Rule Name*

Description

Set rule for

group* + -

asset* + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ for **Data Category** + -


and rows where has

Summary
 Grant access to Marketing
 for [Customer Data](#)
 with Hashing for [Personal Information](#)

Delete a rule

If you have an [author/admin role](#), delete a data access rule that is no longer necessary.

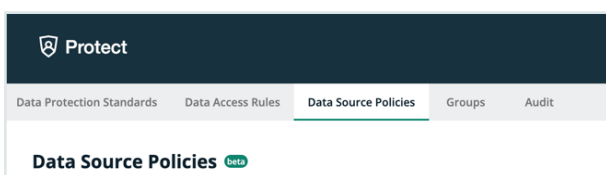
Steps

1. In the rules table, click the  icon in the appropriate row
 - » The **Delete data access rule** dialog box appears.
2. Click the red **Delete** button.



Data source policies

The Data Source Policies page contains an overview of the available policies in your environment.



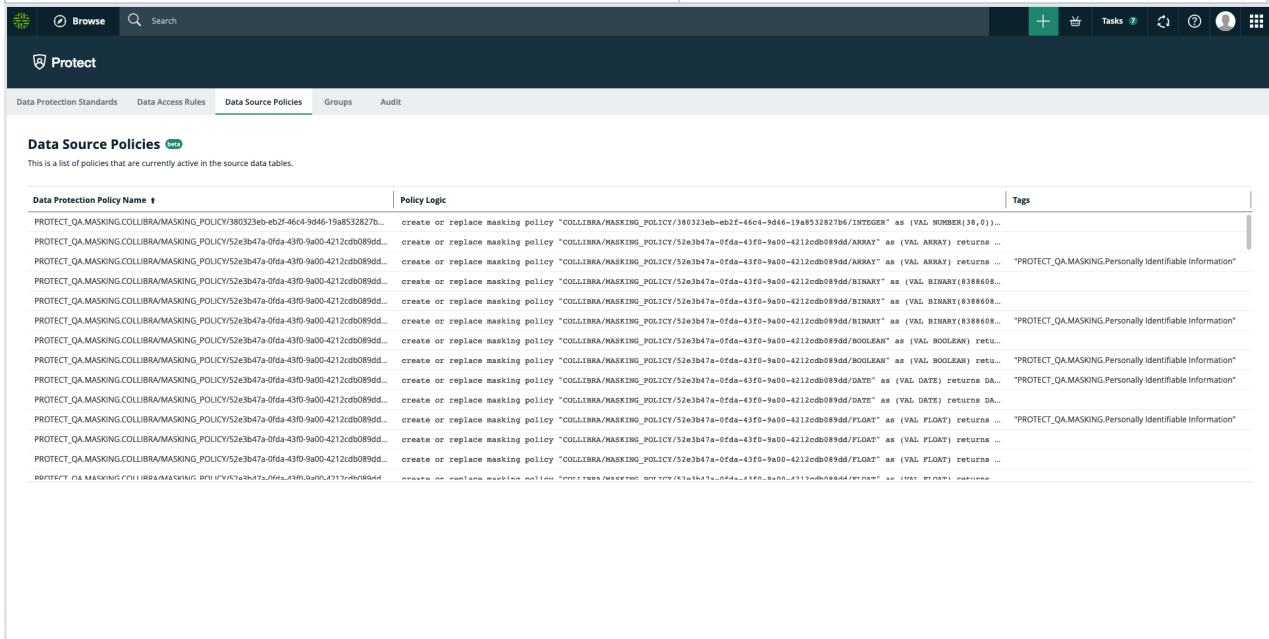
The data protection policy table displays a list of policies that are currently active in the source data tables. This includes policies that were created via Collibra Protect as well as policies that were created in the data source manually.

Note Collibra Protect currently only supports the Snowflake data source.

The table columns include:

| Column name | Description |
|-----------------------------|---|
| Data Protection Policy Name | Policies that originated in Protect have this structure: [DB name].[SCHEMA name].[policy type*].[asset id]. *Policy type can also be masking/row-filtering |
| Policy Logic | This column contains the SQL command that is executed in Snowflake whenever the user tries to access the protected object and will determine how to display the data to the user. |

| Column name | Description |
|-------------|--|
| Tags | For policies that originated in a standard, this column lists the name of the attached tag. The convention is that each tag has the name of the asset that is included in that standard. |



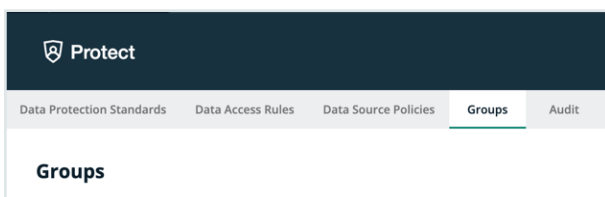
Types of policies on Snowflake

There are three types of policies on Snowflake: Column-based policies, row access policies, and tag-based policies. Each type can be created in Protect or on Snowflake.

For rules, policies are created directly on the column level. Row access policies are created when row filters are specified. For standards, the policy is created, attached to a Snowflake tag, and attached to the tab on any affected column.

Groups

The Groups page contains an overview of the created Collibra Protect groups in your environment.



The groups table displays a list of groups that are currently active in the data source.

A screenshot of the Collibra Protect interface showing a table of active groups. The interface includes the same header and navigation bar as the previous screenshot. Below the 'Groups' heading, there is a section titled 'Adding Groups' with a sub-heading and a note: 'To add a group, you have to use the Collibra Protect Group API. Currently, only Snowflake data sources are supported.' Below this is a table with the following columns: 'Group Name', 'System Reference', 'Created By', and 'Created date'.

| Group Name | System Reference | Created By | Created date |
|-----------------|--------------------------|----------------|------------------------|
| CID | "Snowflake": "string" | Admin Istrator | Jun 16, 2022, 8:52 AM |
| Human Resources | "Snowflake": "HR" | Admin Istrator | May 11, 2022, 11:39 AM |
| Marketing | "Snowflake": "MARKETING" | Admin Istrator | May 11, 2022, 11:39 AM |

Note Collibra Protect currently only supports the Snowflake data source.

The table columns include:

| Column name | Description |
|------------------|------------------------------------|
| Group Name | Name of the Protect group |
| System Reference | |
| Created By | User who created the Protect group |

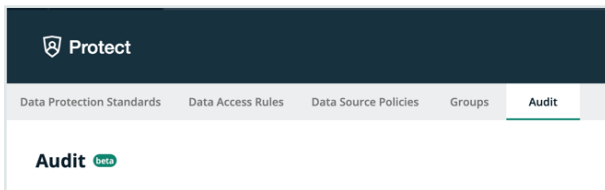
| Column name | Description |
|--------------|----------------------------|
| Created Date | Date the group was created |

Adding groups in Collibra Protect

To add a group, use the [Collibra Protect Group API link](#). This action must be done before any data protection standards or data access rules can be created.

Audit

The Audit page allows you to generate an audit log of access records from the data source .



Generate an audit log

Why would I need to generate an audit log?

Note The actions you take in Collibra Protect only appear in an audit log after one hour.

Steps

1. In Collibra Protect, go to the **Audit** tab.

Click one of the predefined time buttons (Today, Yesterday, A week ago, or 30 days ago) or use the date picker to specify a start date for the audit log.

2. Click the green **Generate Log** button.

» The audit log displays the first 1,000 records after the selected start time.

Note | Important Generating an audit log can take some time. Do not navigate away from the page or your request is canceled. For full details, contact your data source administrator.

Audit 🔍

Today Yesterday A week ago 30 days ago

Start Date: 09/29/2022

Generate Log

For audit log generation, data sources may have latency to summarize access records. Logs generated here for today may not contain information for the most recent access.

| Query ID | Query Start Time | Source User Name | Direct Objects Accessed | Base Objects Accessed |
|--------------------------------------|-----------------------|------------------|---|-----------------------------|
| 01a74800-0501-ec9a-0001-000306f6b19e | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.MAIN_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ec9a-0001-000306f6b1a2 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.MAIN_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ea4f-0001-000306f69dd2 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.DEPENDS_ON_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74800-0501-ec9a-0001-000306f6b1a6 | Sep 29, 2022, 2:00 AM | MELIK | TEST_DB.PUBLIC.NODES_DEPENDS_ON_EXAMPLE | TEST_DB.PUBLIC.MAIN_EXAMPLE |
| 01a74801-0501-ea4f-0001-000306f69dda | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ec9a-0001-000306f6b1b6 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1ba | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ea4f-0001-000306f69de2 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1be | Sep 29, 2022, 2:01 AM | CERTIFICATION | DQ.PUBLIC.EMPLOYEES | DQ.PUBLIC.EMPLOYEES |
| 01a74801-0501-ec9a-0001-000306f6b1c2 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |
| 01a74801-0501-ec9a-0001-000306f6b1c6 | Sep 29, 2022, 2:01 AM | CERTIFICATION | COLLIBRATPCH_SF001.LINEITEM | COLLIBRATPCH_SF001.LINEITEM |

1-50 51-100 101-150 ... 701-712

The audit log table columns include:

| Column name | Description |
|------------------------|---|
| Query ID | The ID of the query in the source DB. |
| Query Start Time | Date and time of the query in the source DB. |
| Source User Name | Name of the user in the source DB that conducted the query (accessed the data). |
| Direct Object Accessed | The DB object that was used to access and view the data (a table or a view). |
| Base Object Accessed | The DB object that was accessed and viewed. |

Why rules or standards fail

Certain rules or standards may fail due to logical errors. This section describes some of the common scenarios that cause them to fail.



Different types of masking affecting the same column

Note In this topic, the term *agent* refers to a data category or a data classification.

Masking within a rule

Scenario

A rule that is set for a group masks multiple agents using different types of masking, and the agents share the same column. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group. The rule masks the **Personal Information** data category by hashing and masks the **Personal and family details** data category by showing only the last two digits. Suppose that both these data categories share the same column. Then, the rule will fail because the same column cannot be masked using two different masking types for a given group.

Rule Name*
Masking within a rule

Description

Set rule for

group* Marketing

asset* Customer Data

and the asset Audit & Internal Controls

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with Hashing for Data Category Data Classification Personal Information

with Show last 2 for Data Category Data Classification Personal and family details

and Select an action rows where Select a data classification has Select a code set Select a code value

Summary
Grant access to Marketing for Customer Data and Audit & Internal Controls with Hashing for Personal Information and with Show last 2 characters for Personal and family details

Masking between rules

This scenario is similar to the [previous scenario](#) except that this scenario considers two rules, instead of one, that are set for the same group. The masking types for the agents in the two rules are different, and both the agents share the same column. Then, a conflict occurs because the same column cannot be masked using two different masking types for a given group.

When two rules conflict with each other, if the synchronization status of only one of them is **Active**, then the other rule fails. If, however, the synchronization status of both the rules is **Active** or **Pending**, then both of them fail.

This scenario is applicable regardless of whether the agents are the same or different, and regardless of whether the rule applies to a single asset or multiple assets.

Rule Name*
Masking between rules - 1

Description

Set rule for

group* Marketing

asset* Customer Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with Hashing for Data Category Data Classification Personal Information

and Select an action rows where Select a data classification has Select a code set Select a code value

Summary
Grant access to Marketing for Customer Data with Hashing for Personal Information

Rule Name*
Masking between rules - 2

Description

Set rule for

group* Marketing

asset* Audit & Internal Controls

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with Show last 2 for Data Category Data Classification Personal and family details

and Select an action rows where Select a data classification has Select a code set Select a code value

Summary
Grant access to Marketing for Audit & Internal Controls with Show last 2 characters for Personal and family details

Masking between standards

Scenario

Two standards mask different agents, and the agents share the same column. This scenario is applicable regardless of whether the groups and the masking types are the same or different.

Example

Consider two standards. The first standard masks the **Personal Information** data category, and the second standard masks the **Name** data classification. Suppose that both the agents share the same column. Then, a conflict occurs because more than one standard cannot be applied to the same column via different agents.

Note This is a limitation on how Collibra Protect implements standards on Snowflake.

When two standards conflict with each other, if the synchronization status of only one of them is **Active**, then the other standard fails. If, however, the synchronization status of both the standards is **Active** or **Pending**, then both of them fail.

Standard Name *

Description

for the group * + -

protect * Data Category Data Classification

with * ⓘ

Summary
 For the Group Marketing
 protect [Personal Information](#)
 with Hashing

Standard Name *
 Masking between standards - 2

Description

for the group * Human Resources

protect * Data Category **Data Classification** Name

with * ⓘ Show last 2

Summary
 For the Group Human Resources
 protect Name
 with Show last 2

Conflicting filters affecting the same column

Filtering within a rule for the same data classification

Scenario

A rule that is set for a group contains conflicting filters for the same data classification. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group and the **Customer Data** asset. The rule contains two filters for the **Country** data classification.

Rule Name *
Filtering within a rule for the same data classification

Description

Set rule for

group * Marketing

asset * Customer Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for **Data Category** Data Classification Select a data category

and Show rows where Country has Country code BE

and Hide rows where Country has Country code PL

Summary
Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE
and Hide rows where Country has Country code: PL

If any of the tables in the asset contain a column that is classified as **Country**:

- The first filter shows the rows that contain **BE** in that column.
- The second filter hides the rows that contain **PL** in that column.

Then, this rule will fail because two conflicting filters affect the same column.

When applying a filter for a specific data classification, you must select only one type of action. That is, you can choose to either show rows based on one or more values or hide rows based on one or more values. You must not use the show and hide filter actions together for the same data classification.

Filtering within a rule for different data classifications

Scenario

A rule that is set for a group contains conflicting filters for different data classifications that share the same column. This scenario is applicable regardless of whether the rule applies to a single asset or multiple assets.

Example

Consider a rule that is set for the **Marketing** group and the **Customer Data** asset. The rule contains two filters: one for the **Country** data classification, and another for the **State** data classification.

Rule Name *
Filtering within a rule for different data classifications

Description

Set rule for

group * Marketing + -

asset * Customer Data + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option Select a data category

for **Data Category** **Data Classification**

and Show + -
rows where Country Country code BE + -

and Hide + -
rows where State Country code PL + -

Summary

Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE
and Hide rows where State has Country code: PL

If any of the tables in the asset contain columns that are classified as **Country**, the first filter shows only the rows that contain **BE** in those columns.

If any of the tables in the asset contain columns that are classified as **State**, the second filter hides only the rows that contain **PL** in those columns.

Suppose that a column is classified as both **Country** and **State**. That is, data classifications **Country** and **State** share the same column. Then, this rule will fail because two conflicting filters affect the same column.

Filtering between rules for same or different data classifications

This scenario is similar to the [previous scenarios](#) except that this scenario considers two rules, instead of one, that are set for the same group. The filter in one rule is different from the filter in

the other rule, and both the filters affect the same column. Then, a conflict occurs because two conflicting filters affect the same column.

When two rules conflict with each other, if the synchronization status of only one of them is **Active**, then the other rule fails. If, however, the synchronization status of both the rules is **Active** or **Pending**, then both of them fail.

Rule Name*
Filtering between rules for same or different data classifications - 1

Description

Set rule for

group* Marketing + -

asset* Customer Data + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for **Data Category** **Data Classification** Select a data category

and Show rows where Country has Country code BE + -

Summary
Grant access to Marketing
for Customer Data
and Show rows where Country has Country code: BE

Rule Name*
Filtering between rules for same or different data classifications - 2

Description

Set rule for

group* Marketing + -

asset* Personal Information + -

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to tables in the database that are linked to the selected assets. If this box is unchecked, no access will be given to the selected tables. Note: Once the rule granting access is saved and synchronized, access to these tables cannot be revoked through Collibra Protect. It can only be revoked by direct action on the data source.

with ⓘ Select a masking option for **Data Category** **Data Classification** Select a data category

and Hide rows where Country has Country code PL + -

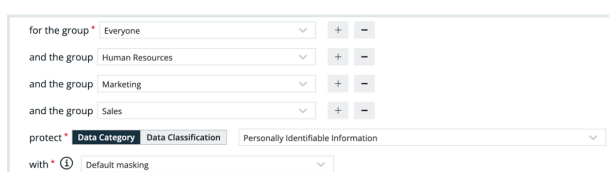
Summary
Grant access to Marketing
for Personal Information
and Hide rows where Country has Country code: PL

Reference documentation

As described in the [DB synchronization](#) section, Collibra Protect periodically does an aggregation of all data protection standards and data access rules available. These standards and rules prepare a representation containing all databases, schemas, tables, and columns involved as well as their protections and accesses. The synchronization process then triggers Edge capabilities, like Collibra Protect for Snowflake, that are responsible for translating the representation to actions toward the data source provider using their technology. This process might involve JDBC and REST calls to perform low-level operations to guarantee that the protections and accesses are applied.

Collibra Protect for Snowflake

Data protection standards rely on [tag-based masking policies](#) available in Snowflake. The name of the data category or data classification specified in the standard becomes a tag, which is applied to all affected columns to enforce data protection. For example, let's say a standard is created on the Personally Identifiable Information data category to restrict access for different groups with the organization.



for the group + -
and the group + -
and the group + -
and the group + -
protect
with

When synchronized and active, the standard resolves to 14 masking policies, which is one for each Snowflake data type. The masking policies are created at the schema level and use the following naming convention: COLLIBRA/MASKING_POLICY/<asset ID>/<snowflake type>.

Chapter 12

| Row | created_on | name ↑ | database_name | schema_name | kind | owner |
|-----|------------------------|--|---------------|-------------|----------------|--------------|
| 1 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/ARRAY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 2 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/BINARY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 3 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/BOOLEAN | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 4 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/DATE | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 5 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/FLOAT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 6 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/GEOGRAPHY | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 7 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/NUMBER | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 8 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/OBJECT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 9 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/STRING | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 10 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIME | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 11 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 12 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP_LTZ | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 13 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/TIMESTAMP_TZ | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 14 | 2022-09-06 03:41:13... | COLLIBRA/MASKING_POLICY/28d226cc-0ab0-4d23-b912-985312fb36b1/VARIANT | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |

The example below shows a masking policy created for the type STRING.

Note The data the consumers see depends on the masking option selected. Go to the Masking and Data Types page to learn more.

```
1 CASE
2   WHEN CURRENT_ROLE() = 'PUBLIC' THEN '*'
3   WHEN CURRENT_ROLE() = 'HR' THEN '*'
4   WHEN CURRENT_ROLE() = 'MARKETING' THEN '*'
5   WHEN CURRENT_ROLE() = 'SALES' THEN '*'
6   ELSE val
7 END
```

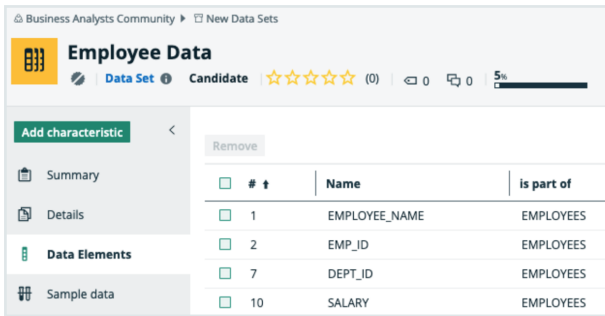
Done

All masking policies are then associated with the Personally Identifiable Information tab, which is created at the schema level and assigned to all columns where the protection needs to be applied. At runtime, Snowflake fetches the right masking policy based on column data type.

| Row | created_on | name | database_name | schema_name | owner | comment |
|-----|------------------------|-------------------------------------|---------------|-------------|--------------|---|
| 1 | 2022-09-06 03:41:13... | Personally Identifiable Information | PROTECT_QA | DEMO | ACCOUNTADMIN | Generated by Collibra: 28d226cc-0ab0-4d23-b912-985312fb36b1 |

Data Access Rules are translated as a combination of [grant instructions](#), [dynamic masking](#), and [row access policies](#) when specified in the rule. For example, a data set named Employee Data has sensitive columns categorized as Personally Identifiable Information.

Chapter 12



| # | Name | is part of |
|----|---------------|------------|
| 1 | EMPLOYEE_NAME | EMPLOYEES |
| 2 | EMP_ID | EMPLOYEES |
| 7 | DEPT_ID | EMPLOYEES |
| 10 | SALARY | EMPLOYEES |

In Collibra Protect, a rule is created to grant access of that data set to Human Resources. Since the Grant Access checkbox is enabled, each database, schema, and table in that data set received a grant for the Snowflake role specified and each column that has protection received a column masking policy.



Set rule for

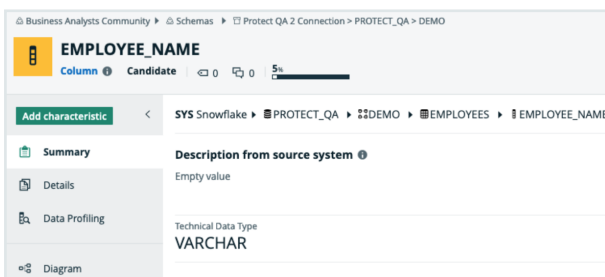
group: Human Resources

asset: Employee Data

Grant access to all data tables linked to these asset columns.
By checking this box, access will be given to the tables in the database with columns linked to the selected assets. If this box is unchecked, no access will be given to these columns.

with: No masking for Data Category: Data Classification, Personally identifiable information

Let's look closer at one of the columns, such as EMPLOYEE_NAME. It belongs to the EMPLOYEES table within the DEMO schema within the PROTECT_QA database.



EMPLOYEE_NAME

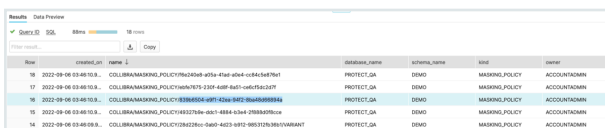
Column Candidate

Add characteristic

Summary: Description from source system: Empty value

Details: Technical Data Type: VARCHAR

In Snowflake, each column has a masking policy assigned to it. The masking policies created at the schema level follow the naming convention: COLLIBRA/MASKING_POLICY/<asset ID>.



| Row | created_at | name | database_name | schema_name | kind | owner |
|-----|-----------------------|--|---------------|-------------|----------------|--------------|
| 19 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY78620268-9734-47af-8044-c684-c0487961 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 17 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY74061679-2337-468f-804f-c6c0793237 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 18 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY93869506-6947-424e-8042-9646909984 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 15 | 2022-09-08 03:48:10.3 | COLLIBRAMASKING_POLICY48327806-0651-4884-8344-2388808004 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |
| 14 | 2022-09-08 03:48:09.8 | COLLIBRAMASKING_POLICY32822200-0467-4423-8012-963372938510481847 | PROTECT_QA | DEMO | MASKING_POLICY | ACCOUNTADMIN |

The content of the masking policy created for the column EMPLOYEE_NAME is shown below.


```
Details
1 CASE
2     WHEN CURRENT_ROLE() = 'HR' THEN va1
3     WHEN CURRENT_ROLE() = 'PUBLIC' THEN '*'
4     WHEN CURRENT_ROLE() = 'MARKETING' THEN '*'
5     WHEN CURRENT_ROLE() = 'SALES' THEN '*'
6     ELSE va1
7 END
```

Done

The Human Resources group has access to the plain data without any masking while the other groups have masked access as created in the data protection standard.

Important In this example, the column EMPLOYEE_NAME has the policy tag and a column masking policy assigned to it. In Snowflake, when both are present, the column masking policy takes precedence and the policy tag is not executed. To mitigate this behavior and ensure that the protection defined in the standard is applied, we prepare the column masking policy with the conditions defined in the policy tag.

Masking and data types

Snowflake provides several functions to transform the data. In Collibra Protect, we support four masking options.

1. Default Masking is not supported by Snowflake. This implementation was added in our Protect capability, so protection can be applied to a wide range of data types. Each column received a default value according to the column data type. Below is a list of Snowflake data types and their default values.

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|-----------------------|
| 1 | NUMBER | NUMBER | 0 |
| 2 | DECIMAL | NUMBER | 0 |
| 3 | NUMERIC | NUMBER | 0 |
| 4 | INT | NUMBER | 0 |
| 5 | INTEGER | NUMBER | 0 |
| 6 | BIGINT | NUMBER | 0 |
| 7 | SMALLINT | NUMBER | 0 |
| 8 | TINYINT | NUMBER | 0 |
| 9 | BYTEINT | FLOAT | 0 |
| 10 | FLOAT | FLOAT | 0 |
| 11 | FLOAT4 | FLOAT | 0 |
| 12 | FLOAT8 | FLOAT | 0 |

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|----------------------------|
| 13 | DOUBLE | FLOAT | 0 |
| 14 | DOUBLE PRECISION | FLOAT | 0 |
| 15 | REAL | FLOAT | 0 |
| 16 | VARCHAR | VARCHAR | * |
| 17 | CHAR | VARCHAR | * |
| 18 | CHARACTER | VARCHAR | * |
| 19 | STRING | VARCHAR | * |
| 20 | TEXT | VARCHAR | * |
| 21 | BINARY | BINARY | 00 |
| 22 | VARBINARY | BINARY | 00 |
| 23 | BOOLEAN | BOOLEAN | false |
| 24 | DATE | DATE | 1970-01-01 |
| 25 | DATETIME | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |
| 26 | TIME | TIME | 00:00:00 |
| 27 | TIMESTAMP | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |

| | Column Data Type | Snowflake Data Type | Default Masking Value |
|----|------------------|---------------------|--|
| 28 | TIMESTAMP_LTZ | TIMESTAMP_LTZ | 1969-12-31 16:00:00.000-0800 <i>Might change based on user TZ</i> |
| 29 | TIMESTAMP_NTZ | TIMESTAMP_NTZ | 1970-01-01 00:00:00.000 |
| 30 | TIMESTAMP_TZ | TIMESTAMP_TZ | 1969-12-31 16:00:00.000-0800 <i>Might change based on user TZ</i> |
| 31 | VARIANT | VARIANT | 0 |
| 32 | OBJECT | OBJECT | {} |
| 33 | ARRAY | ARRAY | [] |
| 34 | GEOGRAPHY | GEOGRAPHY | {"coordinates": [0,0], "type": "Point"} (aka point(0, 0) and visualization can change based on user preferences) |

In Collibra Protect, we also support the hashing and show last masking options. These can only be applied to Snowflake data types STRING, NUMBER, and FLOAT.

2. Hashing allows us to use Snowflake's SHA2 value function for strings, and the HASH value for numbers
3. Show Last allows us to use the substr(to_varchar(value), length(value) - n, n) expression for strings, and mod(value, power(10,n)) for numbers. Value is the content and n is the number of characters to be shown.
4. No Masking is when the raw content is returned.

Note Whenever a masking option cannot be applied, like hashing on the DATE type, default masking is applied, so protection is guaranteed.