



Collibra Data Intelligence Cloud
Collibra Protect

Collibra Data Intelligence CloudCollibra Data Governance Center - Collibra Protect

Release date: Thursday, February 2, 2023

Revision date: Thu Feb 02, 2023

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 Protect	i
Scenarios for using Protect	i
Install Collibra Protect	xxi
Configure Collibra Protect	xxiii
Essentials for Collibra Protect	xxviii
Open Protect	xlvii
Data protection standards	xliv
Data access rules	lv
Data source policies	lxiv
Groups	lxvi
Audit	lxviii
Data protection in the asset pages	lxix
Why rules or standards fail	lxxii
Reference	lxxx
Collibra Protect	xcv
About Protect	i
Scenarios for using Protect	i
Install Collibra Protect	xxi
Configure Collibra Protect	xxiii
Essentials for Collibra Protect	xxviii
Open Protect	xlvii
Data protection standards	xliv

Data access rules	lv
Data source policies	lxiv
Groups	lxvi
Audit	lxviii
Data protection in the asset pages	lxix
Why rules or standards fail	lxxii
Reference	lxxx

About Protect

Collibra Protect is a capability of the Collibra Data Intelligence Cloud to protect sensitive data and grant varying levels of access to the data to specific groups of people through policies that do not require you to code. You can enforce data protection at the source database level directly from the Collibra Protect interface, and apply advanced data protection through masking, redacting, and hashing. Protect simplifies access governance and eliminates the need for repetitive actions and approvals. By providing permission to view information to only those who need it, Protect minimizes risk and promotes a safe data culture in your organization.

You can use Protect to protect the data in the assets of the packaged asset types, such as Business Process, Data Category, and Data Set, in addition to the assets of any new or modified asset types. In addition, you can use Protect to provide differential access, for example, to give everyone access to a data set but allow certain type of access to only certain groups of people based on data categories.

Scenarios for using Protect

This topic describes how Collibra Protect helps you to:

- Use the metamodel graph to build and enforce protection policies on Business Processes, Data Categories, and Data Sets.
- Use classifications to apply a broad coverage of protection mechanisms at the data source.
- Support privacy preferences such as consent management, data subject requests such as access requests, and the right to be forgotten through row-filtering mechanisms.
- Perform an audit of applicable protection at the data source and use reporting to demonstrate compliance where data is stored and consumed.

Discover and classify personal information

Suppose that you want to help your organization find personal information.

To achieve this, typically, your Privacy team sets up the Data Classification Policy, where they classify the data used in the organization based on the sensitivity or the business criticality of the data. This determines the required levels of security for the applications that store that data or the applications that are used for the transit of the data.

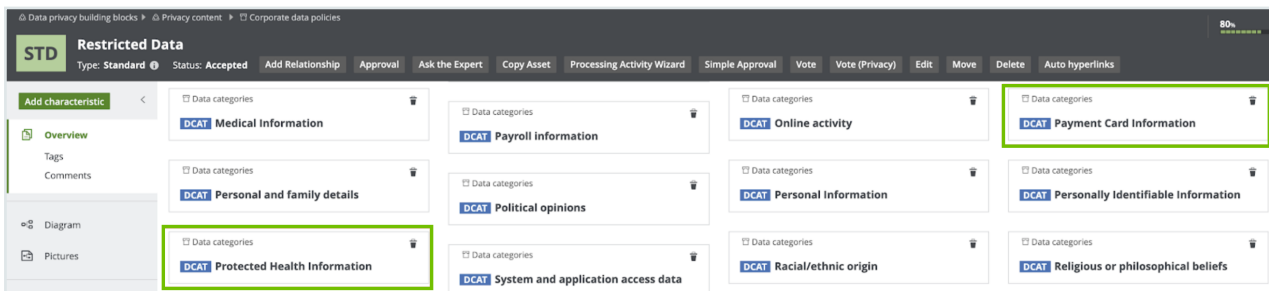
Consider the following three classifications for sensitivity:

- Public data, which is least sensitive.
- Private data, which is slightly more sensitive than the public data.
- Restricted data, which is the most sensitive data and therefore requires the highest level of access controls and security protection.

The following image shows the standard subassets of the Data Classification policy.

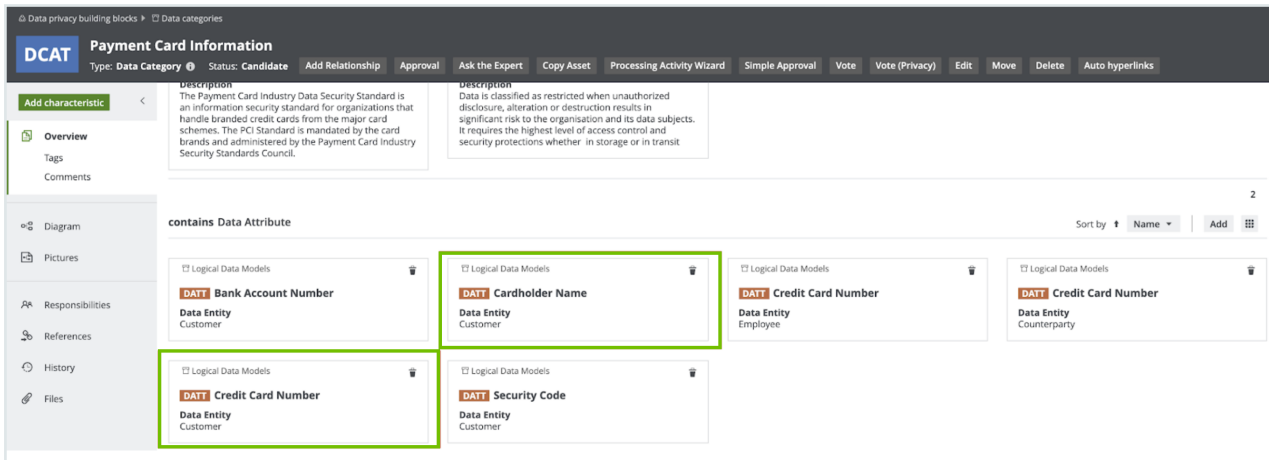


The Privacy team determines the data categories to which these subassets apply. For example, they can determine that Restricted Data applies to the following data categories: Gender, Social Security Number, Payment Card Information.

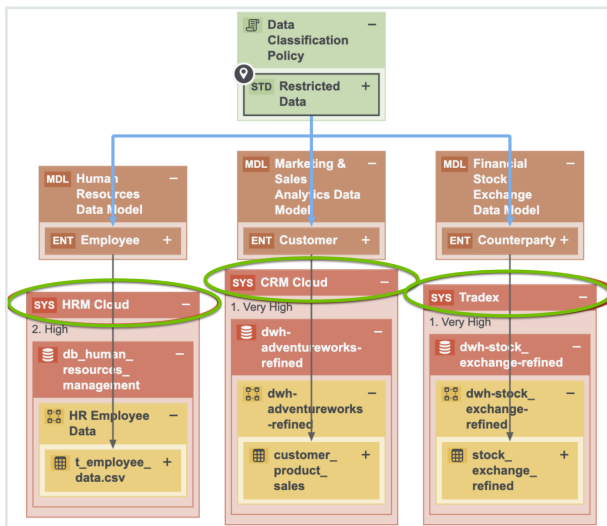


The Privacy team determines the sensitivity and the required security at the data category level as opposed to the column level. At the data category level, the Privacy team then determines what data elements belong to the identified data categories. For example, the

Payment Card Information data category groups the Cardholder Name and the Credit Card Number, among other information.



In this model, Data Attributes are grouped under the Data Category. This is how the Privacy layer is linked to the logical data model. This promotes collaboration between the Privacy team and the Governance team. In addition, this allows the automated data classification of the organisation's personal information, which makes views such as the Restricted Data Overview diagram available at the most sensitive data category, Standard Restricted Data.

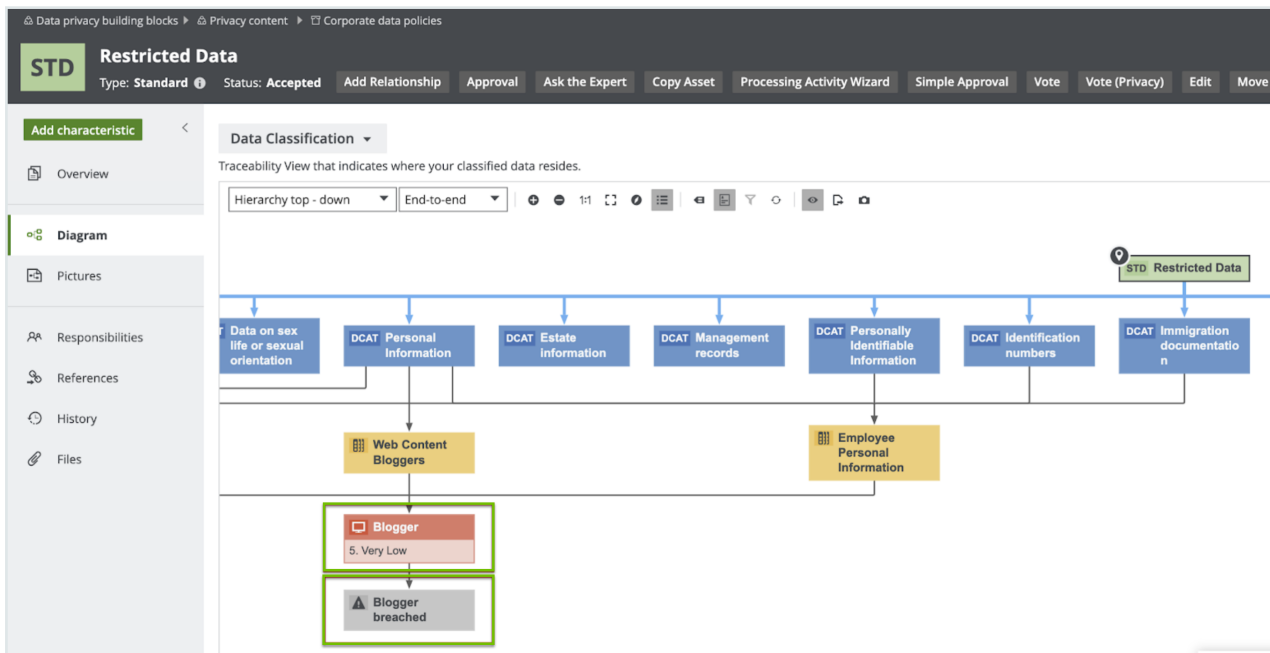


In the above image, the applications in which the restricted data resides are highlighted.

The Privacy team determines the Policies and Standards that determine which data categories are sensitive to the organization and what the required levels of protection are. The Data Governance team maps those data categories to the applications where that

data resides. The Security team determines what the security levels on those applications are. Thus, the view captured in the above image requires collaboration among teams.

Consider the traceability diagram called Data Classification under the Restricted Data standard. This standard contains the most sensitive information and thus requires the highest level of security controls; however, it resides on an application that has very low security. Because of this, the Information Security team needs to take the necessary remediation actions and improve the security levels on Blogger. As shown in the image, an investigation is already ongoing on the potential data breach on Blogger.



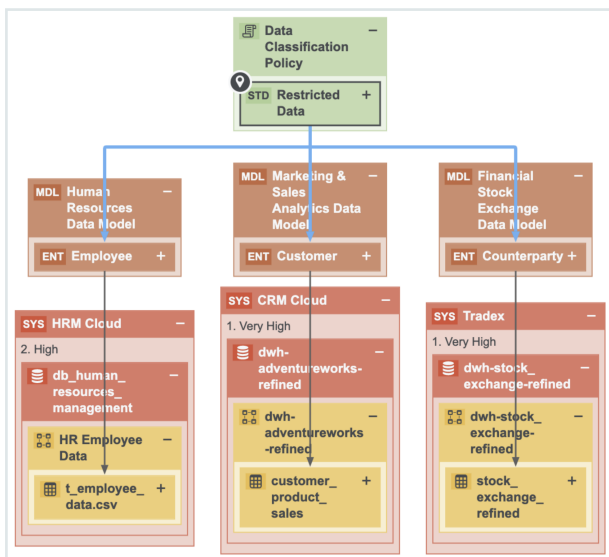
Data classification capabilities and guided stewardship

This section describes how Collibra Privacy and Risk leverages the data classification capabilities in Catalog. Thus far, we learned that the Restricted Data standard groups Data Categories, which group Data Attributes. In the example, the Payment Card Information data category contains the Credit Card Number data attribute.

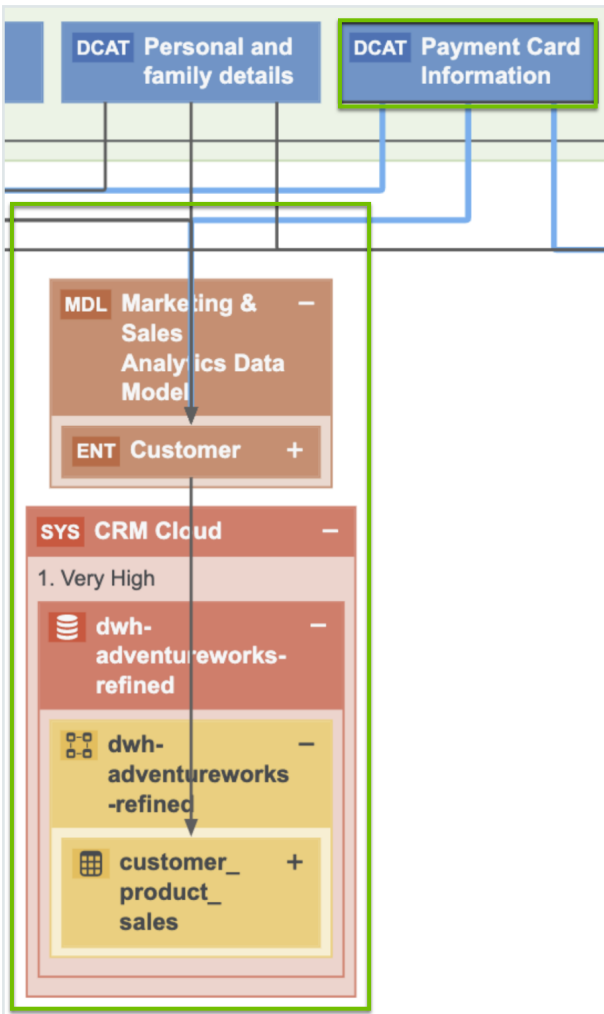
Guided stewardship is a semi-automated process of mapping columns and tables to logical data attributes. It enables content tables to be mapped to data attributes. After scanning a table and then applying guided stewardship in which the steward selects

attributes from the suggestions coming from the automated mapping, the column is mapped to the Credit Card Number. Moreover, when a column is mapped to a data attribute, the column is also mapped to a data category because of the relation between the data category and the data attribute.

The result of classifying one application with the Catalog's Data Classification is shown in the following image.

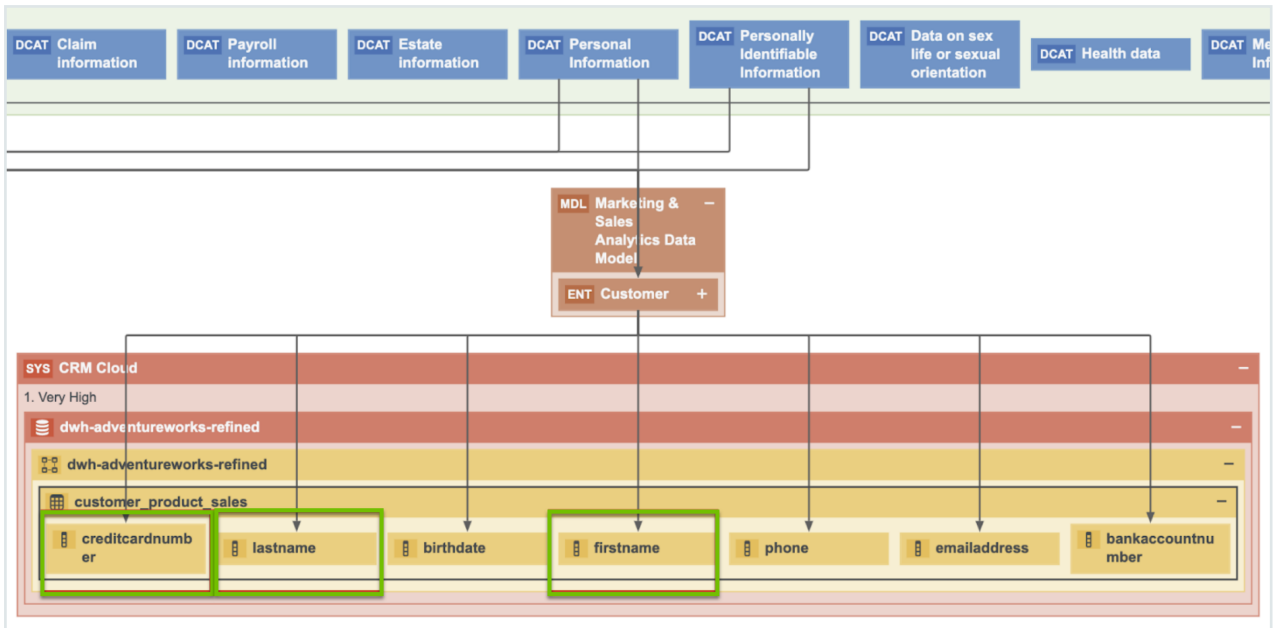


Restricted Data groups multiple data categories. The following image shows the data attributes that the Payment Card Information data category groups.



By applying guided stewardship and data classification, the data attributes are mapped to the columns. Thus, by using Catalog’s data classification capabilities, the Data Governance team can find personal information and sensitive personal information.

It is important to know the context to determine which information is considered personal information. For example, Name can be the name of a customer or an employee, in which case Name is considered personal information. Name can also be the name of another organization. This context can be provided only by a steward. Therefore, data classification and guided stewardship will help the steward mapping customer’s names to the Name column. Because the Privacy team has mapped names and family details, you can safely assume that this is Personal Information. Similarly, Credit Card Number can be the credit card number of another organization, but it is the steward who has mapped the number to the Credit Card Number data attribute belonging to the Customer data entity, and as a result, we know that the payment card information is very restricted data.



This is an example of how guided stewardship, Catalog’s data classification combined with guided stewardship and CollibraPrivacy and Risk, gives you a vertical view on where Personal Information resides.

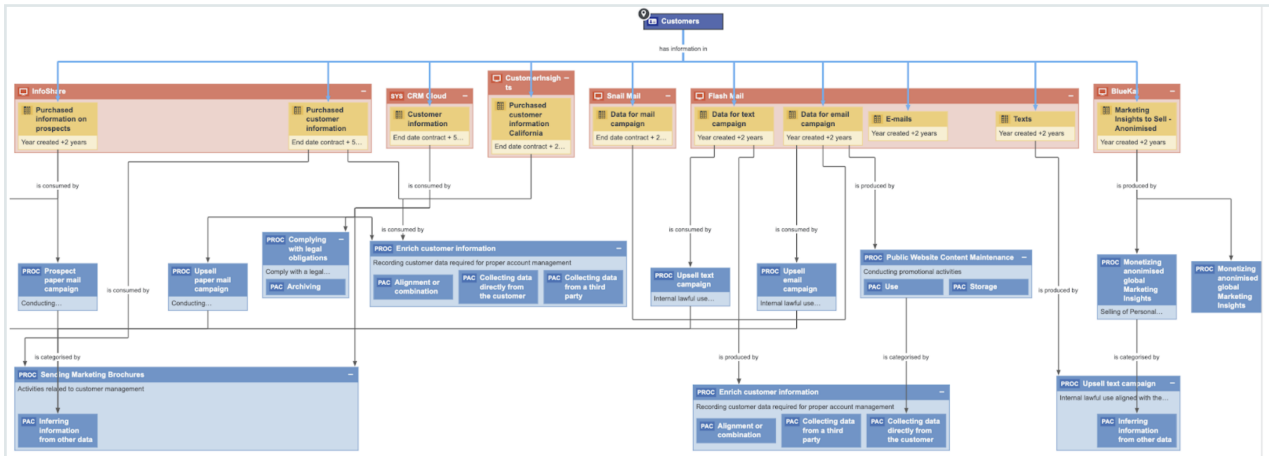
Customer requests and consent management

The previous sections described how we help customers find their Personal Information across applications. This section describes how we help customers manage data subject requests and consent. Collibra has the relevant metadata that is necessary for a partner application that fulfils the data subject requests or manages consent to operate. These applications need the metadata about where the data resides, where you store customer information, how you use the information, why you use the information, and what your legal basis is, so that they can determine for which applications you need consent and for which processes you need instance for a consent. Collibra has and governs the required metadata. In addition, through APIs, Collibra can integrate with those applications to feed them with the metadata that they need to function.

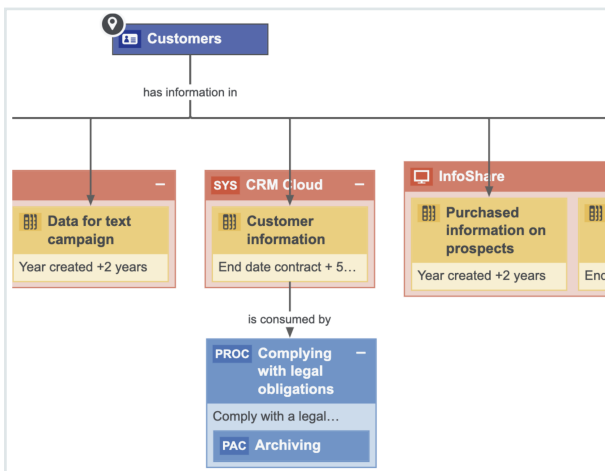
Consider the customer data. Collibra knows where this data resides and how it is being used. This is an outcome of obtaining input from the business users during the onboard of the Business Processes where the users are asked what data they use, which applications they use, for what purpose the data is used. When further onboarding of those business processes by the stewards takes place, one of these steps is mapping the Business

Processes to the data, and then also helping those business stewards with the mapping through the data classification capabilities in Catalog.

The following image shows a traceability view, which is a result of collaboration with the business team, data governance team, and other teams.



The above image shows where data resides and why it is used. It shows all the applications that contain customer data, and also the related retention periods, which can be imported when a customer wants to exercise their right to be forgotten. Collibra knows in which applications the data resides and the business processes that use that data. Thus, we know why and how we are using our customer data. This determines how to respond to the right to be forgotten because there are often Business Processes where you have the real legitimate reason to retain the customer's personal information.



When a customer wants to exercise their right to be forgotten, we can remove the information in these applications; however, we need to store the customer information in

the above table in order to comply with the legal obligation. Therefore, it is not only important to know where your personal information resides, but also why you are using it. Such information is important information for applications that process data subject requests (DSRs). You can integrate with the application that does the DSRs and create a workflow to process the data subject requests. Based on the input of the information and metadata that you will find in Collibra, you can validate the request. When the request is approved, you can point the applications to the stewards and send them a task to perform the action that appears in the data subject request, such as, removing the data or extracting the data and sending it to a customer.

The same approach can be applied to the integrated consent management applications. These applications need to know the processes for reaching the consent, and such applications reside in the process register, so that you can see all the processes that rely on the consent and the data categories for which you need consent.

The screenshot shows the 'Marketing Process Register' interface. At the top, there are navigation buttons: 'Type: Process Register', 'Export Metamodel', 'Go to the Business User Interface', 'Request input', 'Edit', 'Move', 'Delete', and 'Auto hyperlinks'. Below this, there is a section for 'CCPA Default View' with a dropdown arrow and a description: 'The view presents the inventory of Business Processes describing the data flows in your organization.' There are also buttons for 'Delete', 'Move', and 'Validate'. The main content is a table with two columns: 'Name' and 'legal basis'. The table lists various business processes and their corresponding legal bases.

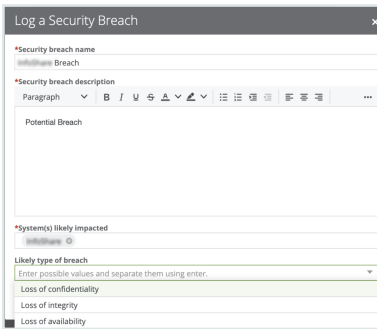
Name	legal basis
Direct Marketing	Legitimate interest
Market Research	Legitimate interest
Monetizing Marketing Insights	Consent, Consent from the minor towards selling of PI
Monetizing anonimised global Marketing Insights	Consent, Opt-out (from selling)
Monetizing Marketing Insights EU customers	Consent
Monetizing Marketing Insights US customers	Consent provided towards selling of PI due to financial incentive received,
Print media advertisement	Legitimate interest
Public Website Management	Consent provided towards selling of PI due to financial incentive received,
Public Website Content Maintenance	Consent, Substantial Public Interest
Create online contest	Consent

These are stored in the data sets that can also contain granular information, such as the individual data elements for which you want to obtain consent—this combines the information about which business processes require consent and the data categories for which you need consent to process all information in Collibra. The information governed in Collibra can be then sent to the consent management application that is used to manage consent.

Potential data breach workflow

This section describes how Collibra helps when a data breach occurs.

With Collibra Data Privacy, Collibra for Desktop, or Collibra for Mobile, you can report any suspicious behavior by logging a potential data breach.



If your organization has suffered a potential data breach, you can determine the application that needs to be investigated and the type of breach that may have occurred, and then log a potential data breach. The related workflow will require the community manager on the data governance counsel to assign the issue manager who will investigate the breach. The issue manager will then investigate the issue, assess the potential impact of the breach, determine the reporting requirements (for example, to whom the incident must be reported), and plan the remediation actions to address the risks. The reporting evidence needs to be stored. If you go to the data helpdesk, you can find an overview of all the breaches that are being investigated.

Name ↑	Description	Assignee	Requester	Reviewer
BigSuite - sent credentials ove...	Employee accidentally sent	Preston Sterling	William Parker	Dora Perreman
Data Breach Blogger	Today it is mentioned in the new	Preston Sterling	David English	Dora Perreman
Example of Breach	Description			

Collibra can help with investigating the impact of the breach because of the knowledge of which data resides in the applications and the processes that use those applications. Such a holistic view on where the data resides, which applications are involved, and the processes that rely on these applications can help in assessing the impact on customers following a data breach. Collibra can not only help an organization log and investigate a data breach but also help analyze the impact of the breaches because Collibra knows

where the data resides and how it is being used. In addition, it contains a history of all the breaches (including potential ones) that would have been logged.

How do we get there?

This section describes the Process register and Business Process discovery capabilities, data categorization and classification, and different prescriptive paths for reaching out from the logical data layer envisioned in the metamodel graph and connected data sets to a physical data layer present in columns located directly at the data source.

Create and maintain Process Register

Process Register is an essential part of privacy compliance, foreseen directly by GDPR article 30 as a Record of Processing Activities and derived from CCPA requirements for performing data mapping in the organization. Process Register enables to store assets of the Business Process type that describes processes in the organization that involve personal data. In Collibra, Business Processes reflect the requirements stated by Processing Activity in GDPR.

Business Process onboarding

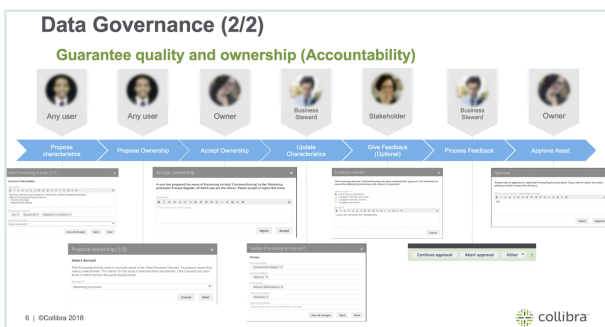
Business Processes may be onboarded by business users as well as privacy stewards through dedicated workflow implementing guided stewardship principle in Collibra Data Privacy. During onboarding, multiple roles collaborate in providing content to the onboarded Business Process. Because of the dedicated tasks and required approval and feedback, assets are onboarded in a governed way.

In the scenario on the Personal Information (PI) Discovery, it was described how Collibra helps with discovering Personal Information. But equally important to knowing where you are storing personal information is knowing why you are using personal information. That is, what the legal context of using that PI is. This context is created within Process Registers, throughout the usage of Business Processes that describe the processes conducted by organization relating to the usage of personal information.

Typically, that information does not reside with one person that can help you document that knowledge. That information is stored within multiple areas across the organization

and it may not be easy to centralize this information and ensure that the information is up to date. To help you with this task, CollibraData Privacy comes with the Business Process discovery capabilities.

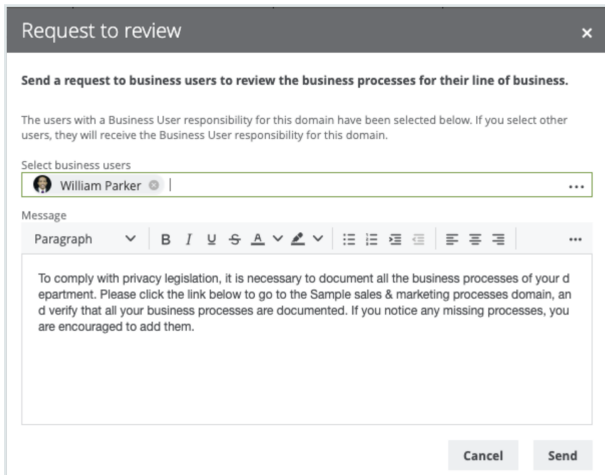
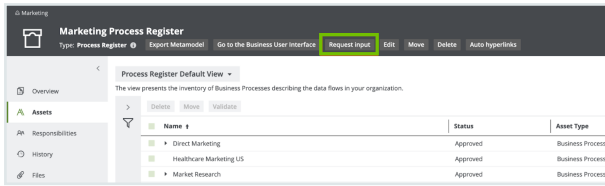
Consider a high-level overview of Collibra Privacy and Risk Business Process discovery capabilities. It commences with the Business Users describing the Business Processes in their terms. They will describe the data being used, applications being used, and any third parties with which they share information. After describing the Business Process, the owner of the Business Process will accept the ownership of that particular Business Process. When the ownership is accepted, the experts or the stewards will further onboard the proposed Business Process. This means that they will ensure that the Business Process is accurate and actionable because that Business Process provides business context on how we use personal information and we must ensure that the description is accurate. Therefore, in principle, you will have the Business Steward, Privacy Steward, and Data Steward, each adding business metadata, adding privacy metadata, and performing data mapping, respectively. After the stewards have updated the characteristics, you can optionally obtain feedback from the stakeholders. The following sections describe each step involved in the process.



Requesting business users' input with a dedicated interface

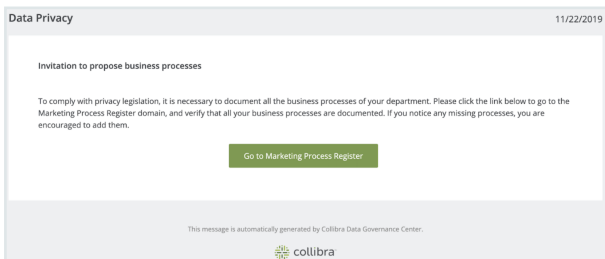
The information related to Business Processes may be requested from the Business User directly from Collibra Privacy and Risk Process Register. Typically, this will be done by those who work on the Privacy program. With the Request input button, email will be generated for the selected business users, which can provide relevant information on the business side of the process through a dedicated interface. You can have a guiding text that explains the purpose of your request. If you click Send, an email is sent to the business user with an invitation to contribute to the Process Register.

Chapter 1

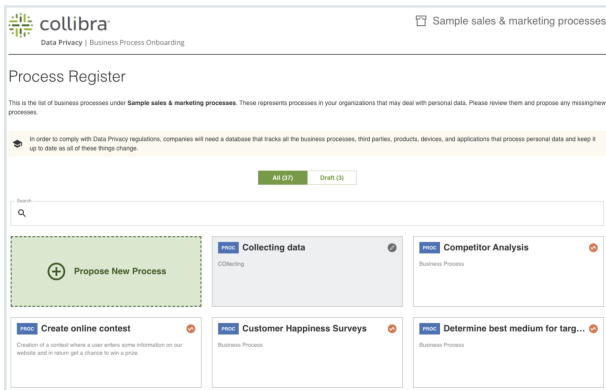


Onboarding Business Process with a business user interface

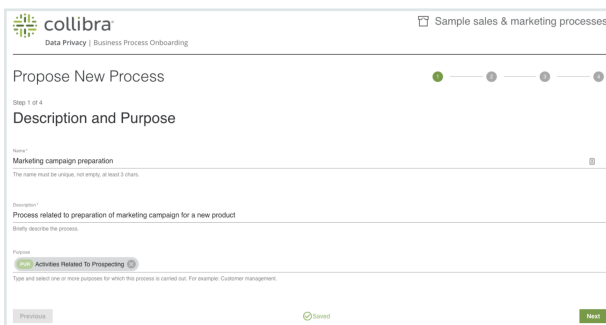
The Business Register User receives an email message asking them to verify that all the processes are in their domain.



When the Business User clicks Go to Marketing Process Register in the email, a page showing all the Business Processes for their department appears to allow the Business User to contribute to the Process Register.



The link provided in the email message directs the User to a survey where they can describe the business processes that they perform on a daily basis. If the Business User cannot find the Business Process that was onboarded was in the process of being onboarded, they can propose a new Business Process using the Propose Business Process button. When proposing a Business Process, they can describe the Business Process, provide a unique name, description, and purpose.

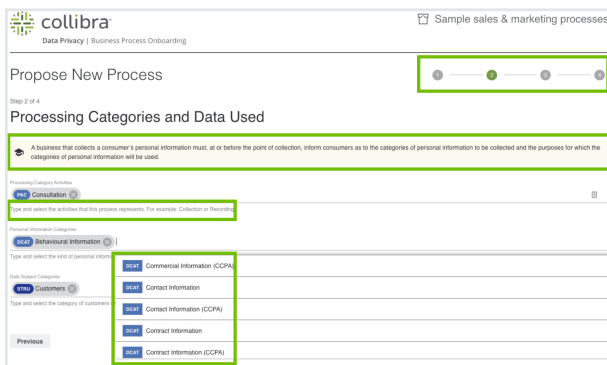


The next step involves covering Process Categories such as use, collection, adaptation, and alteration. The Business User defines the types of data that they are using, for example, behavioral information, contact information, or contract information. Finally, they determine what type of customer's data they are using, such as the customers covered by CCPA or GDPR. There can be also an indication on other types of data subjects, such as, employees and candidates. The Business User can select values only from predefined lists—this reduces the scope of errors as there is no ambiguity around the values that the Business User is able to provide. These values have been predefined by the Privacy team and have legal implications. They show how the organization complies with the privacy regulations. Because, when you collect data directly from customers or from a third parties—by using sensitive information, public information, or customer or employee information—the distinctions will have an impact on how you comply with the regulations.

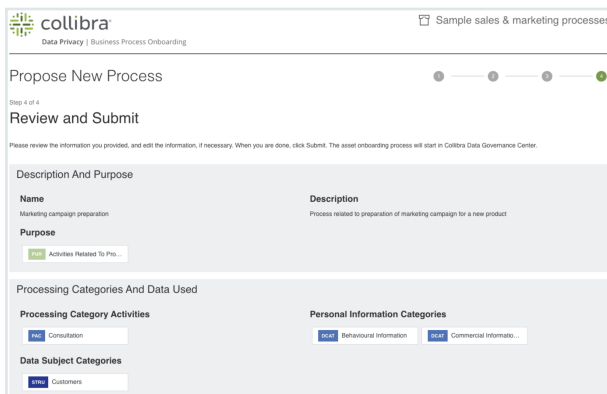
For example, the employee information is temporarily exempted from the CCPA. Therefore, it is considered better for the Business User to select from the drop-down list, as opposed to providing free text. This also prevents common issues such as spelling errors. In addition, if there is any uncertainty about the meaning of these values, the Business User can look up the definitions of these values in Collibra. In the next steps, the lines of business and third parties involved can be described, applications used can be indicated, and the level of automation in the Process can be determined.

The wizard is prescriptive:

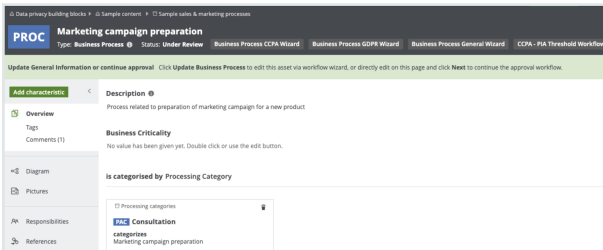
- It shows the user the steps that they have completed and how many steps are remaining, by visually indicating the progress.
- The help text below the question describes what is required from a particular question.
- The ability to open a side panel that provides additional educational information such as the wordings from the law or video content from the Collibra university.
- Smart suggestion based on what the user has already filled and the domain to which they belong.



After entering the information, the Business User can review it before submitting it.

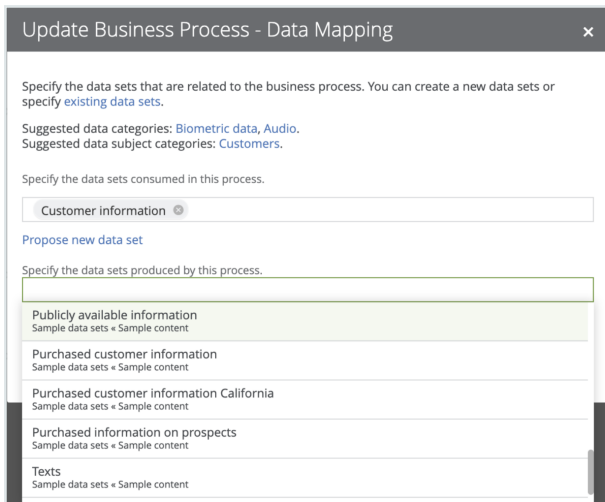


After the Business User provides their Business Process, they can submit the process for further onboarding. The next step is for the owner to accept the ownership of that Business Process. A new task is generated for the owner after they accept the ownership of the Business Process in the Process Register. Based on the metadata, the owner can determine that the Business Process belongs to their Process Register. The ownership can be accepted or rejected. As a result, the status of the asset is changed and the justification is added in the Comments section of the Business Process.

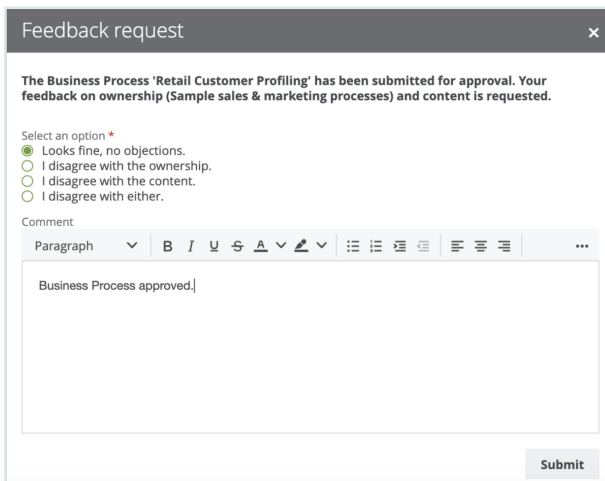


As a next step, the experts or the stewards will be consulted to ensure that the metadata is accurate and complete and the Business Process is going to be mapped to the data. In the following steps, relevant tasks will be created with the request to review and update attributes when necessary. Among others, the Data Mapping task is performed by Data Stewards. The form contains contextual help with suggestions on the relevant Data Sets consumed and produced by the Business Process. Because one of the data categories used in the process is behavioral information, you can click and review it and view the related data sets in the categories. Based on this, Data Stewards can ensure that Data Mapping has been correctly performed.

The last expert who needs to contribute to the Business Process is the Privacy Steward. After opening the task, the first step is to define the regulation that applies, be it GDPR, CCPA, or others. In addition, a purpose needs to be validated, and legal bases, controllers and processors need to be defined. Very specific information on regulation shall be specified, for example, on GDPR, we define cross-border transfers, safeguards, consent collection method, and automated decision-making confirmation. On the CCPA side, we are asked about the collection directly from customers or third parties and whether the data is being sold to third parties.



After the Stewards finish updating the Business Process, we ask the Stakeholders for final feedback. If the feedback is positive, we send the task to the Owner for the final approval.

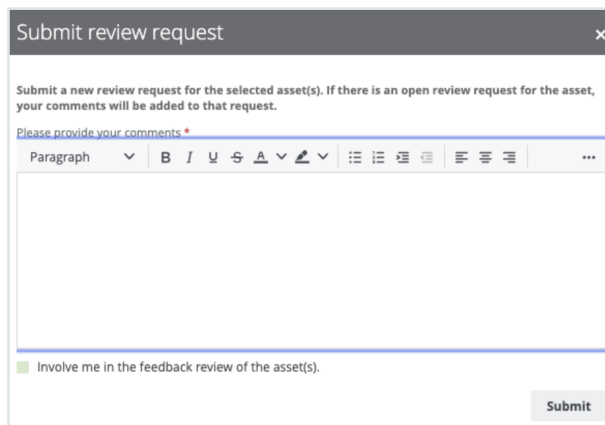
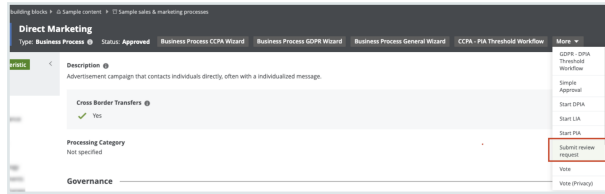


Maintain Process Register over time with review requests

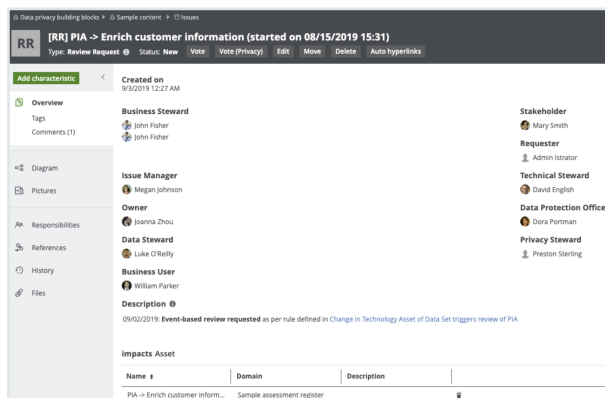
Whereas the successful result of the asset onboarding process is a new asset with the status Approved, asset change management is the standardized procedure for making changes to such approved assets.

You may have many reasons to review an approved asset. Collibra Data Privacy groups such reasons into three categories and offers three corresponding means to trigger a review request:

- **Manual:** A trigger that is manually actioned by a user if, for example, the user wants to request a review of a Business Process asset considered to be incomplete or inaccurate. Any user can manually request a review of an approved asset.



- **Time-based:** A trigger that is automatically actioned at a specified frequency. This is useful for assessment assets for which you might be required to review periodically to comply with a regulation.



- **Event-based:** A trigger that is automatically actioned by the fact of changes made to specified characteristics of the related asset.

All of the review requests are available in the Data Helpdesk.

Name	ID	Description
[R] Customer Information - 2019/09/02 22:03	09/02/2019	Manual review requested by Admin labator, refer to comments below.
	09/02/2019	Request accepted by Admin
	09/02/2019	Review request implemented
[R] Direct Marketing - 2019/08/08 15:52	08/08/2019	Manual review requested by Admin labator, refer to comments below.
[R] Enrich customer information - 2019/09/02	09/02/2019	Manual review requested by Admin labator, refer to comments below.
[R] PIA - Enrich customer information start...	09/02/2019	Event-based review requested as per rule defined in Change in Technology Asset of Data Set triggers review of PIA.
[R] Travel & Expenses - 2019/09/10 08:51	09/10/2019	Manual review requested by Admin labator, refer to comments below.
	09/10/2019	Request accepted by john.fisher

Perform Assessments

Conduct PIA and DPIA

If a business process is likely to introduce a level of risk to the rights and freedoms of natural persons, the Business Steward or the Data Protection Officer must perform the following:

- Privacy Impact Assessment (PIA), if complying with CCPA
- Data Privacy Impact Assessment (DPIA), if complying with GDPR

To determine whether or not you need to perform such an assessment for a Business Process asset, you must run a Threshold workflow.

The potential for business processes to expose the rights and freedoms of natural persons to risk is significant. Privacy Impact Assessments (PIA) and Data Privacy Impact Assessments (DPIA) assess the risks to the rights and freedoms of data subjects, born of a specific business process.

After onboarding a Business Process asset, the relevant Threshold workflow helps you determine whether or not a PIA or DPIA is needed. If it is determined that an assessment is necessary, the Owner or the Business Steward for the Business Process asset must complete the relevant workflow:

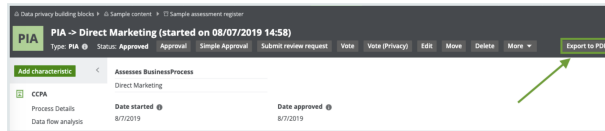
- PIA, if complying with CCPA
- DPIA, if complying with GDPR

Print assessment results

Assessments are a way for an organization to demonstrate compliance. You can export and print the PIA results in a unified way. You can also download a PIA asset page as a printable PDF, regardless of the status of the PIA asset.

Steps

1. Go to the relevant PIA asset page.



2. Click **Export to PDF**.
 - » The PDF is downloaded to your computer.

Data privacy building blocks > Sample content > Sample assessment register Print date: 2019-11-04

PIA PIA -> PIA -> Direct Marketing (started on 08/07/2019 14:58)

Status: Approved Date started: 8/7/2019 | Last modified: 11/4/2019 | Modified by: Istrator Admin

Final decision: 1. Processing allowed

Business Process assessed by PIA
[Direct Marketing](#)

General Description
 In the Direct Marketing Process, we send target marketing materials to our customers and prospects. We profile our customers to categorize our customers in 4 categories, to which we can send marketing materials that are customized to the category the customers belong to.

Details

Personal information usage
 We are processing Personal Information for Direct Marketing Purposes. We are not selling Personal Information. We are in full control of the PI

Personal information source
 Directly from the custom
 From a third par

Purpose of personal information usage defined
 Yes
 Justification not provided

Data flow analysis

Personal information categories <input checked="" type="checkbox"/> Yes Justification not provided	Third parties <input checked="" type="checkbox"/> Yes Justification not provided
Sharing of personal information <input checked="" type="checkbox"/> Yes Justification not provided	Data sharing agreements <input checked="" type="checkbox"/> No We still need to update the Data Sharing Agreements

Controls analysis

Minimization <input checked="" type="checkbox"/> Yes We have minimized the PI to what is absolutely	Quality <input checked="" type="checkbox"/> No No Data Quality process implemented yet. Not the
--	--

1 of 3

Install Collibra Protect

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

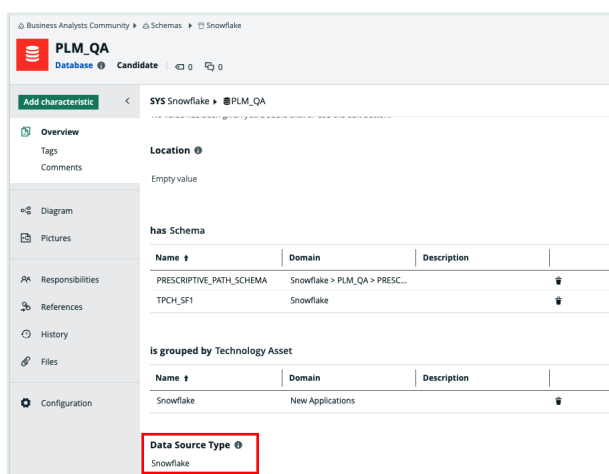
Prerequisites

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.

Name	Description	Required license	Members
Catalog Author		Standard	Admin Istrator
Data Dictionary		Read-only	Everyone
DataSteward	Allows usage of Data Steward...	Read-only	Everyone
DataSteward Author		Standard	
Edge integration engineer	Allows managing connections...	Standard	
Edge manager	Allows creating and deleting E...	Standard	
Edge site	Allows connection from Edge ...	Standard	Edge privacy-risk-qa-25-07
Edge site administrator	Allows downloading Edge site...	Standard	
Glossary	Allows usage of Business Glo...	Read-only	
Helpdesk		Read-only	Everyone
Insights		Standard	Everyone
Policy Manager		Read-only	Everyone
Protect Admin	In this role, you have the sam...	Standard	Admin Istrator
Protect Author	In this role, you can create rul...	Standard	Admin Istrator, Author User
Protect Manager	This is a role for our system u...	Read-only	Edge privacy-risk-qa-25-07, Policy Lifecycle Management API User
Protect Reader	In this role, you can view Colli...	Read-only	Reader User
ReferenceData	Allows usage of Reference Da...	Read-only	Everyone
Sysadmin	Allows for configuring and m...	Standard	Admin Istrator

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.
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. Collibra 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.

Groups

Adding Groups
 To add a group, you have to use the [Collibra Protect Group API](#). Currently, only Snowflake data sources are supported.

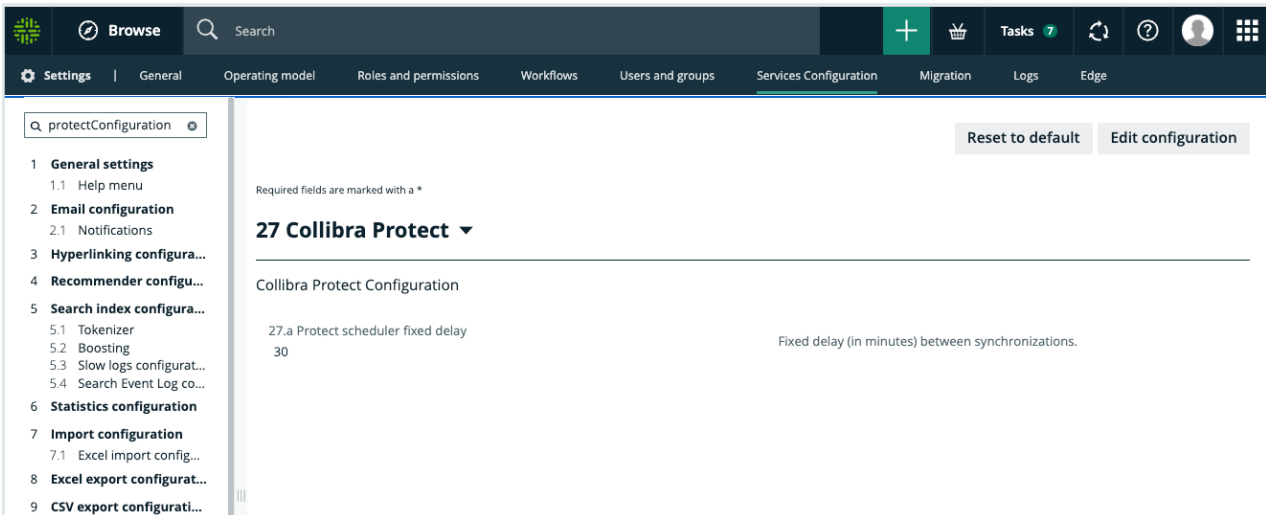
Group Name	System Reference	Created By	Created Date
------------	------------------	------------	--------------

The screenshot shows the Snowflake query editor interface. The query executed is `SHOW roles;`. The results are displayed in a table with the following columns: Row, created_on, name, is_default, is_current, is_inherited, assigned_to_users, granted_to_roles, granted_roles, owner, and comment. There are 12 rows of data.

Row	created_on	name	is_default	is_current	is_inherited	assigned_to_users	granted_to_roles	granted_roles	owner	comment
1	2019-09-17 16:47:2...	ACCOUNTADMIN	N	Y	N	35	0	3		Account administrat...
2	2022-06-27 01:10:4...	ANTONIO	N	N	N	1	1	0	SBL_TEMPLATE_SN...	
3	2022-06-02 07:07:...	BILLING	N	N	N	1	0	0	ACCOUNTADMIN	
4	2020-04-15 05:12:2...	CERTIFICATION	N	N	Y	1	1	0	ACCOUNTADMIN	
5	2022-06-02 07:05:...	CUSTOMER_SERVICE	N	N	N	1	0	0	ACCOUNTADMIN	
6	2020-05-06 00:56:...	DATALIFT_ROLE	N	N	Y	1	2	0	ACCOUNTADMIN	
7	2022-06-27 01:12:4...	Direct Marketing	N	N	N	1	0	1	SBL_TEMPLATE_SN...	
8	2022-01-27 13:27:5...	FIVETRAN_ROLE	N	N	Y	3	1	0	SECURITYADMIN	
9	2021-09-27 05:36:1...	GLOBAL_PS	N	N	N	1	0	0	ACCOUNTADMIN	
10	2021-10-22 04:38:4...	HR	N	N	Y	10	1	0	ACCOUNTADMIN	
11	2022-03-03 00:00:...	LAW	N	N	N	0	0	0	ACCOUNTADMIN	
12	2021-09-29 04:59:...	MARKETING	N	N	Y	11	1	0	ACCOUNTADMIN	

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.

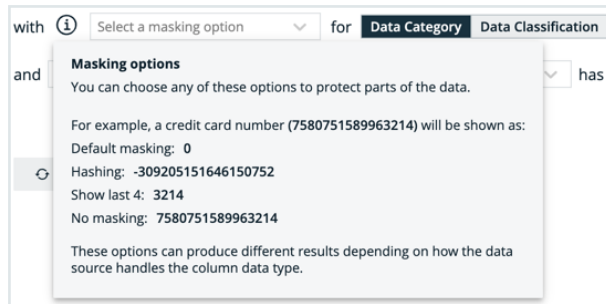
2. Column-based protection

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

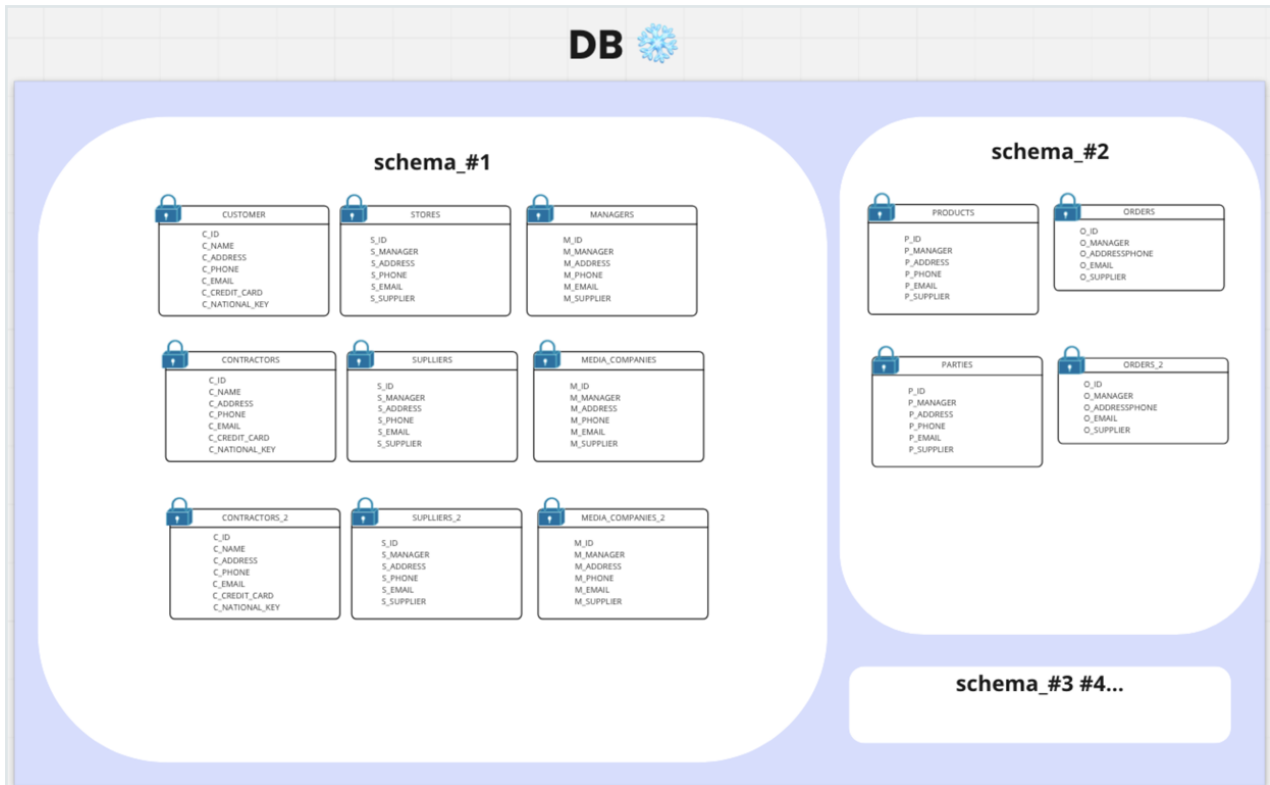
Another way to protect your data is to filter rows of a specific table. If you do not want to expose all of the existing items in a table because one of the columns is part of a certain data classification, you can easily leverage the Collibra operating model to do so.

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**.

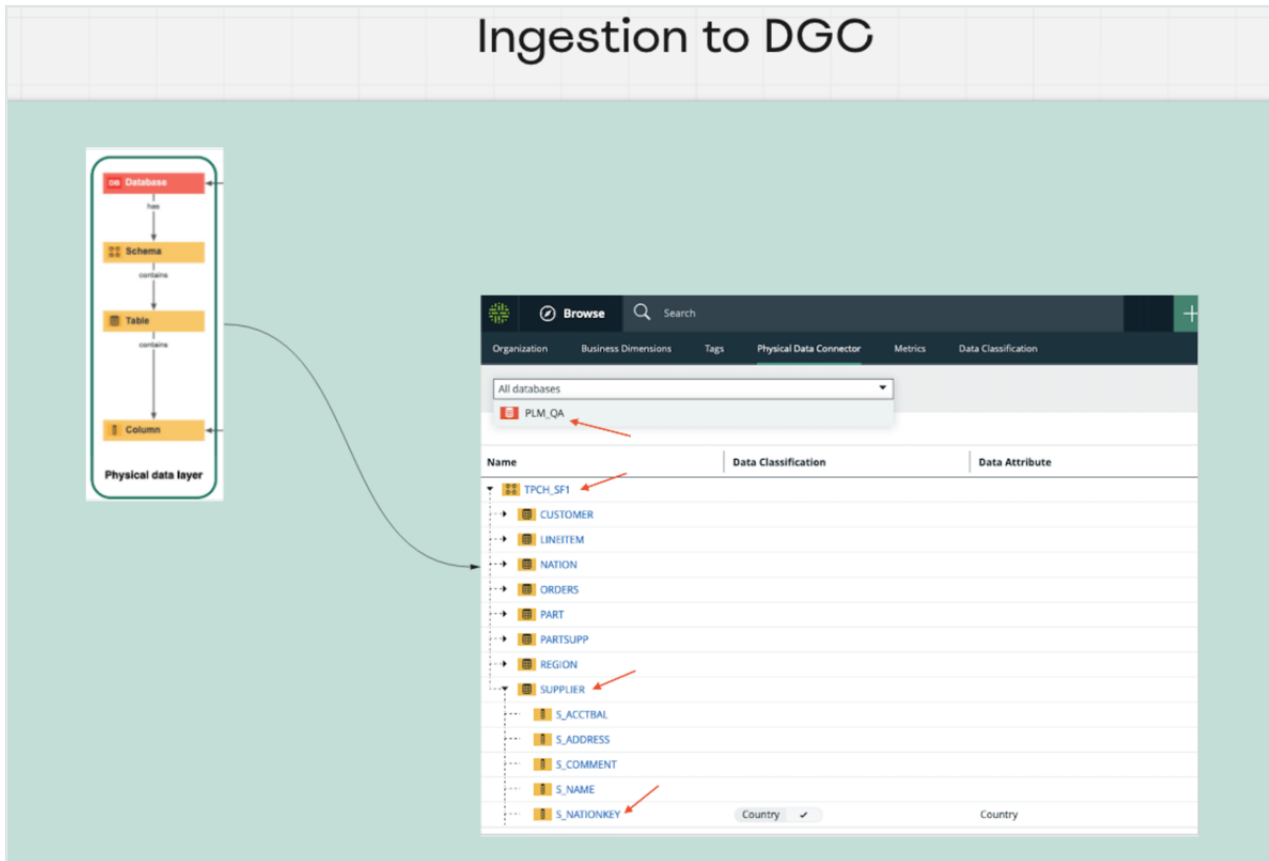
Technical background

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 and data access rules

[Data protection standards](#) and [data access rules](#) govern your data with ease and clarity.

Data protection standards

Data protection standards create a layer of protection for similar types of data by masking them wherever they are.

For example, if columns with the first and last names are a part of the PII data category, regardless of the tables, schemas, and databases to which they belong, you can create a data protection standard that targets all of these columns, by choosing the PII data category and masking it.

Data access rules

After establishing this primary layer (blanket) of protection to your most sensitive data, you can use data access rules to manage access and enhance protection for specific usages.

For example, you can create a data access 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 data protection standard that you created earlier.

Tip When creating [data protection standards](#) or [data access rules](#) for assets, consider how those assets are grouped. Suppose that you have a Business Process asset, BP, which contains the following Data Set assets: DS1, DS2, and DS3. Instead of creating a [data protection standard](#) or [data access rule](#) for each of the three Data Set assets (DS1, DS2, and DS3), consider creating a standard or rule that targets the Business Process asset (BP), to save time.

Frequently asked questions

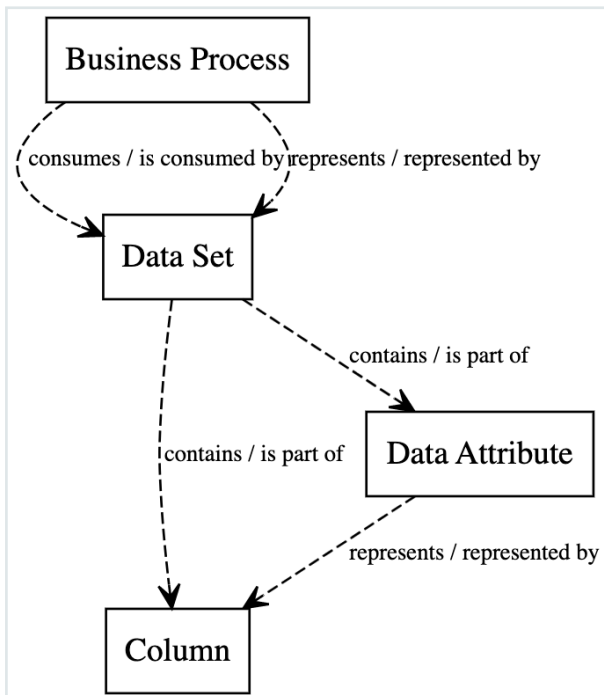
- What if I want to grant access to a group without having the PII masked?
 - » When creating a rule for an asset that contains data masked by a standard, choose to override it by unmasking it or changing its masking type.
- What if I want to grant access to a group, but the protection from the standard is not enough because there might also be other sensitive data within a 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

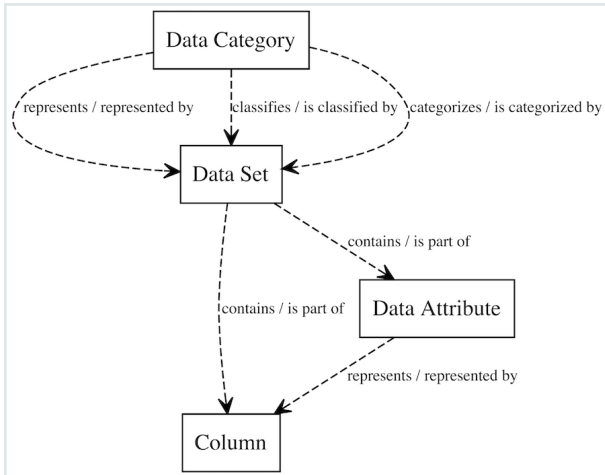
The assets that you use to create data protection standards and data access rules are related to the physical data layer, such as tables and columns, through a set of relations and intermediate assets. Collibra Protect uses these relationships and intermediate assets to search the knowledge graph to find the physical data layer assets that it needs to protect. The traversal of the knowledge graph follows a set of prescriptive paths. Each asset type has a set of prescriptive paths for traversing to the Column asset, as illustrated in the following sections.

Note Depending on your permission, you can customize the prescriptive paths. For more information, go to [Customization of prescriptive paths](#).

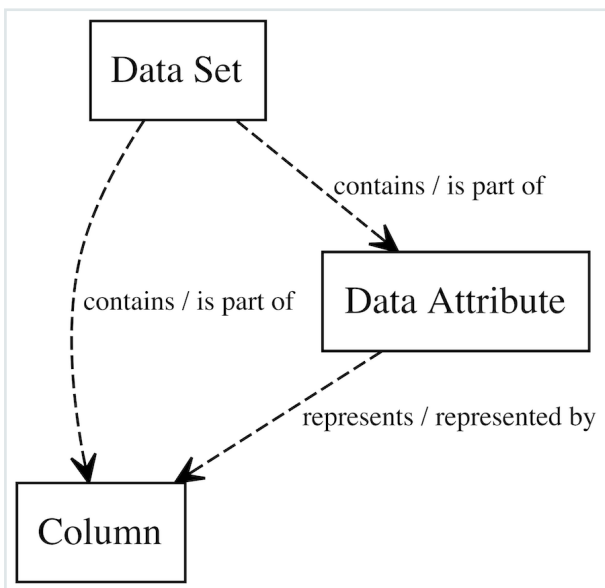
Business Process



Data Category



Data Set



Customization of prescriptive paths

Collibra Protect supports the following asset types:

- Packaged asset types: Business Process, Data Category, and Data Set
- Custom asset types: These are the packaged asset types that you have modified or the asset types that you have created. If you modify the attributes and relations of a packaged asset type, then the packaged asset type becomes a custom asset type.

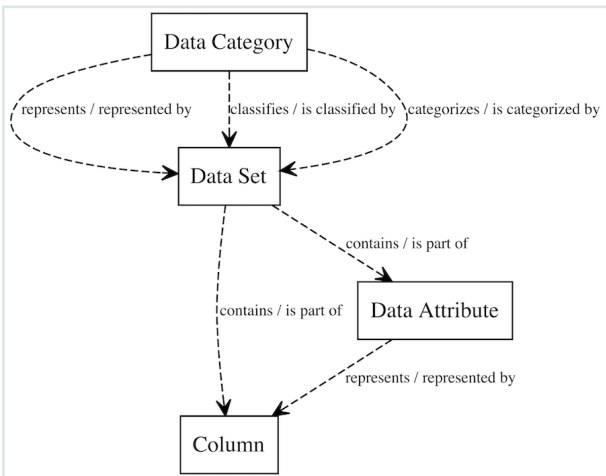
If you have the **Protect > Administration** global permission, you can customize the **prescriptive paths** for the asset types through APIs. The customization may include creating, modifying, or deleting the prescriptive paths: for example, adding or modifying the prescriptive paths for packaged and custom asset types, defining how the asset types relate to columns, removing any obsolete prescriptive paths.

The customized prescriptive paths are applied to data protection standards and data access rules.

Note You cannot remove a customized prescriptive path if an asset type linked to the prescriptive path is used in a data protection standard or a data access rule.

Collibra Protect supports a maximum of 10 asset types. Each asset type can have a maximum of 6 relations and a maximum depth of 3. However, when customizing the prescriptive path for an asset type, we recommend that you provide only one relation for the asset type. Prescriptive paths must always end in a Column asset type (that is, 00000000-0000-0000-0000-000000031008).

The following image is an example of a prescriptive path that has 6 relations and a depth of 3.



If you want to restore the default asset types defined by Collibra, a PATCH operation must be performed on each asset type. The list of asset types and their specifications are as follows.

If Data Privacy is not installed

Data Set (00000000-0000-0000-0001-000400000001)

```

    {
      "description": "Prescriptive path from Data Set to Column",
      "relations": [
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000000031008"
          }
        },
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000000031005",
            "relation": {
              "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
              "relationTypeDirection": "SOURCE",
              "assetType": {
                "assetTypeId": "00000000-0000-0000-0000-0000000031008"
              }
            }
          }
        }
      ],
      "assetTypeId": "00000000-0000-0000-0001-000400000001"
    }

```

Data Category (00000000-0000-0000-0000-0000000031109)

```

    {
      "description": "Prescriptive path from Data Category to Column",
      "relations": [
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
          "relationTypeDirection": "SOURCE",

```

```

    "assetType": {
      "assetTypeId": "00000000-0000-0000-0001-
000400000001",
      "relation": {
        "relationTypeId": "00000000-0000-0000-0000-
000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-
000000031008"
        }
      }
    }
  },
  {
    "relationTypeId": "00000000-0000-0000-0000-
000000007038",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0001-
000400000001",
      "relation": {
        "relationTypeId": "00000000-0000-0000-0000-
000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-
000000031005",
          "relation": {
            "relationTypeId": "00000000-0000-0000-0000-
000000007094",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-
000000031008"
            }
          }
        }
      }
    }
  }
},
{
  "relationTypeId": "00000000-0000-0000-0000-
000000007007",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0001-
000400000001",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-

```

```

000000007062",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0000-
000000031008"
    }
  }
},
{
  "relationTypeId": "00000000-0000-0000-0000-
000000007007",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0001-
000400000001",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-
000000031005",
        "relation": {
          "relationTypeId": "00000000-0000-0000-0000-
000000007094",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-
000000031008"
          }
        }
      }
    }
  }
},
{
  "assetTypeId": "00000000-0000-0000-0000-000000031109"
}

```

Business Process (00000000-0000-0000-0000-000000031103)

```

{
  "description": "Prescriptive path from Data Set to Column",
  "relations": [
    {
      "relationTypeId": "00000000-0000-0000-0000-

```

```

000000007062",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0000-000000031008"
  }
},
{
  "relationTypeId": "00000000-0000-0000-0000-
000000007062",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0000-
000000031005",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-
000000007094",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-
000000031008"
      }
    }
  }
}
],
"assetTypeId": "00000000-0000-0000-0001-000400000001"
}

```

If Data Privacy is installed

Data Set (00000000-0000-0000-0001-000400000001)

```

{
  "description": "Prescriptive path from Data Set to Column",
  "relations": [
    {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-000000031008"
      }
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",

```

```

        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
                }
            }
        }
    },
    ],
    "assetTypeId": "00000000-0000-0000-0001-000400000001"
}

```

Data Category (00000000-0000-0000-0000-000000031109)

```

    {
        "description": "Prescriptive path from Data Category to Column",
        "relations": [
            {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0001-000400000001",
                    "relation": {
                        "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
                        "relationTypeDirection": "SOURCE",
                        "assetType": {
                            "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
                        }
                    }
                }
            }
        ],
        {
            "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
            "relationTypeDirection": "SOURCE",
            "assetType": {

```

```

        "assetTypeId": "00000000-0000-0000-0001-
000400000001",
        "relation": {
            "relationTypeId": "00000000-0000-0000-0000-
000000007062",
            "relationTypeDirection": "SOURCE",
            "assetType": {
                "assetTypeId": "00000000-0000-0000-0000-
000000031005",
                "relation": {
                    "relationTypeId": "00000000-0000-0000-0000-
000000007094",
                    "relationTypeDirection": "SOURCE",
                    "assetType": {
                        "assetTypeId": "00000000-0000-0000-0000-
000000031008"
                    }
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007007",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-
000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007007",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",

```

```

        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
          "relation": {
            "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
            }
          }
        }
      },
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-0000-000000007315",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0001-0004000000001",
        "relation": {
          "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
          }
        }
      }
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-0000-000000007315",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0001-0004000000001",
        "relation": {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
            "relation": {

```

```

    "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007094",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
    }
  }
}
],
"assetTypeId": "00000000-0000-0000-0000-000000031109"
}

```

Business Process (00000000-0000-0000-0000-000000031103)

```

  {
    "description": "Prescriptive path from Business Process to Column",
    "relations": [
      {
        "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0001-000400000001",
          "relation": {
            "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-000000031008"
            }
          }
        }
      },
      {
        "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0001-000400000001",
          "relation": {

```



```

        "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-0000000031005",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-0000-0000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-0004000000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-0000-0000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-0004000000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-0000-0000000031008"
                }
            }
        }
    }
]

```

```
000000031005",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-0000-00000007094",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
      }
    }
  },
  "assetTypeId": "00000000-0000-0000-0000-000000031103"
}
```

Open Protect

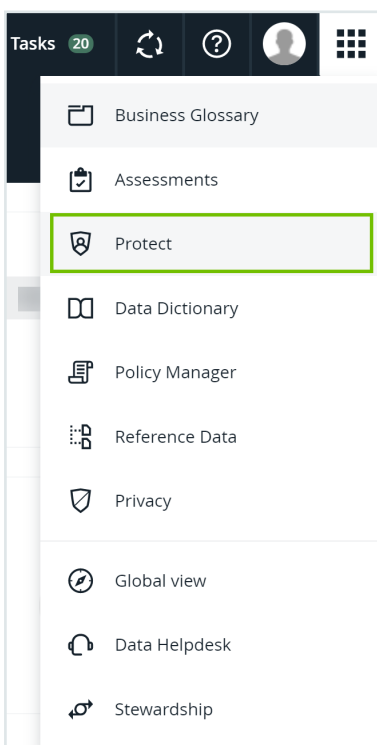
This topic describes how to open Protect, including how you can use the [tabs](#) on the Protect landing page.

Requirements and permissions

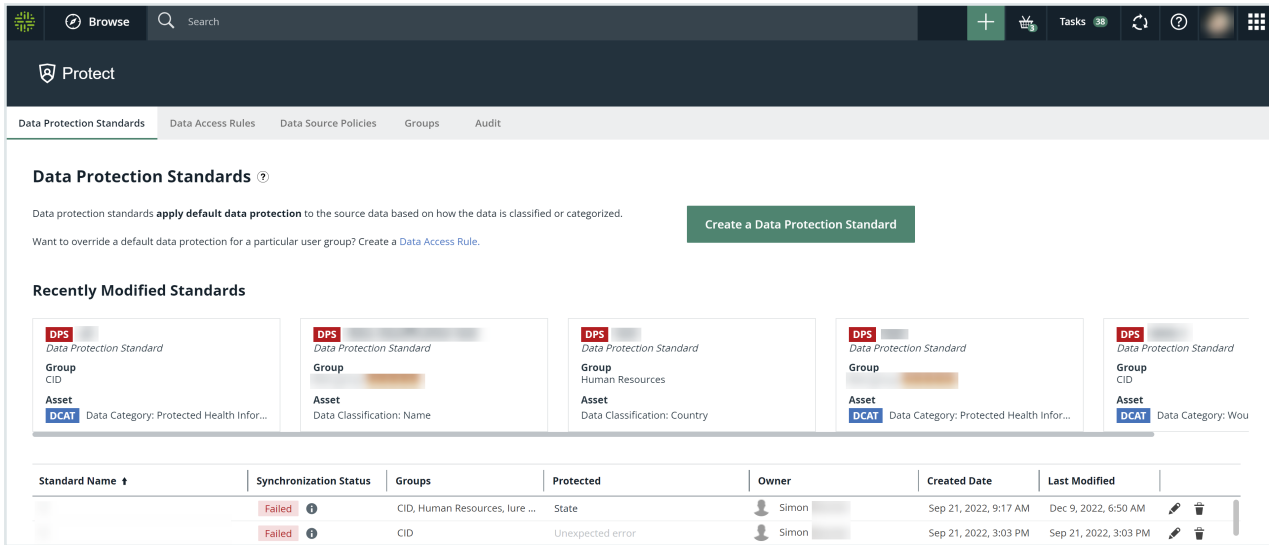
You have a global role that has the [Protect global permission](#).

Steps

On the main menu, click , and then click **Protect**.



» The Protect landing page is shown.



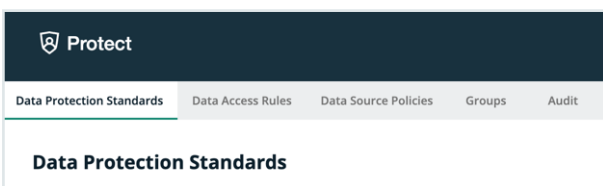
Tabs on the Protect landing page

On the Protect landing page, depending on your role, the following tabs are shown.

Tab	Description
Data Protection Standards	View or create standards to define data source access to data types based on data categories, data attributes, or data classifications.
Data Access Rules	View or create rules to grant specific groups different access to the same data in data sets, business processes, or data categories. <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note Data access rules take precedence over data protection standards.</p> </div>
Data Source Policies	View the policies that are active in the data source tables. These include the policies that were manually created in the data source and the policies that were generated in the data source due to data protection standards and data access rules in Protect. <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Tip Contact Collibra support to import policies from the data source using the Collibra Protect Data Source Policies API.</p> </div>
Groups	View or create groups for data protection standards and data access rules. You can create groups using the Collibra Protect Group API . <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note You must create at least one group before creating a standard or a rule.</p> </div>
Audit	Generate an audit log of the 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.
Pending	This standard has been created or edited, and is pending synchronization.

Synchronization Status	Description
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; background-color: #f9f9f9; 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 * 

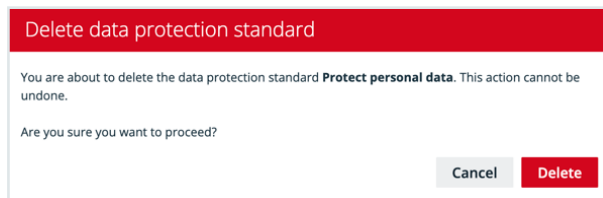
Summary
For the Group Human Resources and Marketing
protect [GDPR data related to criminal convictions and offences](#)
with Default masking

Delete a standard

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

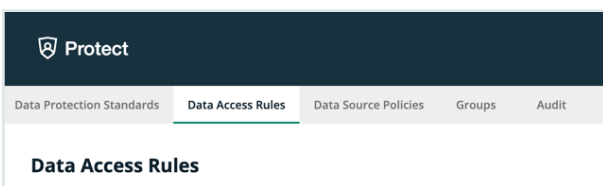
Steps

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 of protection to your most sensitive data using data protection standards, you can create data access rules to manage access to the data sources and enhance protection for specific usages.

Steps

1. In Collibra Protect, click the **Data Access Rules** tab.
2. Click **Create a Data Access Rule**.
 - » The **Create a Data Access Rule** dialog box appears.
3. Enter the required information.

Field descriptions

Field	Description
Rule Name	The name to identify the data access rule.
Description (optional)	A description for the rule.
Group	Group for the rule.

Field	Description
Asset	<p>The data asset that the rule is protecting. This field contains Business Process, Data Category, and Data Set assets, as well as assets of custom asset types.</p> <div data-bbox="1118 618 1422 1144"><p>Tip</p><ul style="list-style-type: none">○ For more information, go to Technical background and Prescriptive paths.○ You can add more groups using the plus icon.</div>

Field	Description
Optional: Select a masking option	<p>The type of masking that you want to apply to a data category or data classification. The following options are available:</p> <ul style="list-style-type: none">○ Default masking○ Hashing○ Show last○ No masking <p>In the Select a data category/data classification field, select the data category or data classification for the masking option that you selected.</p> <div data-bbox="1118 1037 1417 1335"><p>Tip You can add more data categories and data classifications for masking using the plus icon.</p></div>

Field	Description
Optional: Select an action	<p>The type of row-filtering action that you want to apply to a data classification with a specific code set and code value. The following actions are available:</p> <ul style="list-style-type: none">○ Show○ Hide <ol style="list-style-type: none">a. In the Select a data classification field, select the data classification that you want to show or hide.b. In the Select a code set field, select the code set for the data classification.c. In the Select a code value field, select the code value for the code set. <div data-bbox="1118 1279 1418 1509" style="border-left: 2px solid #008000; padding-left: 10px; margin-top: 10px;"><p>Tip You can add more data classifications for row-filtering using the plus icon.</p></div>

Tip

- The grant access checkbox is selected by default. The selected checkbox indicates that you are granting access to the tables and the columns in the database that are linked to the selected assets to the groups that you selected in the rule. If you do not want to grant this level of access to the selected groups, clear the checkbox.
- The **Summary** section shows a summary of the rule.

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 and apply default masking to Genetic data

Set rule for

group * Marketing + -

asset * Geographic Information + -

Grant access to the data linked to these assets.
By checking this box, additional access is given to the data tables or columns linked with the selected assets. If this box is unchecked, no access is given to the selected assets, but they can still be protected. **Note: once the rule granting access is saved and synchronized, access to these assets 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 Select a data classification 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

4. To preview the rule, click **Generate Preview**.

Tip The preview shows only the first 1,000 affected columns. The drop-down list box below the **Generate Preview** button is used to switch between the assets that you selected in the rule. Each asset has its own preview table.

5. Click **Save Rule**.

- » A message stating that the data access rule is sent to source appears, and the rule is shown in the table containing rules.

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 Data Classification + -


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.

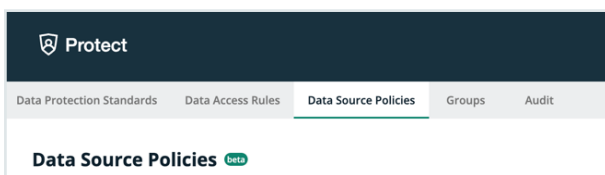
Delete data access rule

You are about to delete the data access rule **Rule 1**. This action cannot be undone.

Are you sure you want to proceed?

Data source policies

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

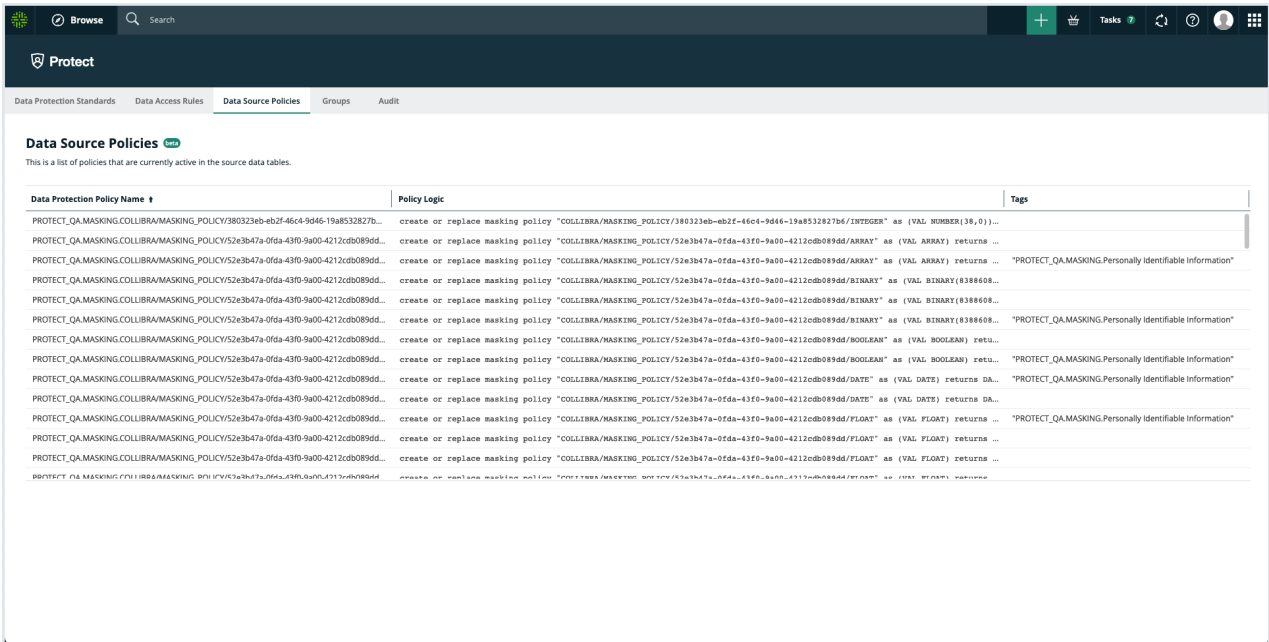


The data protection policy table displays a list of policies that are currently active in the source data tables. This includes policies that were created via Collibra Protect as well as policies that were created in the data source manually.

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 Collibra 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.
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.



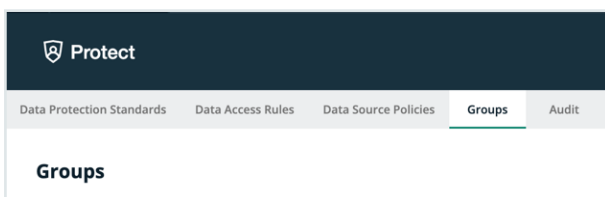
Types of policies on Snowflake

There are three types of policies on Snowflake: Column-based policies, row access policies, and tag-based policies. Each type can be created in Collibra 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 sub-header 'Groups' and 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.' Below this is a table listing active groups.

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 Collibra Protect group
System Reference	
Created By	User who created the Collibra Protect group
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

An audit log contains information about the queries that were run to access the data and the data that was accessed.

Generate an audit log

You can generate an audit log of access records from the data source on the **Audit** page.

Note The time that it takes for the actions performed in a data source to appear in an audit log in Collibra Protect varies from several minutes to hours, depending on the data source.

Steps

1. In Collibra Protect, click the **Audit** tab.
2. Depending on your data source, click **BigQuery** or **Snowflake**.
3. Click one of the following buttons: **Today**, **Yesterday**, **A week ago**, **30 days ago**.

Tip The start date corresponding to the button that you clicked is shown in the **Start Date** field. Alternatively, you can enter or select a date in the **Start Date** field, and then click **Generate Log**.

» The audit log is generated.

Important

- Generating an audit log may take up to a minute. After clicking **Generate Log**, do not navigate away from the **Audit** page because doing so cancels the audit log generation.



- The audit log contains the first 1,000 records from the selected start date. If you want to view the remaining records, contact your data source administrator.

The screenshot shows the 'Audit' interface with a 'Generate Log' button and a table of audit records. The table has five columns: Query ID, Query Start Time, Source User Name, Direct Objects Accessed, and Base Objects Accessed. The records show queries executed by user 'MELIK' on Sep 29, 2022, at 2:00 AM, accessing various database objects.

Query ID	Query Start Time	Source User Name	Direct Objects Accessed	Base Objects Accessed
01a74800-0501-ec9a-0001-0003066b19e	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.MAIN_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ec9a-0001-0003066b1a2	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.MAIN_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ea46-0001-0003066f9d42	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.DEPENDS_ON_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ec9a-0001-0003066b1a6	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.NODES_DEPENDS_ON_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE

Audit log data

The following table describes the columns that are shown in an audit log.

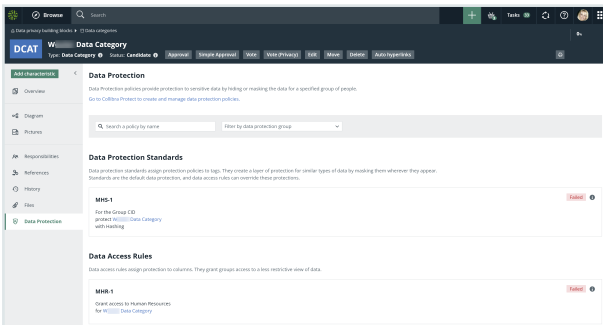
Column	Description
Query ID	The ID of the query in the source database.
Query Start Time	The date and time of the query in the source database.
Source User Name	The name of the user in the source database who ran the query to access the data.
Direct Object Accessed	The database object (a table or a view) that was used to access the data.
Base Object Accessed	The database object that was accessed.

Data protection in the asset pages

The asset pages for the following asset types contain the **Data Protection** tab to allow you to view, filter, create, and manage data protection standards and data access rules:

- Business Process
- Data Category
- Data Set
- Custom asset types such as Column, Database, Schema, and Table, derived from the aforementioned asset types via prescriptive paths

Note Data protection standards support only Data Category assets and data classifications.



View or filter standards and rules

Requirements and permissions

You have the **Protect Reader** global role.

Steps

On the asset page (for the one of the **aforementioned** asset types), click the **Data Protection** tab.

» Data protection standards and data access rules that are linked to the asset are shown.

Tip

- To filter the standards and rules by name, in the **Search a policy by name** field, enter the name of the standard or rule that you want to view.
- To filter the standards and rules by group, in the **Filter by data protection group** field, select the group for which you want to view the standard or rule.

Create or manage standards and rules

Requirements and permissions

You have the **Protect Author** and **Protect Admin** global roles.

Steps

1. On the asset page (for the one of the [aforementioned](#) asset types), click the **Data Protection** tab.
2. Click the following link: **Go to Collibra Protect to create and manage data protection policies.**

Tip For information about how to create and manage data protection standards and data access rules, go to [Data protection standards](#) and [Data access rules](#).

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

This topic contains examples to describe how data protection standards and data access rules behave when different types of masking affect the same column.

Note In the 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 + - for **Data Category** **Data Classification** Personal and family details

and rows where has

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

Conflicting filters affecting the same column

This topic contains examples to describe how data protection standards and data access rules behave when conflicting filters affect 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

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

and Hide + - rows where State 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 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 Select a data category

for Data Category Data Classification

and Show + -

rows where Country Country code BE

Summary

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

Chapter 11

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

Collibra Protect periodically synchronizes with an aggregation of all data protection standards and data access rules. These standards and rules form a data source-agnostic representation containing all databases, schemas, tables, and columns, as well as their protections and accesses. The synchronization process then triggers the [Edge capabilities](#), such as Collibra Protect for Snowflake. These Edge capabilities 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.

Protect for Snowflake

Data protection standards in Collibra Protect rely on the [tag-based masking policies](#) of Snowflake. The name of the data category or data classification selected in a standard becomes a tag with the same name. The tag is applied to all the affected columns to enforce data protection. For more information, go to [Examples](#).

Examples

This topic contains examples to describe how Snowflake behaves in relation to certain data protection standards and data access rules.

Example 1



Introduction

This example describes the behavior in Snowflake when a standard is applied to a data category and a rule is applied to a data set with categorized columns in Protect.

The example considers the following:

- A standard created for the **Everyone**, **Human Resources**, **Marketing**, and **Sales** groups, to protect the columns in the **Personally Identifiable Information** data category by default masking.

The screenshot shows the configuration for a standard in the Snowflake Protect console. It is set for the groups 'Everyone', 'Human Resources', 'Marketing', and 'Sales'. The standard is applied to the 'Personally Identifiable Information' data category using 'Default masking'.

- A rule created for the **Human Resources** group and the **Employee Data** asset, without any protection applied to the columns in the **Personally Identifiable Information** data category.

The screenshot shows the configuration for a rule in the Snowflake Protect console. It is set for the 'Human Resources' group and the 'Employee Data' asset. The rule is configured with 'No masking' for the 'Personally Identifiable Information' data category. A checkbox for 'Grant access to all data tables linked to these asset columns' is checked.

Standard

When the **standard** is synchronized and active, the standard results in 14 masking policies—one policy for each **Snowflake data type**. The masking policies are created at the schema level with the following naming convention: `COLLIBRA/MASKING_POLICY/<asset ID>/<snowflake type>`.

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

All the masking policies are then associated with the **Personally Identifiable Information** tag, which is created at the schema level and assigned to those columns that need to be protected. At runtime, Snowflake fetches the right masking policy based on the **column data type**.

Row	created_on	name	database_name	schema_name	owner	comment
1	2022-09-06 03:46:10.054...	Personally Identifiable Information	PROTECT_QA	DEMO	ACCOUNTADMIN	Generated by Collibra: 28d226cc-0ab0-4d23-b912-985312fb36b1

The following image shows a masking policy for the STRING data type. The data that is shown in the policy depends on the masking type selected in the standard. In the policy, `val` indicates the value as it is stored in the table.

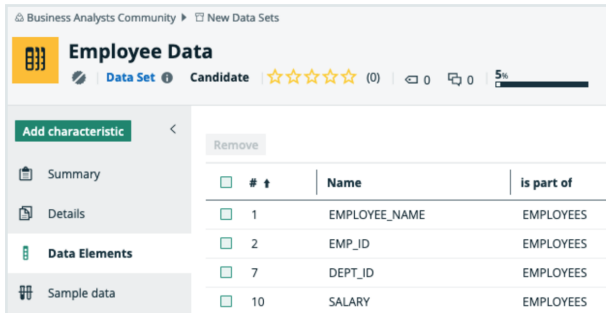
```

Details
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
    
```

Rule

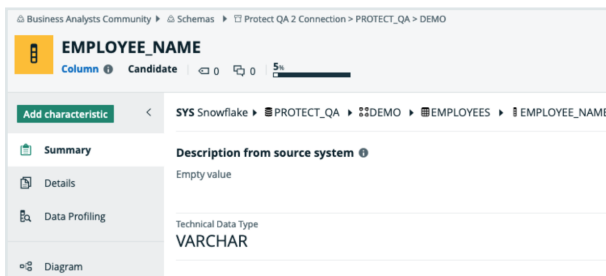
A rule results in a combination of **grant instructions**, **dynamic masking**, and **row access policies**.

Suppose that the **Employee Data** data set selected in the **rule** contains sensitive columns categorized as **Personally Identifiable Information**.



The **rule** grants access of the **Employee Data** data set to the **Human Resources** group, as indicated by the selected **Grant access...** checkbox in the rule. Then, the corresponding Snowflake role for the group can access each database, schema, and table in the data set. In addition, the column masking policy is applied to those columns that need to be protected.

Consider the **EMPLOYEE_NAME** column in the **Employee Data** data set. This column belongs to the **EMPLOYEES** table within the **DEMO** schema in the **PROTECT_QA** database.



In Snowflake, each column that is categorized as **Personally Identifiable Information** within the **Employee Data** dataset inherits the masking policy that is applied to the column in Protect. The masking policies created at the schema level use the following naming convention: **COLLIBRA/MASKING_POLICY/<asset ID>**.

Row	created_on	name	database_name	schema_name	kind	owner
18	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY7f62b08a-af5a-41ef-af64-c684-6487961	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
17	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY4667075-210f-488f-8461-46670752107	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
16	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY898806689f1688-9f08-9f08-9f08-9f08988984	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
15	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY8832796-8647-884-834-298839510e	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
14	2022-09-06 03:46:08.9...	COLLIBRAMASKING_POLICY082282c0-0607-4423-8912-9931283615981ANT	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN

The following image shows the masking policy created for the **EMPLOYEE_NAME** column.

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

```

Behavior

According to the [standard](#), the **Everyone**, **Human Resources**, **Marketing**, and **Sales** groups have masked access to the data. However, according to the [rule](#), the **Human Resources** group has unmasked access to the data. As a result, the **EMPLOYEE_NAME** column has both a policy tag and a column masking policy applied to it via the standard and the rule, respectively.

In Snowflake, if both a policy tag and a column masking policy exist for a column, the column masking policy takes precedence and the policy tag is not assigned to the column. To mitigate this behavior and ensure that the protection defined in the standard is not ignored, the column masking policy also considers the conditions defined in the standard (policy tag).

Thus, when a standard is created for the **Human Resources**, **Marketing**, and **Sales** groups to mask the **Personally Identifiable Information** column by default masking, and when a rule is created for the **Human Resources** group to not mask the same column, the result is as follows:

- The column is not masked for the **Human Resources** group.
- The column is masked for the **Marketing** and **Sales** groups via default masking.

Example 2

Introduction

This example describes the behavior in Snowflake when multiple standards affect the same column without conflict.

The example considers the following:

- A standard created for the **HR** group to protect the columns in the **Personally Identifiable Information** data category by hashing.
- A standard created for the **Marketing** group to protect the columns in the **Personal Information** data category by default masking.
- The **Personally Identifiable Information** and **Personal Information** data categories share the same column named **DOB**.

Behavior

Protect creates a tag for each standard and adds a policy to each tag. The two tags are then linked to the **DOB** column. In addition, Protect creates a masking policy that is an aggregation of the policies from the two tags. This aggregated masking policy, which is then applied to the **DOB** column, thus contains the content of both the tag policies.

```
1 CASE
2     WHEN CURRENT_ROLE() = 'HR' THEN hash(val)::NUMBER
3     WHEN CURRENT_ROLE() = 'MARKETING' THEN 0
4     ELSE val
5 END
```

When a policy exists for the **DOB** column, Snowflake considers only the column masking policy, ignoring all the tag policies associated with the column. Because the column masking policy is an aggregation of all the tag policies, the protection that is defined in the two standards is not ignored.

Thus, Protect handles multiple standards with tag policies for Snowflake by creating a column masking policy, which considers the protection defined in the standards.

Masking and data types

Snowflake provides several functions to transform the data. This topic describes how Snowflake transforms the data for a given Protect masking type.

- **Default masking:** Snowflake does not support this masking type. Protect, however, uses the default masking type to apply protection to a wide range of data types. A default masking value is applied to each column according to the data type of the column.

Default masking values for data types

Column data type	Snowflake data type	Default masking value
NUMBER	NUMBER	0
DECIMAL	NUMBER	0
NUMERIC	NUMBER	0
INT	NUMBER	0
INTEGER	NUMBER	0
BIGINT	NUMBER	0
SMALLINT	NUMBER	0
TINYINT	NUMBER	0
BYTEINT	FLOAT	0
FLOAT	FLOAT	0
FLOAT4	FLOAT	0
FLOAT8	FLOAT	0
DOUBLE	FLOAT	0

Column data type	Snowflake data type	Default masking value
DOUBLE PRECISION	FLOAT	0
REAL	FLOAT	0
VARCHAR	VARCHAR	*
CHAR	VARCHAR	*
CHARACTER	VARCHAR	*
STRING	VARCHAR	*
TEXT	VARCHAR	*
BINARY	BINARY	00
VARBINARY	BINARY	00
BOOLEAN	BOOLEAN	false
DATE	DATE	1970-01-01
DATETIME	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIME	TIME	00:00:00
TIMESTAMP	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIMESTAMP_LTZ	TIMESTAMP_LTZ	1969-12-31 16:00:00.000-0800
		<p>Note This may change depending on the time zone.</p>

Column data type	Snowflake data type	Default masking value
TIMESTAMP_NTZ	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIMESTAMP_TZ	TIMESTAMP_TZ	1969-12-31 16:00:00.000-0800 <div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> <p>Note This may change depending on the time zone.</p> </div>
VARIANT	VARIANT	0
OBJECT	OBJECT	{}
ARRAY	ARRAY	[]
GEOGRAPHY	GEOGRAPHY	{"coordinates": [0,0], "type": "Point"} (aka point(0, 0) and visualization can change based on user preferences)

- **Hashing:** Uses the following Snowflake functions:
 - *SHA2* (for strings)
 - *HASH* (for numbers)
- **Show last:** Uses the following expressions:
 - *substr(to_varchar(value), length(value) - n, n)* (for strings)
 - *mod(value, power(10,n))* (for numbers)

Tip In the expressions, *value* indicates the content and *n* indicates the number of characters to be shown.

- **No masking:** Returns the raw content.

Note

- You can apply the **Hashing** and **Show last** masking types to only the following Snowflake data types: FLOAT, NUMBER, and STRING.
- If a selected masking type cannot be applied to a certain data type—for example, when you attempt to apply the **Hashing** masking type to the DATE data type—the **Default masking** type is applied to the data type to guarantee protection.

Snowflake privileges

To perform actions in Snowflake, Collibra Protect uses an Edge connection that must be configured with a user and a role that can manage grants; create and assign masking policies, row access policies, and tags; and manage usage access on databases and schemas involved in the protection. This enforcement role requires the following Snowflake privileges.

Snowflake privilege	Description
[APPLY MASKING], [APPLY ROW ACCESS], [APPLY TAG], [MANAGE GRANTS], [IMPORTED PRIVILEGES ON DATABASE SNOWFLAKE]	Required for the role performing the actions.
[USAGE]	Required on each database and schemas where policies are applied to the role performing the actions.

Snowflake privilege	Description
<pre>[CREATE MASKING POLICY], [CREATE ROW ACCESS POLICY], [CREATE TAG]</pre>	<p>Required on each schema where policies are applied to the role performing the actions.</p>

Example

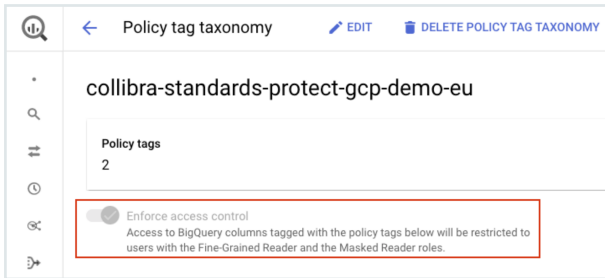
Suppose that a role named PROTECT exists in Snowflake and is responsible for managing access on all schemas within a database named DEMO. Then, the following statements can be used to enable the Snowflake PROTECT role to perform the enforcement.

```
GRANT APPLY MASKING POLICY ON ACCOUNT TO ROLE PROTECT;
GRANT APPLY ROW ACCESS POLICY ON ACCOUNT TO ROLE PROTECT;
GRANT APPLY TAG ON ACCOUNT TO ROLE PROTECT;
GRANT MANAGE GRANTS ON ACCOUNT TO ROLE PROTECT;
GRANT IMPORTED PRIVILEGES ON DATABASE SNOWFLAKE TO ROLE PROTECT;
GRANT USAGE ON DATABASE DEMO TO ROLE PROTECT;
GRANT USAGE ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE MASKING POLICY ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE ROW ACCESS POLICY ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE TAG ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
```

Protect for BigQuery

Collibra Protect uses Google's Policy tag taxonomies to create tags and assign the tags to your BigQuery columns. Policy tag taxonomies inherently apply access control. This

means that the tags applied to your BigQuery columns will be accessible only by the Protect groups configured in your data protection standards and data access rules.



BigQuery masking rules

Each Protect masking type has an equivalent counterpart in BigQuery called a [masking rule](#). As such, masking rules in BigQuery correspond to masking types in Protect.

Note The BigQuery masking rules are not the same as the Protect data access rules.

The following table contains the equivalent [BigQuery masking rule](#) for a given Protect masking type.

Protect masking type	Equivalent BigQuery masking rule
Default masking	Default masking value
Hashing	Hash (SHA256) <div style="background-color: #f0f0f0; padding: 10px; margin-top: 10px;"> <p>Note BigQuery supports the Hash (SHA256) masking rule only for certain columns depending on their data types. If Hash (SHA256) cannot be applied to a certain column due to the data type of the column, the following masking rule is applied instead: Default masking value.</p> </div>

Protect masking type	Equivalent BigQuery masking rule
Show last	Default masking value <div style="border: 1px solid #ccc; padding: 5px;"> <p>Note BigQuery does not support the Show last masking type. The Show last masking type is supported only on MadCap:variable name="Protect.Snowflake".</p> </div>
No masking	Fine-Grained Reader <div style="border: 1px solid #ccc; padding: 5px;"> <p>Note Each Protect group to which you assign standards has an equivalent counterpart in BigQuery called a GCP principal. BigQuery grants the Fine-Grained Reader role to the assigned GCP principal to allow the GCP principal to view the data to which no masking is applied in Protect.</p> </div>

BigQuery data types

The following table contains the BigQuery masking rule that Protect supports for a given BigQuery data type.

Summary

- Protect supports the BigQuery **Default masking value** rule for all types of columns.
- Protect does not support the BigQuery **Nullify** rule for any type of column.
- Protect supports the BigQuery **Hash (SHA256)** rule only for the following types of columns: BYTES, STRING.

BigQuery data type	BigQuery masking rule supported by Protect
ARRAY	Default masking value
BIGNUMERIC	Default masking value
BOOL	Default masking value
BYTES	<ul style="list-style-type: none"> • Default masking value • Hash (SHA256)

BigQuery data type	BigQuery masking rule supported by Protect
DATE	Default masking value
DATETIME	Default masking value
FLOAT64	Default masking value
GEOGRAPHY	Default masking value
INT64	Default masking value
INTERVAL	Default masking value
JSON	Default masking value
NUMERIC	Default masking value
STRING	<ul style="list-style-type: none"> • Default masking value • Hash (SHA256)
STRUCT	Default masking value
TIME	Default masking value
TIMESTAMP	Default masking value

BigQuery group mapping

The Collibra Protect group mapping for BigQuery must follow the syntax for principal identifiers. For example, the Protect group, **Sales**, maps to the BigQuery group email address, **sales@example.com**.

```
{
  "name": "Sales",
  "mappings":
  [
    {
      "provider": "GoogleBigQuery",
```

```
    "identity": "group:sales@example.com"  
  }  
]  
}
```

BigQuery permissions

To perform actions in BigQuery, Collibra Protect uses a GCP connection that must be configured with a service account having the following permissions:

- bigquery.dataPolicies.create
- bigquery.dataPolicies.delete
- bigquery.dataPolicies.get
- bigquery.dataPolicies.getIamPolicy
- bigquery.dataPolicies.list
- bigquery.dataPolicies.setIamPolicy
- bigquery.dataPolicies.update
- bigquery.datasets.get
- bigquery.jobs.create
- bigquery.rowAccessPolicies.create
- bigquery.tables.get
- bigquery.tables.list
- bigquery.tables.setCategory
- bigquery.tables.update
- datacatalog.categories.getIamPolicy
- datacatalog.categories.setIamPolicy
- datacatalog.taxonomies.create
- datacatalog.taxonomies.get
- datacatalog.taxonomies.list
- datacatalog.taxonomies.update
- logging.logEntries.list
- resourcemanager.projects.get

Collibra Protect

About Protect

Collibra Protect is a capability of the Collibra Data Intelligence Cloud to protect sensitive data and grant varying levels of access to the data to specific groups of people through policies that do not require you to code. You can enforce data protection at the source database level directly from the Collibra Protect interface, and apply advanced data protection through masking, redacting, and hashing. Protect simplifies access governance and eliminates the need for repetitive actions and approvals. By providing permission to view information to only those who need it, Protect minimizes risk and promotes a safe data culture in your organization.

You can use Protect to protect the data in the assets of the packaged asset types, such as Business Process, Data Category, and Data Set, in addition to the assets of any new or modified asset types. In addition, you can use Protect to provide differential access, for example, to give everyone access to a data set but allow certain type of access to only certain groups of people based on data categories.

Scenarios for using Protect

This topic describes how Collibra Protect helps you to:

- Use the metamodel graph to build and enforce protection policies on Business Processes, Data Categories, and Data Sets.
- Use classifications to apply a broad coverage of protection mechanisms at the data source.
- Support privacy preferences such as consent management, data subject requests such as access requests, and the right to be forgotten through row-filtering mechanisms.
- Perform an audit of applicable protection at the data source and use reporting to demonstrate compliance where data is stored and consumed.

Discover and classify personal information

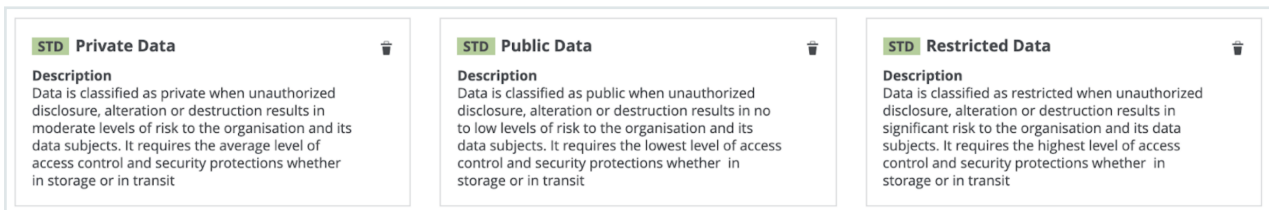
Suppose that you want to help your organization find personal information.

To achieve this, typically, your Privacy team sets up the Data Classification Policy, where they classify the data used in the organization based on the sensitivity or the business criticality of the data. This determines the required levels of security for the applications that store that data or the applications that are used for the transit of the data.

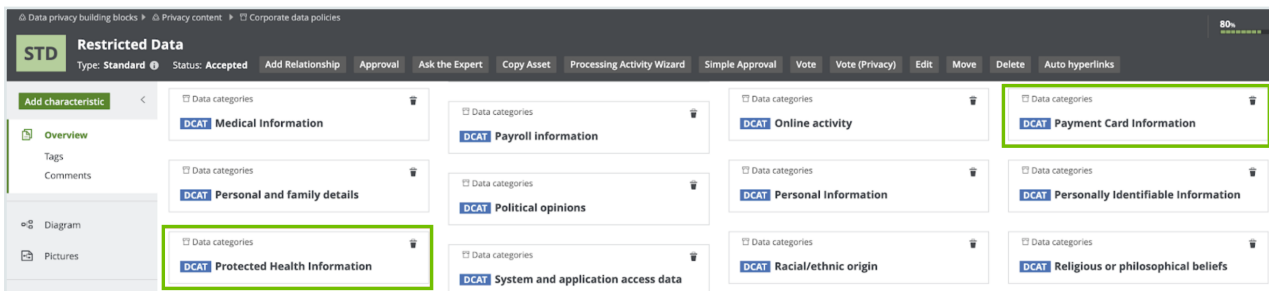
Consider the following three classifications for sensitivity:

- Public data, which is least sensitive.
- Private data, which is slightly more sensitive than the public data.
- Restricted data, which is the most sensitive data and therefore requires the highest level of access controls and security protection.

The following image shows the standard subassets of the Data Classification policy.

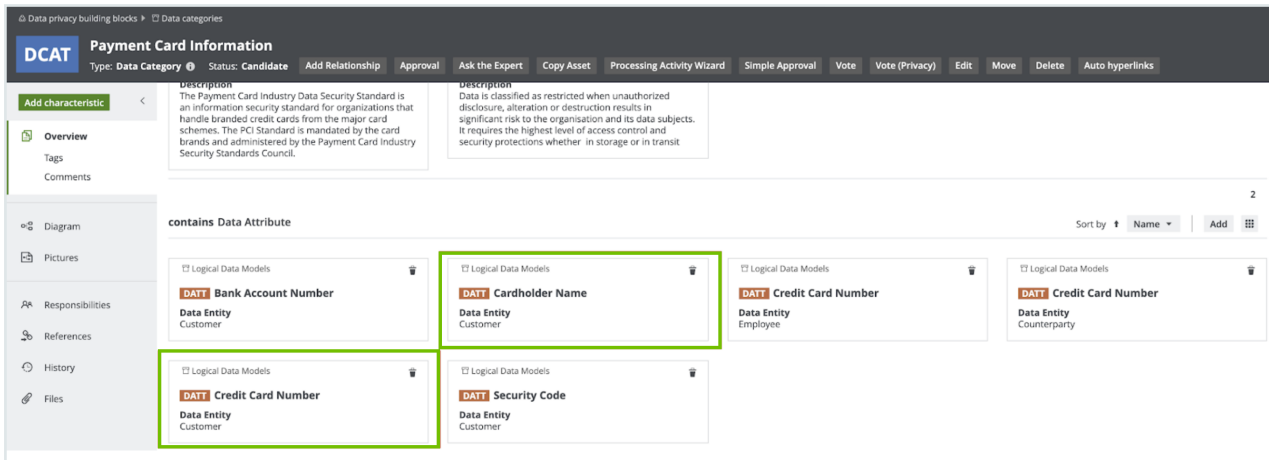


The Privacy team determines the data categories to which these subassets apply. For example, they can determine that Restricted Data applies to the following data categories: Gender, Social Security Number, Payment Card Information.

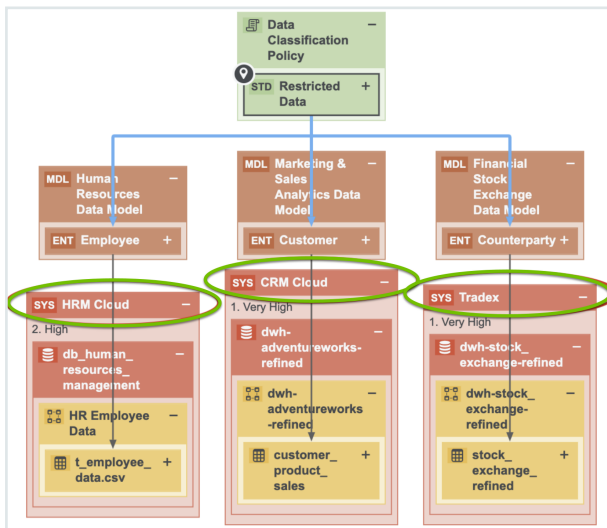


The Privacy team determines the sensitivity and the required security at the data category level as opposed to the column level. At the data category level, the Privacy team then determines what data elements belong to the identified data categories. For example, the

Payment Card Information data category groups the Cardholder Name and the Credit Card Number, among other information.



In this model, Data Attributes are grouped under the Data Category. This is how the Privacy layer is linked to the logical data model. This promotes collaboration between the Privacy team and the Governance team. In addition, this allows the automated data classification of the organisation's personal information, which makes views such as the Restricted Data Overview diagram available at the most sensitive data category, Standard Restricted Data.

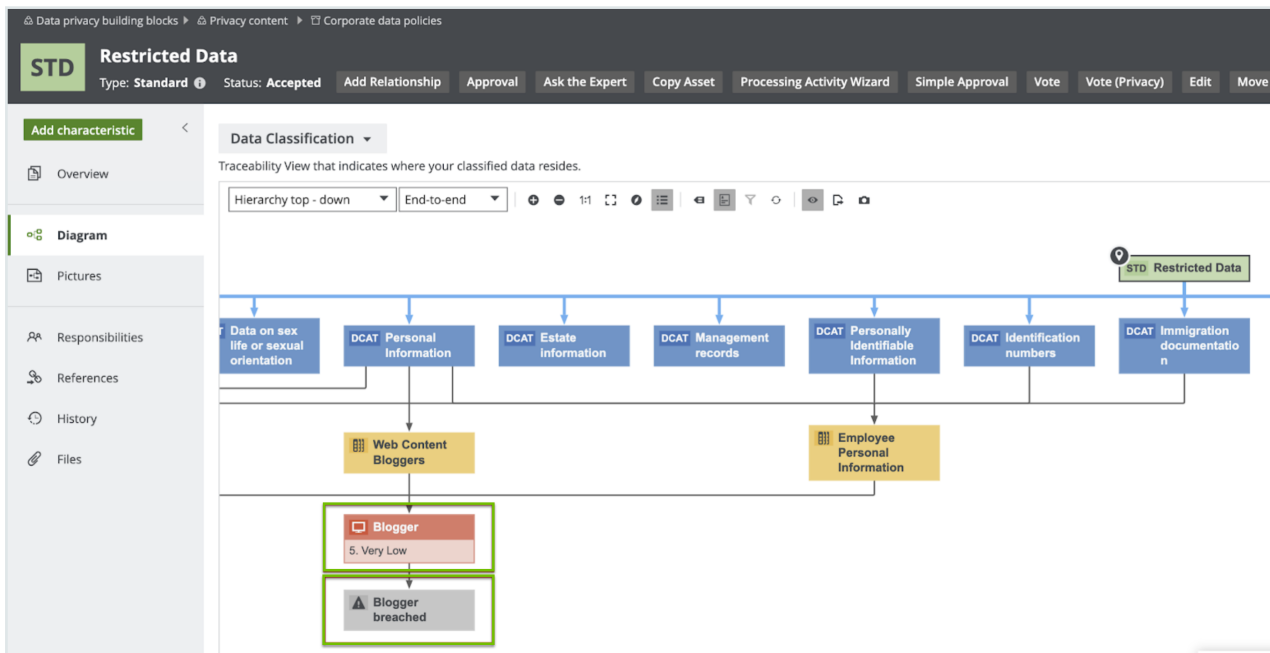


In the above image, the applications in which the restricted data resides are highlighted.

The Privacy team determines the Policies and Standards that determine which data categories are sensitive to the organization and what the required levels of protection are. The Data Governance team maps those data categories to the applications where that

data resides. The Security team determines what the security levels on those applications are. Thus, the view captured in the above image requires collaboration among teams.

Consider the traceability diagram called Data Classification under the Restricted Data standard. This standard contains the most sensitive information and thus requires the highest level of security controls; however, it resides on an application that has very low security. Because of this, the Information Security team needs to take the necessary remediation actions and improve the security levels on Blogger. As shown in the image, an investigation is already ongoing on the potential data breach on Blogger.



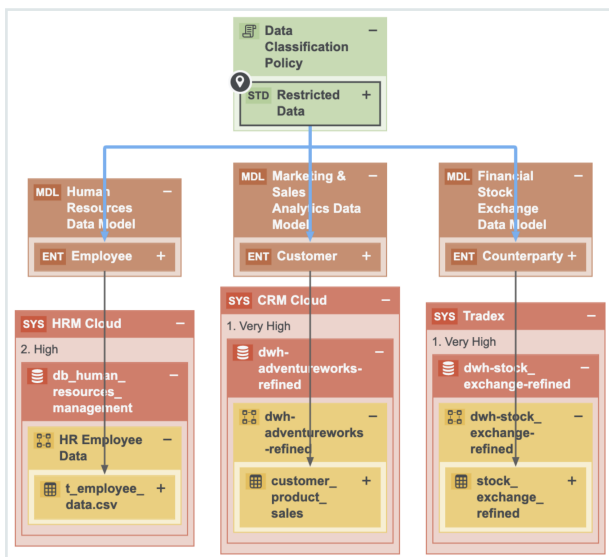
Data classification capabilities and guided stewardship

This section describes how Collibra Privacy and Risk leverages the data classification capabilities in Catalog. Thus far, we learned that the Restricted Data standard groups Data Categories, which group Data Attributes. In the example, the Payment Card Information data category contains the Credit Card Number data attribute.

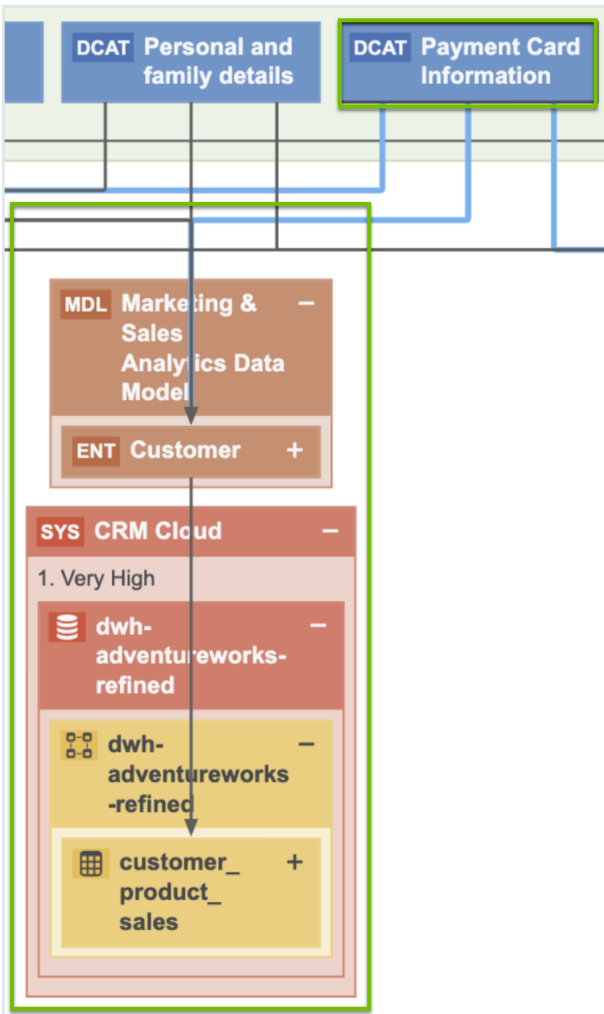
Guided stewardship is a semi-automated process of mapping columns and tables to logical data attributes. It enables content tables to be mapped to data attributes. After scanning a table and then applying guided stewardship in which the steward selects

attributes from the suggestions coming from the automated mapping, the column is mapped to the Credit Card Number. Moreover, when a column is mapped to a data attribute, the column is also mapped to a data category because of the relation between the data category and the data attribute.

The result of classifying one application with the Catalog's Data Classification is shown in the following image.

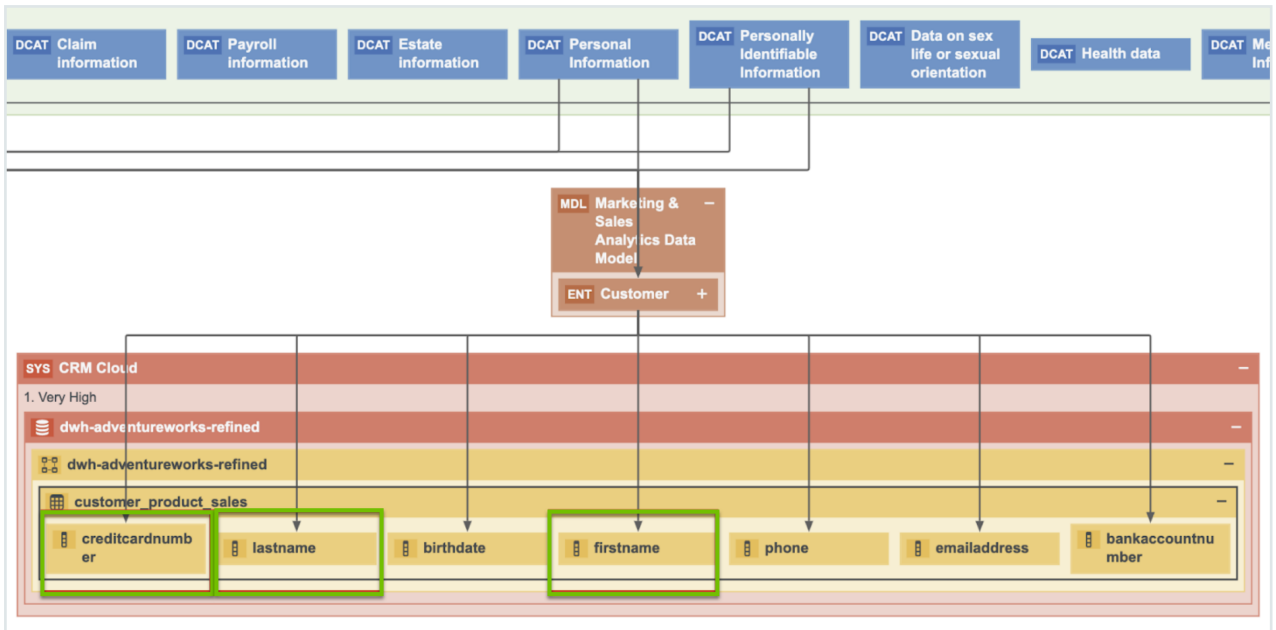


Restricted Data groups multiple data categories. The following image shows the data attributes that the Payment Card Information data category groups.



By applying guided stewardship and data classification, the data attributes are mapped to the columns. Thus, by using Catalog’s data classification capabilities, the Data Governance team can find personal information and sensitive personal information.

It is important to know the context to determine which information is considered personal information. For example, Name can be the name of a customer or an employee, in which case Name is considered personal information. Name can also be the name of another organization. This context can be provided only by a steward. Therefore, data classification and guided stewardship will help the steward mapping customer’s names to the Name column. Because the Privacy team has mapped names and family details, you can safely assume that this is Personal Information. Similarly, Credit Card Number can be the credit card number of another organization, but it is the steward who has mapped the number to the Credit Card Number data attribute belonging to the Customer data entity, and as a result, we know that the payment card information is very restricted data.



This is an example of how guided stewardship, Catalog’s data classification combined with guided stewardship and CollibraPrivacy and Risk, gives you a vertical view on where Personal Information resides.

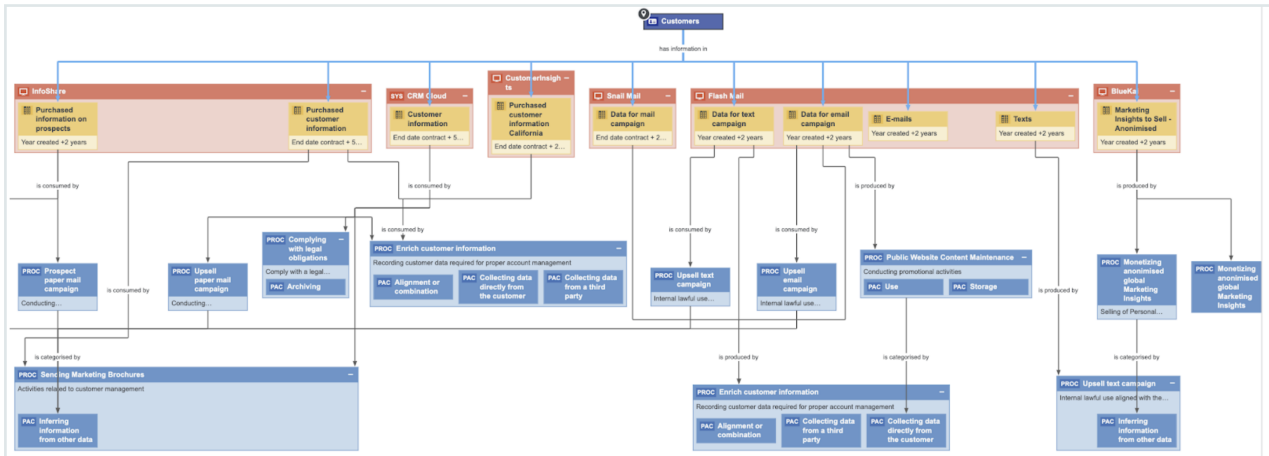
Customer requests and consent management

The previous sections described how we help customers find their Personal Information across applications. This section describes how we help customers manage data subject requests and consent. Collibra has the relevant metadata that is necessary for a partner application that fulfils the data subject requests or manages consent to operate. These applications need the metadata about where the data resides, where you store customer information, how you use the information, why you use the information, and what your legal basis is, so that they can determine for which applications you need consent and for which processes you need instance for a consent. Collibra has and governs the required metadata. In addition, through APIs, Collibra can integrate with those applications to feed them with the metadata that they need to function.

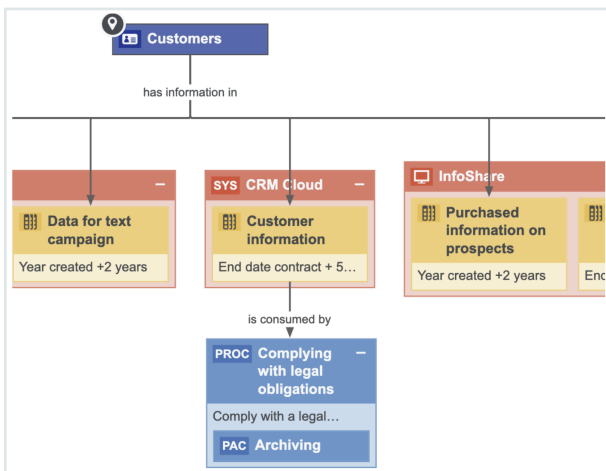
Consider the customer data. Collibra knows where this data resides and how it is being used. This is an outcome of obtaining input from the business users during the onboard of the Business Processes where the users are asked what data they use, which applications they use, for what purpose the data is used. When further onboarding of those business processes by the stewards takes place, one of these steps is mapping the Business

Processes to the data, and then also helping those business stewards with the mapping through the data classification capabilities in Catalog.

The following image shows a traceability view, which is a result of collaboration with the business team, data governance team, and other teams.



The above image shows where data resides and why it is used. It shows all the applications that contain customer data, and also the related retention periods, which can be imported when a customer wants to exercise their right to be forgotten. Collibra knows in which applications the data resides and the business processes that use that data. Thus, we know why and how we are using our customer data. This determines how to respond to the right to be forgotten because there are often Business Processes where you have the real legitimate reason to retain the customer's personal information.



When a customer wants to exercise their right to be forgotten, we can remove the information in these applications; however, we need to store the customer information in

the above table in order to comply with the legal obligation. Therefore, it is not only important to know where your personal information resides, but also why you are using it. Such information is important information for applications that process data subject requests (DSRs). You can integrate with the application that does the DSRs and create a workflow to process the data subject requests. Based on the input of the information and metadata that you will find in Collibra, you can validate the request. When the request is approved, you can point the applications to the stewards and send them a task to perform the action that appears in the data subject request, such as, removing the data or extracting the data and sending it to a customer.

The same approach can be applied to the integrated consent management applications. These applications need to know the processes for reaching the consent, and such applications reside in the process register, so that you can see all the processes that rely on the consent and the data categories for which you need consent.

The screenshot shows the 'Marketing Process Register' interface. At the top, there are navigation buttons: 'Type: Process Register', 'Export Metamodel', 'Go to the Business User Interface', 'Request input', 'Edit', 'Move', 'Delete', and 'Auto hyperlinks'. Below this, there is a section for 'CCPA Default View' with a description: 'The view presents the inventory of Business Processes describing the data flows in your organization.' There are also buttons for 'Delete', 'Move', and 'Validate'. The main content is a table with two columns: 'Name' and 'legal basis'.

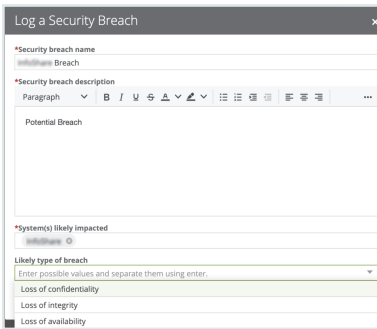
Name	legal basis
Direct Marketing	Legitimate interest
Market Research	Legitimate interest
Monetizing Marketing Insights	Consent, Consent from the minor towards selling of PI
Monetizing anonimised global Marketing Insights	Consent, Opt-out (from selling)
Monetizing Marketing Insights EU customers	Consent
Monetizing Marketing Insights US customers	Consent provided towards selling of PI due to financial incentive received,
Print media advertisement	Legitimate interest
Public Website Management	Consent provided towards selling of PI due to financial incentive received,
Public Website Content Maintenance	Consent, Substantial Public Interest
Create online contest	Consent

These are stored in the data sets that can also contain granular information, such as the individual data elements for which you want to obtain consent—this combines the information about which business processes require consent and the data categories for which you need consent to process all information in Collibra. The information governed in Collibra can be then sent to the consent management application that is used to manage consent.

Potential data breach workflow

This section describes how Collibra helps when a data breach occurs.

With Collibra Data Privacy, Collibra for Desktop, or Collibra for Mobile, you can report any suspicious behavior by logging a potential data breach.



If your organization has suffered a potential data breach, you can determine the application that needs to be investigated and the type of breach that may have occurred, and then log a potential data breach. The related workflow will require the community manager on the data governance counsel to assign the issue manager who will investigate the breach. The issue manager will then investigate the issue, assess the potential impact of the breach, determine the reporting requirements (for example, to whom the incident must be reported), and plan the remediation actions to address the risks. The reporting evidence needs to be stored. If you go to the data helpdesk, you can find an overview of all the breaches that are being investigated.

Name ↑	Description	Assignee	Requester	Reviewer
BigSuite - sent credentials ove...	Employee accidentally sent	Preston Sterling	William Parker	Dora Perreman
Data Breach Blogger	Today it is mentioned in the new	Preston Sterling	David English	Dora Perreman
Example of Breach	Description			

Collibra can help with investigating the impact of the breach because of the knowledge of which data resides in the applications and the processes that use those applications. Such a holistic view on where the data resides, which applications are involved, and the processes that rely on these applications can help in assessing the impact on customers following a data breach. Collibra can not only help an organization log and investigate a data breach but also help analyze the impact of the breaches because Collibra knows

where the data resides and how it is being used. In addition, it contains a history of all the breaches (including potential ones) that would have been logged.

How do we get there?

This section describes the Process register and Business Process discovery capabilities, data categorization and classification, and different prescriptive paths for reaching out from the logical data layer envisioned in the metamodel graph and connected data sets to a physical data layer present in columns located directly at the data source.

Create and maintain Process Register

Process Register is an essential part of privacy compliance, foreseen directly by GDPR article 30 as a Record of Processing Activities and derived from CCPA requirements for performing data mapping in the organization. Process Register enables to store assets of the Business Process type that describes processes in the organization that involve personal data. In Collibra, Business Processes reflect the requirements stated by Processing Activity in GDPR.

Business Process onboarding

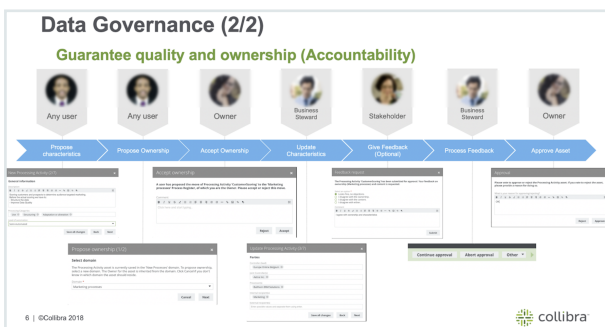
Business Processes may be onboarded by business users as well as privacy stewards through dedicated workflow implementing guided stewardship principle in Collibra Data Privacy. During onboarding, multiple roles collaborate in providing content to the onboarded Business Process. Because of the dedicated tasks and required approval and feedback, assets are onboarded in a governed way.

In the scenario on the Personal Information (PI) Discovery, it was described how Collibra helps with discovering Personal Information. But equally important to knowing where you are storing personal information is knowing why you are using personal information. That is, what the legal context of using that PI is. This context is created within Process Registers, throughout the usage of Business Processes that describe the processes conducted by organization relating to the usage of personal information.

Typically, that information does not reside with one person that can help you document that knowledge. That information is stored within multiple areas across the organization

and it may not be easy to centralize this information and ensure that the information is up to date. To help you with this task, CollibraData Privacy comes with the Business Process discovery capabilities.

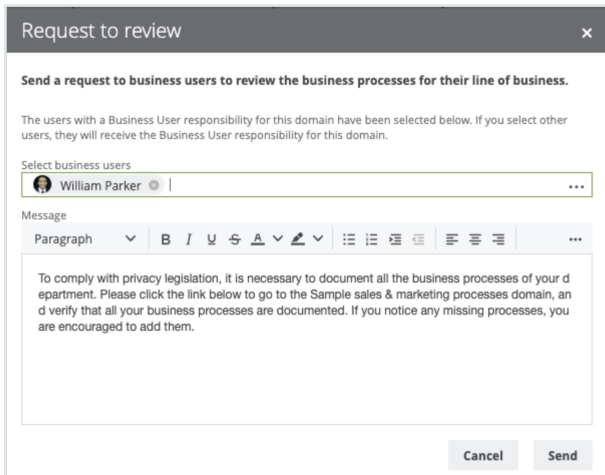
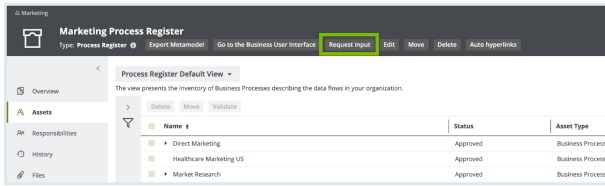
Consider a high-level overview of Collibra Privacy and Risk Business Process discovery capabilities. It commences with the Business Users describing the Business Processes in their terms. They will describe the data being used, applications being used, and any third parties with which they share information. After describing the Business Process, the owner of the Business Process will accept the ownership of that particular Business Process. When the ownership is accepted, the experts or the stewards will further onboard the proposed Business Process. This means that they will ensure that the Business Process is accurate and actionable because that Business Process provides business context on how we use personal information and we must ensure that the description is accurate. Therefore, in principle, you will have the Business Steward, Privacy Steward, and Data Steward, each adding business metadata, adding privacy metadata, and performing data mapping, respectively. After the stewards have updated the characteristics, you can optionally obtain feedback from the stakeholders. The following sections describe each step involved in the process.



Requesting business users' input with a dedicated interface

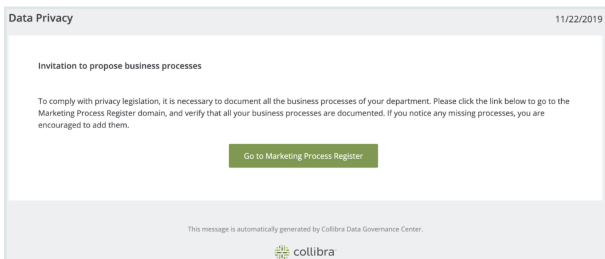
The information related to Business Processes may be requested from the Business User directly from Collibra Privacy and Risk Process Register. Typically, this will be done by those who work on the Privacy program. With the Request input button, email will be generated for the selected business users, which can provide relevant information on the business side of the process through a dedicated interface. You can have a guiding text that explains the purpose of your request. If you click Send, an email is sent to the business user with an invitation to contribute to the Process Register.

Chapter 1

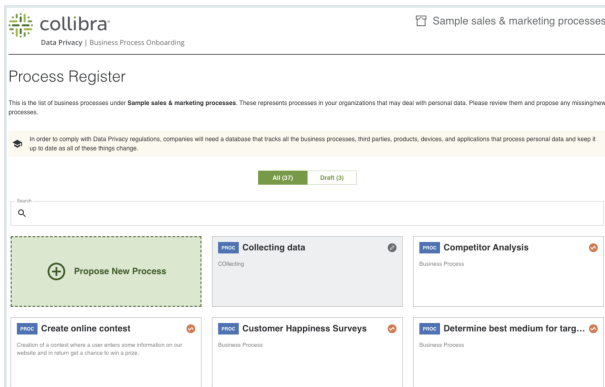


Onboarding Business Process with a business user interface

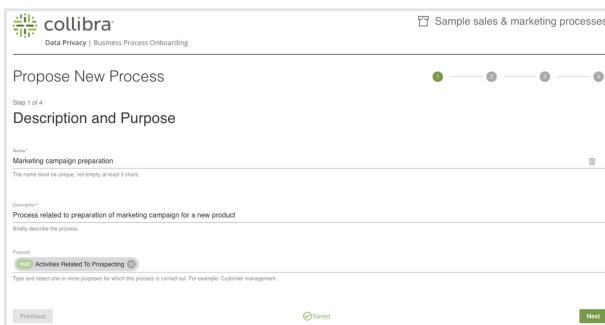
The Business Register User receives an email message asking them to verify that all the processes are in their domain.



When the Business User clicks Go to Marketing Process Register in the email, a page showing all the Business Processes for their department appears to allow the Business User to contribute to the Process Register.



The link provided in the email message directs the User to a survey where they can describe the business processes that they perform on a daily basis. If the Business User cannot find the Business Process that was onboarded was in the process of being onboarded, they can propose a new Business Process using the Propose Business Process button. When proposing a Business Process, they can describe the Business Process, provide a unique name, description, and purpose.

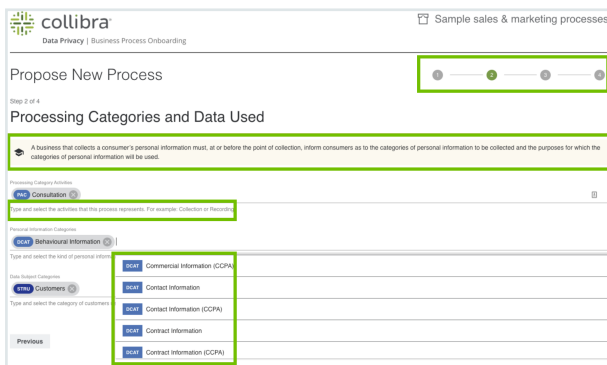


The next step involves covering Process Categories such as use, collection, adaptation, and alteration. The Business User defines the types of data that they are using, for example, behavioral information, contact information, or contract information. Finally, they determine what type of customer's data they are using, such as the customers covered by CCPA or GDPR. There can be also an indication on other types of data subjects, such as, employees and candidates. The Business User can select values only from predefined lists—this reduces the scope of errors as there is no ambiguity around the values that the Business User is able to provide. These values have been predefined by the Privacy team and have legal implications. They show how the organization complies with the privacy regulations. Because, when you collect data directly from customers or from a third parties—by using sensitive information, public information, or customer or employee information—the distinctions will have an impact on how you comply with the regulations.

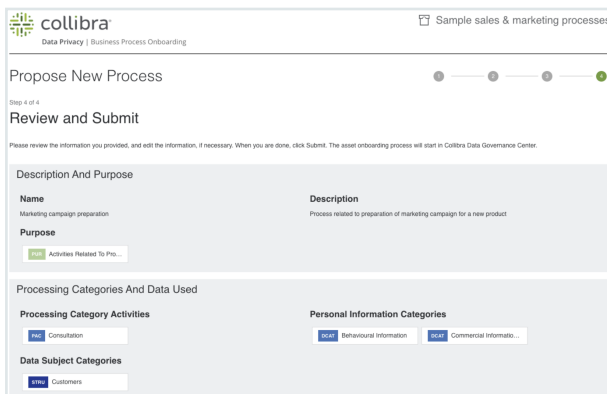
For example, the employee information is temporarily exempted from the CCPA. Therefore, it is considered better for the Business User to select from the drop-down list, as opposed to providing free text. This also prevents common issues such as spelling errors. In addition, if there is any uncertainty about the meaning of these values, the Business User can look up the definitions of these values in Collibra. In the next steps, the lines of business and third parties involved can be described, applications used can be indicated, and the level of automation in the Process can be determined.

The wizard is prescriptive:

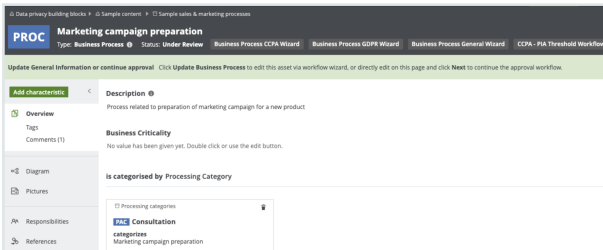
- It shows the user the steps that they have completed and how many steps are remaining, by visually indicating the progress.
- The help text below the question describes what is required from a particular question.
- The ability to open a side panel that provides additional educational information such as the wordings from the law or video content from the Collibra university.
- Smart suggestion based on what the user has already filled and the domain to which they belong.



After entering the information, the Business User can review it before submitting it.

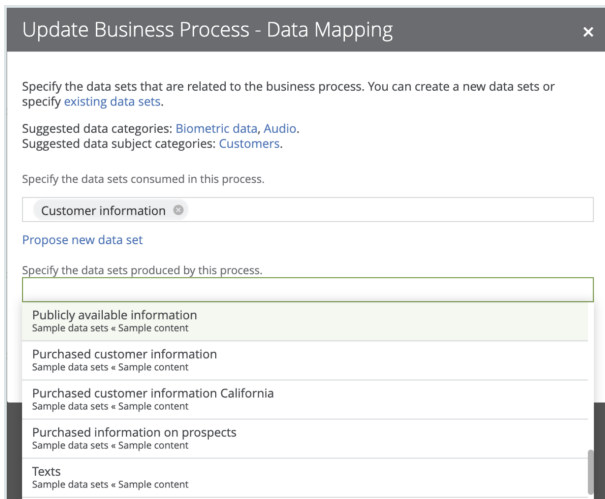


After the Business User provides their Business Process, they can submit the process for further onboarding. The next step is for the owner to accept the ownership of that Business Process. A new task is generated for the owner after they accept the ownership of the Business Process in the Process Register. Based on the metadata, the owner can determine that the Business Process belongs to their Process Register. The ownership can be accepted or rejected. As a result, the status of the asset is changed and the justification is added in the Comments section of the Business Process.

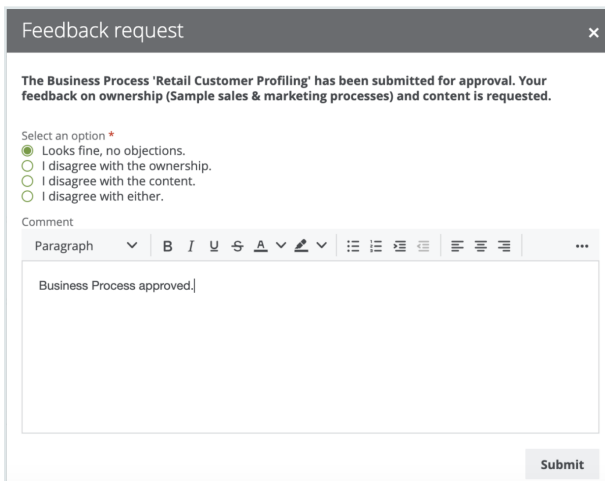


As a next step, the experts or the stewards will be consulted to ensure that the metadata is accurate and complete and the Business Process is going to be mapped to the data. In the following steps, relevant tasks will be created with the request to review and update attributes when necessary. Among others, the Data Mapping task is performed by Data Stewards. The form contains contextual help with suggestions on the relevant Data Sets consumed and produced by the Business Process. Because one of the data categories used in the process is behavioral information, you can click and review it and view the related data sets in the categories. Based on this, Data Stewards can ensure that Data Mapping has been correctly performed.

The last expert who needs to contribute to the Business Process is the Privacy Steward. After opening the task, the first step is to define the regulation that applies, be it GDPR, CCPA, or others. In addition, a purpose needs to be validated, and legal bases, controllers and processors need to be defined. Very specific information on regulation shall be specified, for example, on GDPR, we define cross-border transfers, safeguards, consent collection method, and automated decision-making confirmation. On the CCPA side, we are asked about the collection directly from customers or third parties and whether the data is being sold to third parties.



After the Stewards finish updating the Business Process, we ask the Stakeholders for final feedback. If the feedback is positive, we send the task to the Owner for the final approval.

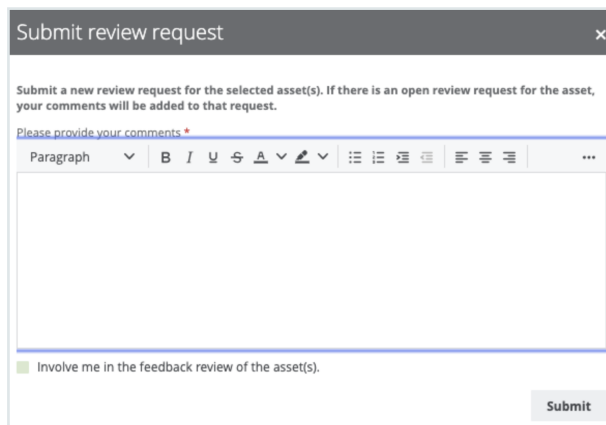
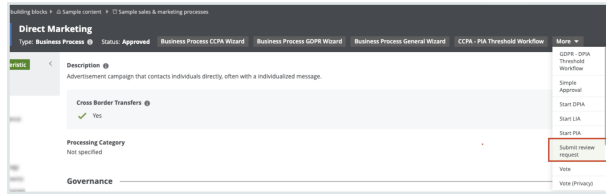


Maintain Process Register over time with review requests

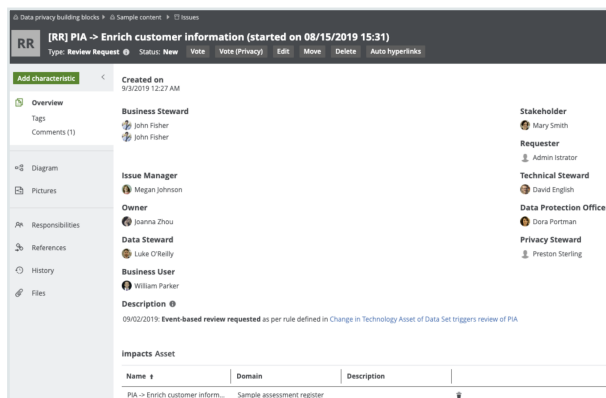
Whereas the successful result of the asset onboarding process is a new asset with the status Approved, asset change management is the standardized procedure for making changes to such approved assets.

You may have many reasons to review an approved asset. Collibra Data Privacy groups such reasons into three categories and offers three corresponding means to trigger a review request:

- **Manual:** A trigger that is manually actioned by a user if, for example, the user wants to request a review of a Business Process asset considered to be incomplete or inaccurate. Any user can manually request a review of an approved asset.



- **Time-based:** A trigger that is automatically actioned at a specified frequency. This is useful for assessment assets for which you might be required to review periodically to comply with a regulation.



- **Event-based:** A trigger that is automatically actioned by the fact of changes made to specified characteristics of the related asset.

All of the review requests are available in the Data Helpdesk.

Name	ID	Description
[R] Customer Information - 2019/09/02 22:03	09/02/2019	Manual review requested by Admin labator, refer to comments below.
	09/02/2019	Request accepted by Admin
	09/02/2019	Review request implemented
[R] Direct Marketing - 2019/08/08 15:52	08/08/2019	Manual review requested by Admin labator, refer to comments below.
[R] Enrich customer information - 2019/09/02	09/02/2019	Manual review requested by Admin labator, refer to comments below.
[R] PIA - Enrich customer information start...	09/02/2019	Event-based review requested as per rule defined in Change in Technology Asset of Data Set triggers review of PIA.
[R] Travel & Expenses - 2019/09/10 08:51	09/10/2019	Manual review requested by Admin labator, refer to comments below.
	09/10/2019	Request accepted by john.fisher

Perform Assessments

Conduct PIA and DPIA

If a business process is likely to introduce a level of risk to the rights and freedoms of natural persons, the Business Steward or the Data Protection Officer must perform the following:

- Privacy Impact Assessment (PIA), if complying with CCPA
- Data Privacy Impact Assessment (DPIA), if complying with GDPR

To determine whether or not you need to perform such an assessment for a Business Process asset, you must run a Threshold workflow.

The potential for business processes to expose the rights and freedoms of natural persons to risk is significant. Privacy Impact Assessments (PIA) and Data Privacy Impact Assessments (DPIA) assess the risks to the rights and freedoms of data subjects, born of a specific business process.

After onboarding a Business Process asset, the relevant Threshold workflow helps you determine whether or not a PIA or DPIA is needed. If it is determined that an assessment is necessary, the Owner or the Business Steward for the Business Process asset must complete the relevant workflow:

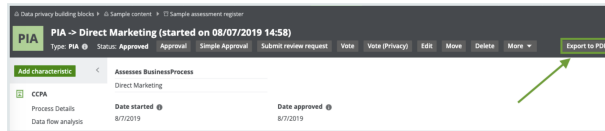
- PIA, if complying with CCPA
- DPIA, if complying with GDPR

Print assessment results

Assessments are a way for an organization to demonstrate compliance. You can export and print the PIA results in a unified way. You can also download a PIA asset page as a printable PDF, regardless of the status of the PIA asset.

Steps

1. Go to the relevant PIA asset page.



2. Click **Export to PDF**.
 - » The PDF is downloaded to your computer.

Data privacy building blocks > Sample content > Sample assessment register Print date: 2019-11-04

PIA PIA -> PIA -> Direct Marketing (started on 08/07/2019 14:58)

Status: Approved Date started: 8/7/2019 | Last modified: 11/4/2019 | Modified by: Istrator Admin

Final decision: 1. Processing allowed

Business Process assessed by PIA
[Direct Marketing](#)

General Description
 In the Direct Marketing Process, we send target marketing materials to our customers and prospects. We profile our customers to categorize our customers in 4 categories, to which we can send marketing materials that are customized to the category the customers belong to.

Details

Personal information usage
 We are processing Personal Information for Direct Marketing Purposes. We are not selling Personal Information. We are in full control of the PI

Personal information source
 Directly from the custom
 From a third par

Purpose of personal information usage defined
 Yes
 Justification not provided

Data flow analysis

Personal information categories <input checked="" type="checkbox"/> Yes Justification not provided	Third parties <input checked="" type="checkbox"/> Yes Justification not provided
Sharing of personal information <input checked="" type="checkbox"/> Yes Justification not provided	Data sharing agreements <input checked="" type="checkbox"/> No We still need to update the Data Sharing Agreements

Controls analysis

Minimization <input checked="" type="checkbox"/> Yes We have minimized the PI to what is absolutely	Quality <input checked="" type="checkbox"/> No No Data Quality process implemented yet. Not the
--	--

1 of 3

Install Collibra Protect

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

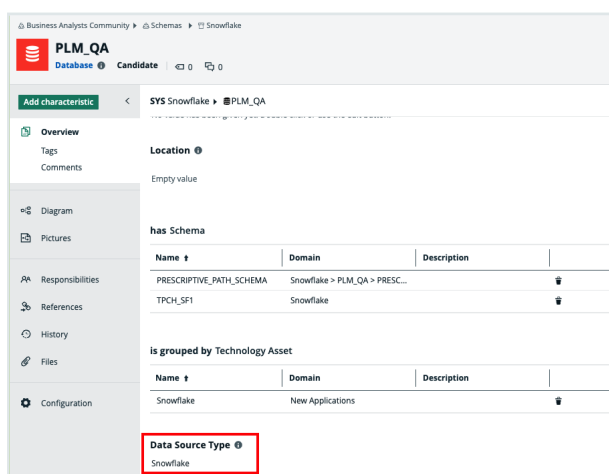
Prerequisites

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.

Name	Description	Required license	Members
Catalog Author		Standard	Admin Istrator
Data Dictionary		Read-only	Everyone
DataSteward	Allows usage of Data Steward...	Read-only	Everyone
DataSteward Author		Standard	
Edge integration engineer	Allows managing connections...	Standard	
Edge manager	Allows creating and deleting E...	Standard	
Edge site	Allows connection from Edge ...	Standard	Edge privacy-risk-qa-25-07
Edge site administrator	Allows downloading Edge site...	Standard	
Glossary	Allows usage of Business Glo...	Read-only	
Helpdesk		Read-only	Everyone
Insights		Standard	Everyone
Policy Manager		Read-only	Everyone
Protect Admin	In this role, you have the sam...	Standard	Admin Istrator
Protect Author	In this role, you can create rul...	Standard	Admin Istrator, Author User
Protect Manager	This is a role for our system u...	Read-only	Edge privacy-risk-qa-25-07, Policy Lifecycle Management API User
Protect Reader	In this role, you can view Colli...	Read-only	Reader User
ReferenceData	Allows usage of Reference Da...	Read-only	Everyone
Sysadmin	Allows for configuring and m...	Standard	Admin Istrator

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.

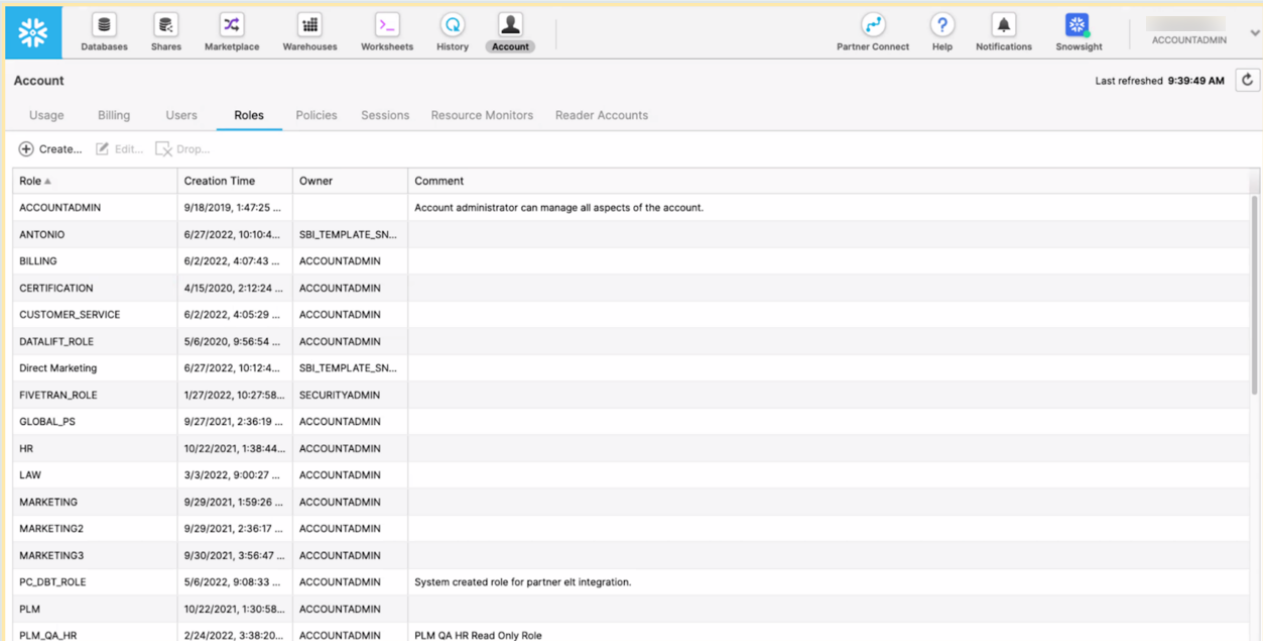
Roles	Description
Protect Reader	Users in this role can view Collibra Protect with read-only access to the content. This role is assigned to 'Everyone' and grants the users the 'protect' permission. Without this permission, users cannot see 'Protect' as an application in the ☰ menu. They also cannot 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.
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. Collibra 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.

Groups

Adding Groups
 To add a group, you have to use the [Collibra Protect Group API](#). Currently, only Snowflake data sources are supported.

Group Name	System Reference	Created By	Created Date
------------	------------------	------------	--------------

Navigation: Databases, Shares, Marketplace, Warehouses, Worksheets, History, Account

Partner Connect, Help, Notifications, Snowsight, ACCOUNTADMIN

Reset | HUMAN RESOURCES | SALES | Roles | New Worksheet | Standard Introspection | BILLING

Run | All Queries | Saved 10 seconds ago

ACCOUNTADMIN | DEMO_WH (M) | PLM_GA | TPC_H_SF1

```
1 SHOW roles;
```

Results | Data Preview | Open History

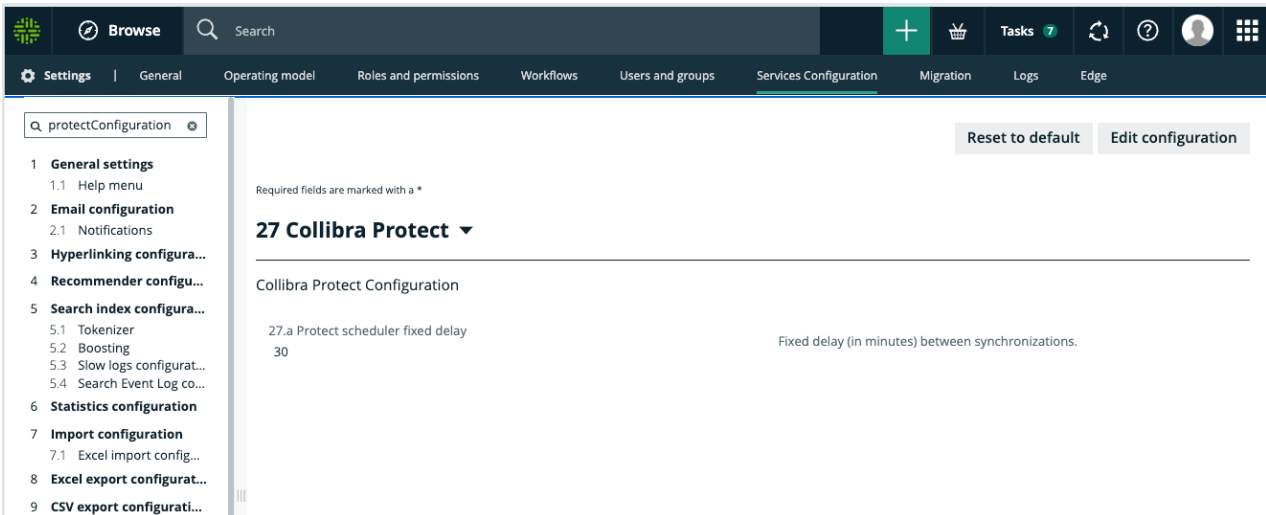
Query ID: SQL | 90ms | 37 rows

Filter result... | Copy

Row	created_on	name	is_default	is_current	is_inherited	assigned_to_users	granted_to_roles	granted_roles	owner	comment
1	2019-09-17 16:47:2...	ACCOUNTADMIN	N	Y	N	35	0	3		Account administrat...
2	2022-06-27 01:10:4...	ANTONIO	N	N	N	1	1	0	SBL_TEMPLATE_SN...	
3	2022-06-02 07:07:...	BILLING	N	N	N	1	0	0	ACCOUNTADMIN	
4	2020-04-15 05:12:2...	CERTIFICATION	N	N	Y	1	1	0	ACCOUNTADMIN	
5	2022-06-02 07:05:...	CUSTOMER_SERVICE	N	N	N	1	0	0	ACCOUNTADMIN	
6	2020-05-06 00:56:...	DATALIFT_ROLE	N	N	Y	1	2	0	ACCOUNTADMIN	
7	2022-06-27 01:12:4...	Direct Marketing	N	N	N	1	0	1	SBL_TEMPLATE_SN...	
8	2022-01-27 13:27:5...	FIVETRAN_ROLE	N	N	Y	3	1	0	SECURITYADMIN	
9	2021-09-27 05:36:1...	GLOBAL_PS	N	N	N	1	0	0	ACCOUNTADMIN	
10	2021-10-22 04:38:4...	HR	N	N	Y	10	1	0	ACCOUNTADMIN	
11	2022-03-03 00:00:...	LAW	N	N	N	0	0	0	ACCOUNTADMIN	
12	2021-09-29 04:59:...	MARKETING	N	N	Y	11	1	0	ACCOUNTADMIN	

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.

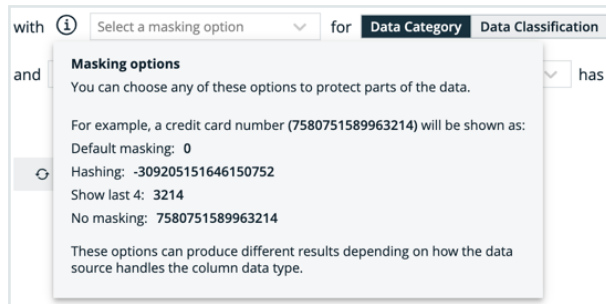
2. Column-based protection

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

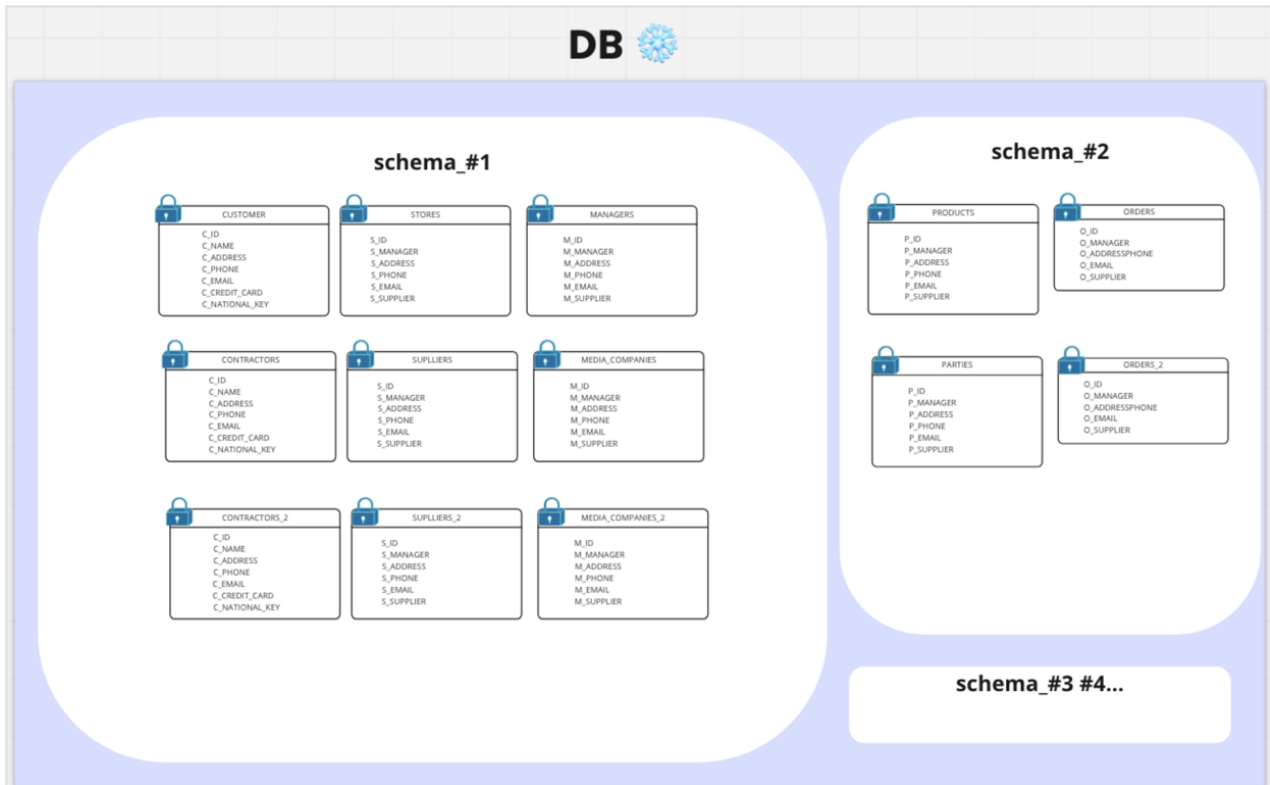
Another way to protect your data is to filter rows of a specific table. If you do not want to expose all of the existing items in a table because one of the columns is part of a certain data classification, you can easily leverage the Collibra operating model to do so.

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**.

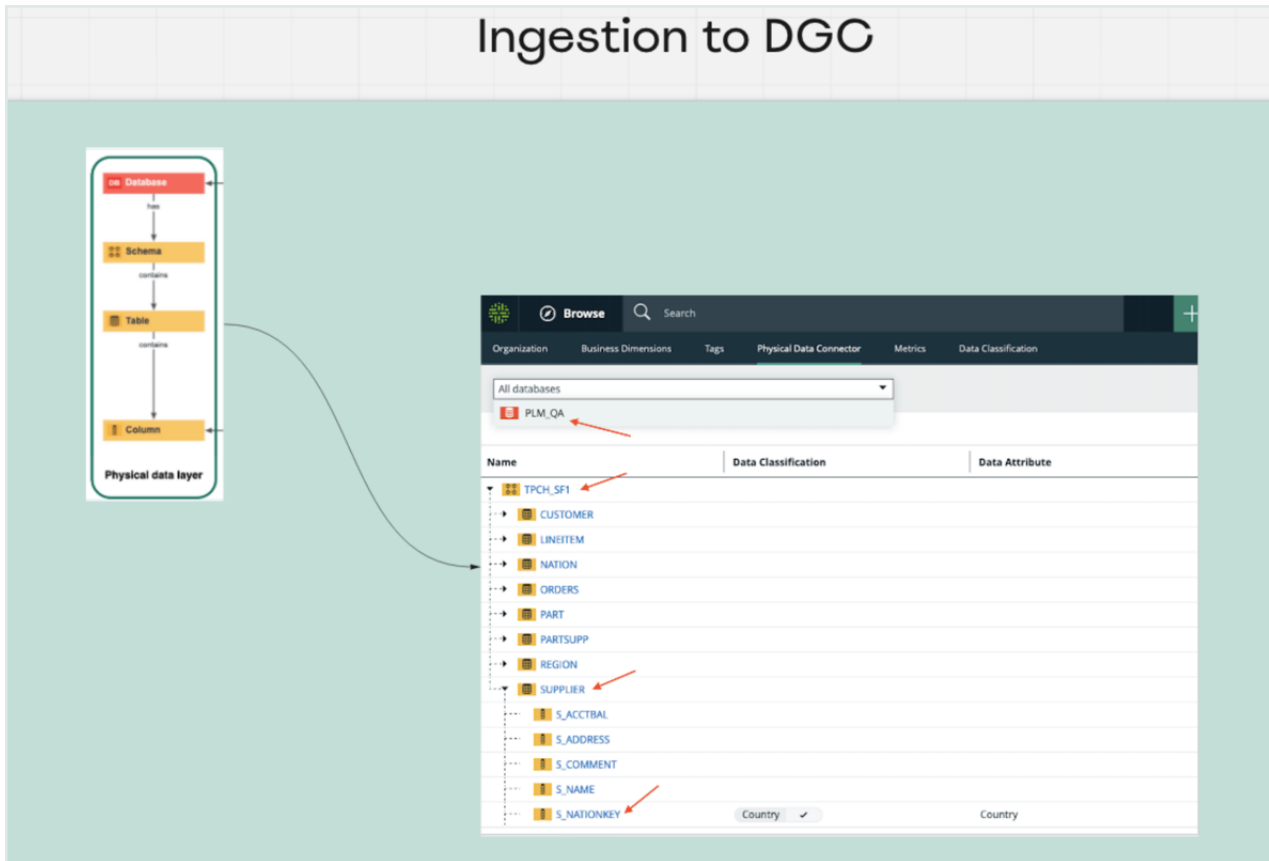
Technical background

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 and data access rules

[Data protection standards](#) and [data access rules](#) govern your data with ease and clarity.

Data protection standards

Data protection standards create a layer of protection for similar types of data by masking them wherever they are.

For example, if columns with the first and last names are a part of the PII data category, regardless of the tables, schemas, and databases to which they belong, you can create a data protection standard that targets all of these columns, by choosing the PII data category and masking it.

Data access rules

After establishing this primary layer (blanket) of protection to your most sensitive data, you can use data access rules to manage access and enhance protection for specific usages.

For example, you can create a data access 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 data protection standard that you created earlier.

Tip When creating [data protection standards](#) or [data access rules](#) for assets, consider how those assets are grouped. Suppose that you have a Business Process asset, BP, which contains the following Data Set assets: DS1, DS2, and DS3. Instead of creating a [data protection standard](#) or [data access rule](#) for each of the three Data Set assets (DS1, DS2, and DS3), consider creating a standard or rule that targets the Business Process asset (BP), to save time.

Frequently asked questions

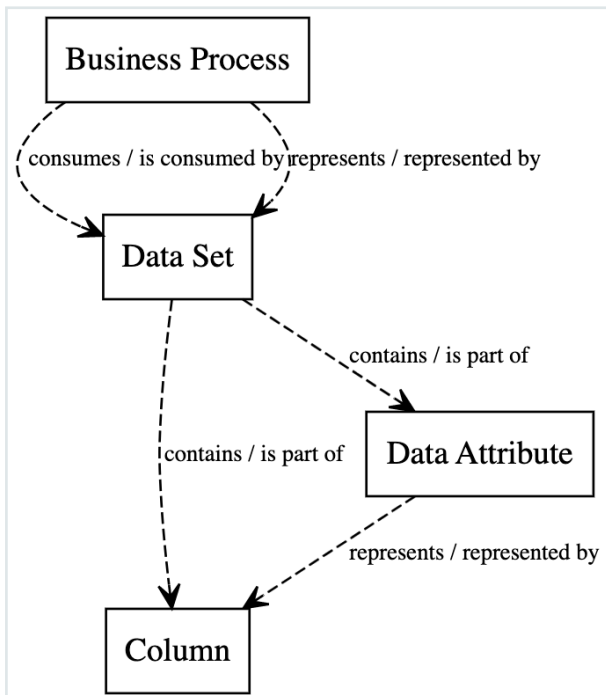
- What if I want to grant access to a group without having the PII masked?
 - » When creating a rule for an asset that contains data masked by a standard, choose to override it by unmasking it or changing its masking type.
- What if I want to grant access to a group, but the protection from the standard is not enough because there might also be other sensitive data within a 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

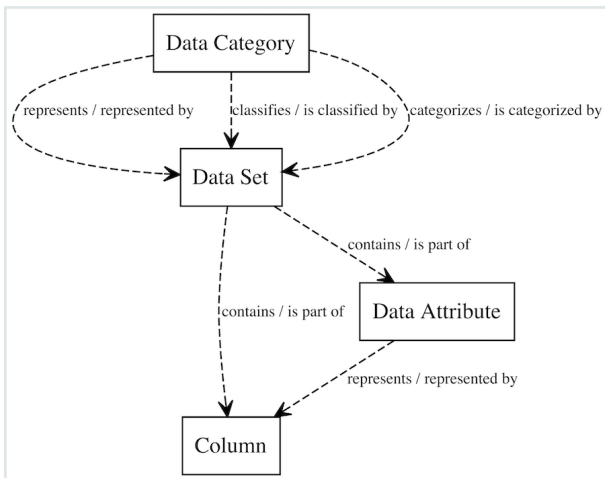
The assets that you use to create data protection standards and data access rules are related to the physical data layer, such as tables and columns, through a set of relations and intermediate assets. Collibra Protect uses these relationships and intermediate assets to search the knowledge graph to find the physical data layer assets that it needs to protect. The traversal of the knowledge graph follows a set of prescriptive paths. Each asset type has a set of prescriptive paths for traversing to the Column asset, as illustrated in the following sections.

Note Depending on your permission, you can customize the prescriptive paths. For more information, go to [Customization of prescriptive paths](#).

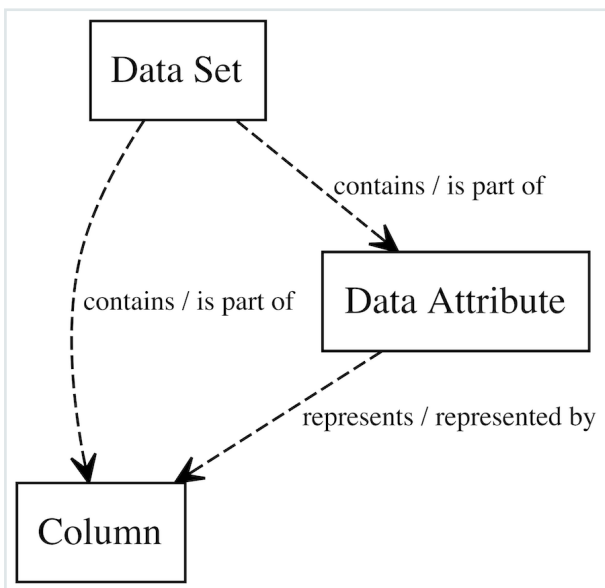
Business Process



Data Category



Data Set



Customization of prescriptive paths

Collibra Protect supports the following asset types:

- Packaged asset types: Business Process, Data Category, and Data Set
- Custom asset types: These are the packaged asset types that you have modified or the asset types that you have created. If you modify the attributes and relations of a packaged asset type, then the packaged asset type becomes a custom asset type.

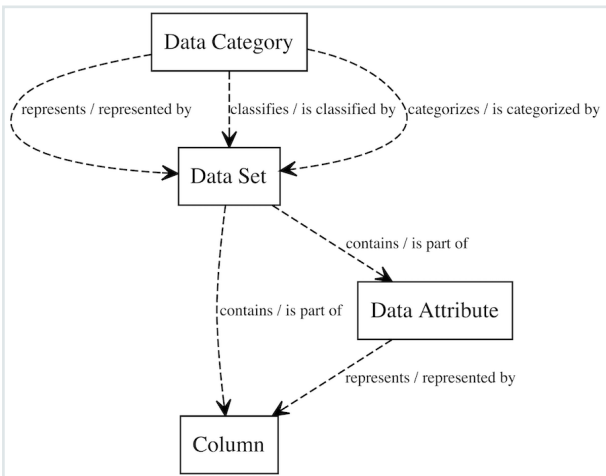
If you have the **Protect > Administration** global permission, you can customize the **prescriptive paths** for the asset types through APIs. The customization may include creating, modifying, or deleting the prescriptive paths: for example, adding or modifying the prescriptive paths for packaged and custom asset types, defining how the asset types relate to columns, removing any obsolete prescriptive paths.

The customized prescriptive paths are applied to data protection standards and data access rules.

Note You cannot remove a customized prescriptive path if an asset type linked to the prescriptive path is used in a data protection standard or a data access rule.

Collibra Protect supports a maximum of 10 asset types. Each asset type can have a maximum of 6 relations and a maximum depth of 3. However, when customizing the prescriptive path for an asset type, we recommend that you provide only one relation for the asset type. Prescriptive paths must always end in a Column asset type (that is, 00000000-0000-0000-0000-000000031008).

The following image is an example of a prescriptive path that has 6 relations and a depth of 3.



If you want to restore the default asset types defined by Collibra, a PATCH operation must be performed on each asset type. The list of asset types and their specifications are as follows.

If Data Privacy is not installed

Data Set (00000000-0000-0000-0001-000400000001)

```

    {
      "description": "Prescriptive path from Data Set to Column",
      "relations": [
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000000031008"
          }
        },
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000000031005",
            "relation": {
              "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
              "relationTypeDirection": "SOURCE",
              "assetType": {
                "assetTypeId": "00000000-0000-0000-0000-0000000031008"
              }
            }
          }
        }
      ],
      "assetTypeId": "00000000-0000-0000-0001-000400000001"
    }
  
```

Data Category (00000000-0000-0000-0000-0000000031109)

```

    {
      "description": "Prescriptive path from Data Category to Column",
      "relations": [
        {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
          "relationTypeDirection": "SOURCE",
        }
      ]
    }
  
```



```

    "assetType": {
      "assetTypeId": "00000000-0000-0000-0001-
000400000001",
      "relation": {
        "relationTypeId": "00000000-0000-0000-0000-
000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-
000000031008"
        }
      }
    }
  },
  {
    "relationTypeId": "00000000-0000-0000-0000-
000000007038",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0001-
000400000001",
      "relation": {
        "relationTypeId": "00000000-0000-0000-0000-
000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-
000000031005",
          "relation": {
            "relationTypeId": "00000000-0000-0000-0000-
000000007094",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-
000000031008"
            }
          }
        }
      }
    }
  },
  {
    "relationTypeId": "00000000-0000-0000-0000-
000000007007",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0001-
000400000001",
      "relation": {
        "relationTypeId": "00000000-0000-0000-0000-

```

```

000000007062",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0000-
000000031008"
    }
  }
},
{
  "relationTypeId": "00000000-0000-0000-0000-
000000007007",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0001-
000400000001",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-
000000031005",
        "relation": {
          "relationTypeId": "00000000-0000-0000-0000-
000000007094",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-
000000031008"
          }
        }
      }
    }
  }
},
{
  "assetTypeId": "00000000-0000-0000-0000-000000031109"
}

```

Business Process (00000000-0000-0000-0000-000000031103)

```

{
  "description": "Prescriptive path from Data Set to Column",
  "relations": [
    {
      "relationTypeId": "00000000-0000-0000-0000-

```

```

000000007062",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0000-000000031008"
  }
},
{
  "relationTypeId": "00000000-0000-0000-0000-
000000007062",
  "relationTypeDirection": "SOURCE",
  "assetType": {
    "assetTypeId": "00000000-0000-0000-0000-
000000031005",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-
000000007094",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-
000000031008"
      }
    }
  }
}
],
"assetTypeId": "00000000-0000-0000-0001-000400000001"
}

```

If Data Privacy is installed

Data Set (00000000-0000-0000-0001-000400000001)

```

{
  "description": "Prescriptive path from Data Set to Column",
  "relations": [
    {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-000000031008"
      }
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-
000000007062",

```

```

        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
                }
            }
        }
    },
    ],
    "assetTypeId": "00000000-0000-0000-0001-000400000001"
}

```

Data Category (00000000-0000-0000-0000-000000031109)

```

    {
        "description": "Prescriptive path from Data Category to Column",
        "relations": [
            {
                "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0001-000400000001",
                    "relation": {
                        "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
                        "relationTypeDirection": "SOURCE",
                        "assetType": {
                            "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
                        }
                    }
                }
            }
        ],
        {
            "relationTypeId": "00000000-0000-0000-0000-0000-000000007038",
            "relationTypeDirection": "SOURCE",
            "assetType": {

```

```

        "assetTypeId": "00000000-0000-0000-0001-
000400000001",
        "relation": {
            "relationTypeId": "00000000-0000-0000-0000-
000000007062",
            "relationTypeDirection": "SOURCE",
            "assetType": {
                "assetTypeId": "00000000-0000-0000-0000-
000000031005",
                "relation": {
                    "relationTypeId": "00000000-0000-0000-0000-
000000007094",
                    "relationTypeDirection": "SOURCE",
                    "assetType": {
                        "assetTypeId": "00000000-0000-0000-0000-
000000031008"
                    }
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007007",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-
000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007007",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",

```

```

        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
          "relation": {
            "relationTypeId": "00000000-0000-0000-0000-0000-000000007094",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
            }
          }
        }
      },
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-0000-000000007315",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0001-0004000000001",
        "relation": {
          "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
          }
        }
      }
    },
    {
      "relationTypeId": "00000000-0000-0000-0000-0000-000000007315",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0001-0004000000001",
        "relation": {
          "relationTypeId": "00000000-0000-0000-0000-0000-000000007062",
          "relationTypeDirection": "SOURCE",
          "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-0000-000000031005",
            "relation": {

```

```

    "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007094",
    "relationTypeDirection": "SOURCE",
    "assetType": {
      "assetTypeId": "00000000-0000-0000-0000-0000-000000031008"
    }
  }
}
],
"assetTypeId": "00000000-0000-0000-0000-000000031109"
}

```

Business Process (00000000-0000-0000-0000-000000031103)

```

  {
    "description": "Prescriptive path from Business Process to Column",
    "relations": [
      {
        "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0001-000400000001",
          "relation": {
            "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
            "relationTypeDirection": "SOURCE",
            "assetType": {
              "assetTypeId": "00000000-0000-0000-0000-000000031008"
            }
          }
        }
      },
      {
        "relationTypeId": "c0e00000-0000-0000-0000-0000-000000007314",
        "relationTypeDirection": "SOURCE",
        "assetType": {
          "assetTypeId": "00000000-0000-0000-0001-000400000001",
          "relation": {

```

```

        "relationTypeId": "00000000-0000-0000-0000-
000000007062",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0000-
0000000031005",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007094",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-
0000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007038",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-
0000000031008"
                }
            }
        }
    },
    {
        "relationTypeId": "00000000-0000-0000-0000-
000000007038",
        "relationTypeDirection": "SOURCE",
        "assetType": {
            "assetTypeId": "00000000-0000-0000-0001-
000400000001",
            "relation": {
                "relationTypeId": "00000000-0000-0000-0000-
000000007062",
                "relationTypeDirection": "SOURCE",
                "assetType": {
                    "assetTypeId": "00000000-0000-0000-0000-

```



```
000000031005",
    "relation": {
      "relationTypeId": "00000000-0000-0000-0000-
000000007094",
      "relationTypeDirection": "SOURCE",
      "assetType": {
        "assetTypeId": "00000000-0000-0000-0000-
000000031008"
      }
    }
  }
},
],
"assetTypeId": "00000000-0000-0000-0000-000000031103"
}
```

Open Protect

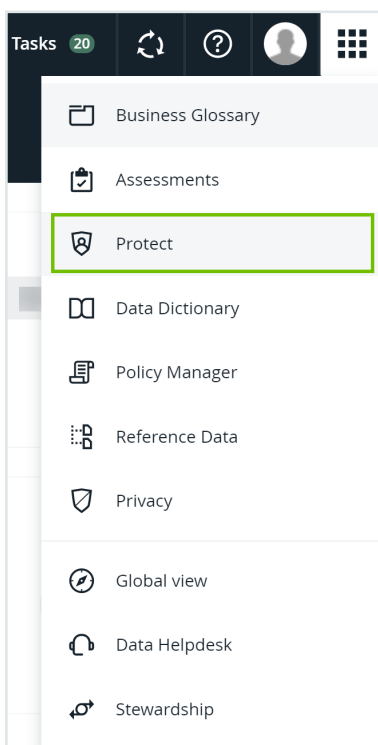
This topic describes how to open Protect, including how you can use the [tabs](#) on the Protect landing page.

Requirements and permissions

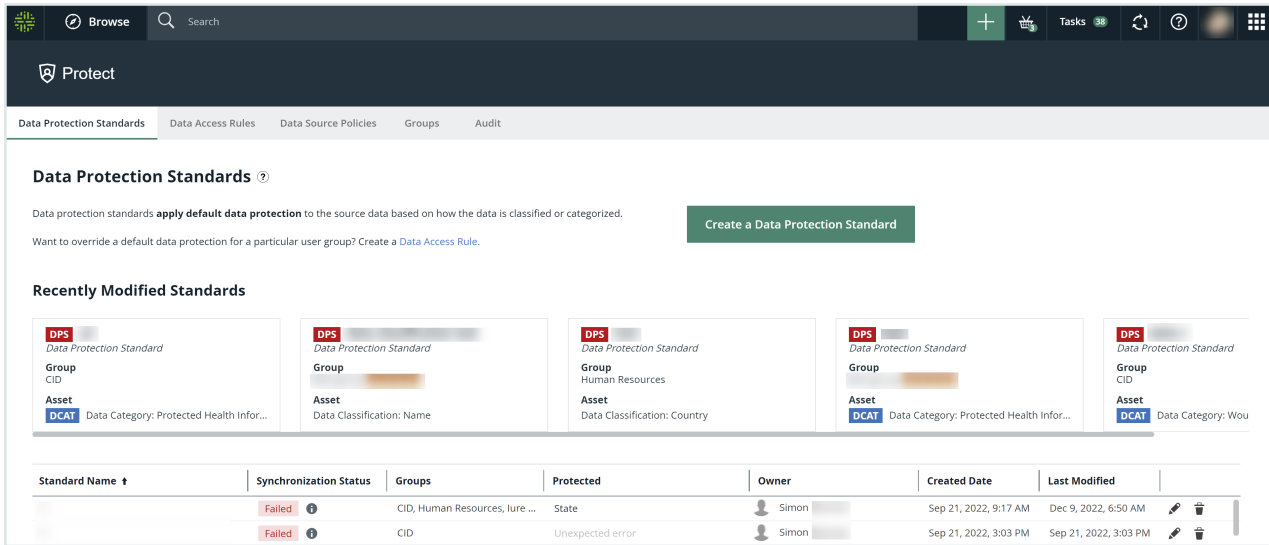
You have a global role that has the [Protect global permission](#).

Steps

On the main menu, click , and then click **Protect**.



» The Protect landing page is shown.



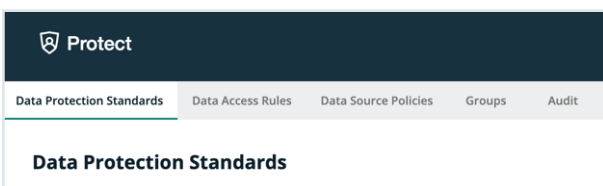
Tabs on the Protect landing page

On the Protect landing page, depending on your role, the following tabs are shown.

Tab	Description
Data Protection Standards	View or create standards to define data source access to data types based on data categories, data attributes, or data classifications.
Data Access Rules	View or create rules to grant specific groups different access to the same data in data sets, business processes, or data categories. <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note Data access rules take precedence over data protection standards.</p> </div>
Data Source Policies	View the policies that are active in the data source tables. These include the policies that were manually created in the data source and the policies that were generated in the data source due to data protection standards and data access rules in Protect. <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Tip Contact Collibra support to import policies from the data source using the Collibra Protect Data Source Policies API.</p> </div>
Groups	View or create groups for data protection standards and data access rules. You can create groups using the Collibra Protect Group API . <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note You must create at least one group before creating a standard or a rule.</p> </div>
Audit	Generate an audit log of the 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.
Pending	This standard has been created or edited, and is pending synchronization.

Synchronization Status	Description
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; background-color: #f9f9f9; 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 * 

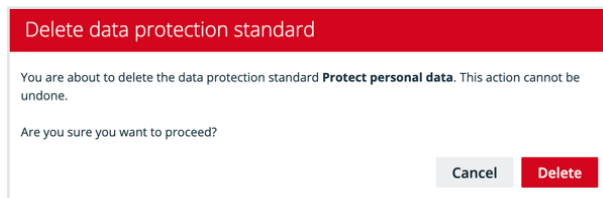
Summary
For the Group Human Resources and Marketing
protect [GDPR data related to criminal convictions and offences](#)
with Default masking

Delete a standard

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

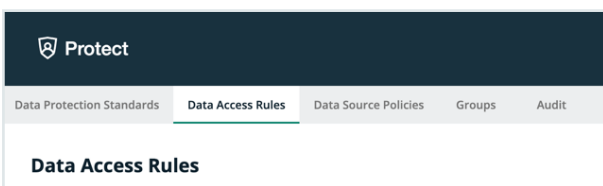
Steps

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 of protection to your most sensitive data using data protection standards, you can create data access rules to manage access to the data sources and enhance protection for specific usages.

Steps

1. In Collibra Protect, click the **Data Access Rules** tab.
2. Click **Create a Data Access Rule**.
 - » The **Create a Data Access Rule** dialog box appears.
3. Enter the required information.

Field descriptions

Field	Description
Rule Name	The name to identify the data access rule.
Description (optional)	A description for the rule.
Group	Group for the rule.

Field	Description
Asset	<p>The data asset that the rule is protecting. This field contains Business Process, Data Category, and Data Set assets, as well as assets of custom asset types.</p> <div data-bbox="1118 618 1422 1144"><p>Tip</p><ul style="list-style-type: none">○ For more information, go to Technical background and Prescriptive paths.○ You can add more groups using the plus icon.</div>

Field	Description
Optional: Select a masking option	<p>The type of masking that you want to apply to a data category or data classification. The following options are available:</p> <ul style="list-style-type: none">○ Default masking○ Hashing○ Show last○ No masking <p>In the Select a data category/data classification field, select the data category or data classification for the masking option that you selected.</p> <div data-bbox="1123 1037 1418 1335"><p>Tip You can add more data categories and data classifications for masking using the plus icon.</p></div>

Field	Description
Optional: Select an action	<p>The type of row-filtering action that you want to apply to a data classification with a specific code set and code value. The following actions are available:</p> <ul style="list-style-type: none">○ Show○ Hide <ol style="list-style-type: none">a. In the Select a data classification field, select the data classification that you want to show or hide.b. In the Select a code set field, select the code set for the data classification.c. In the Select a code value field, select the code value for the code set. <div data-bbox="1118 1279 1420 1509" style="border-left: 2px solid #00A651; padding-left: 10px; margin-top: 10px;"><p>Tip You can add more data classifications for row-filtering using the plus icon.</p></div>

Tip

- The grant access checkbox is selected by default. The selected checkbox indicates that you are granting access to the tables and the columns in the database that are linked to the selected assets to the groups that you selected in the rule. If you do not want to grant this level of access to the selected groups, clear the checkbox.
- The **Summary** section shows a summary of the rule.

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 and apply default masking to Genetic data

Set rule for

group * Marketing + -

asset * Geographic Information + -

Grant access to the data linked to these assets.
By checking this box, additional access is given to the data tables or columns linked with the selected assets. If this box is unchecked, no access is given to the selected assets, but they can still be protected. **Note: once the rule granting access is saved and synchronized, access to these assets 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 Select a data classification 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

4. To preview the rule, click **Generate Preview**.

Tip The preview shows only the first 1,000 affected columns. The drop-down list box below the **Generate Preview** button is used to switch between the assets that you selected in the rule. Each asset has its own preview table.

5. Click **Save Rule**.

- » A message stating that the data access rule is sent to source appears, and the rule is shown in the table containing rules.

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 Data Classification + -

and rows where has

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


↻ Generate Preview

Cancel
Save Rule

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.

Delete data access rule

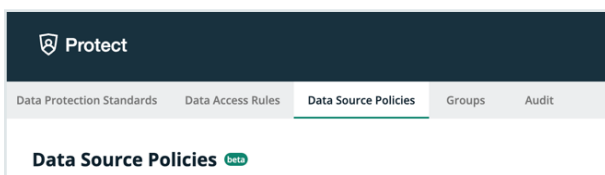
You are about to delete the data access rule **Rule 1**. This action cannot be undone.

Are you sure you want to proceed?

Cancel
Delete

Data source policies

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

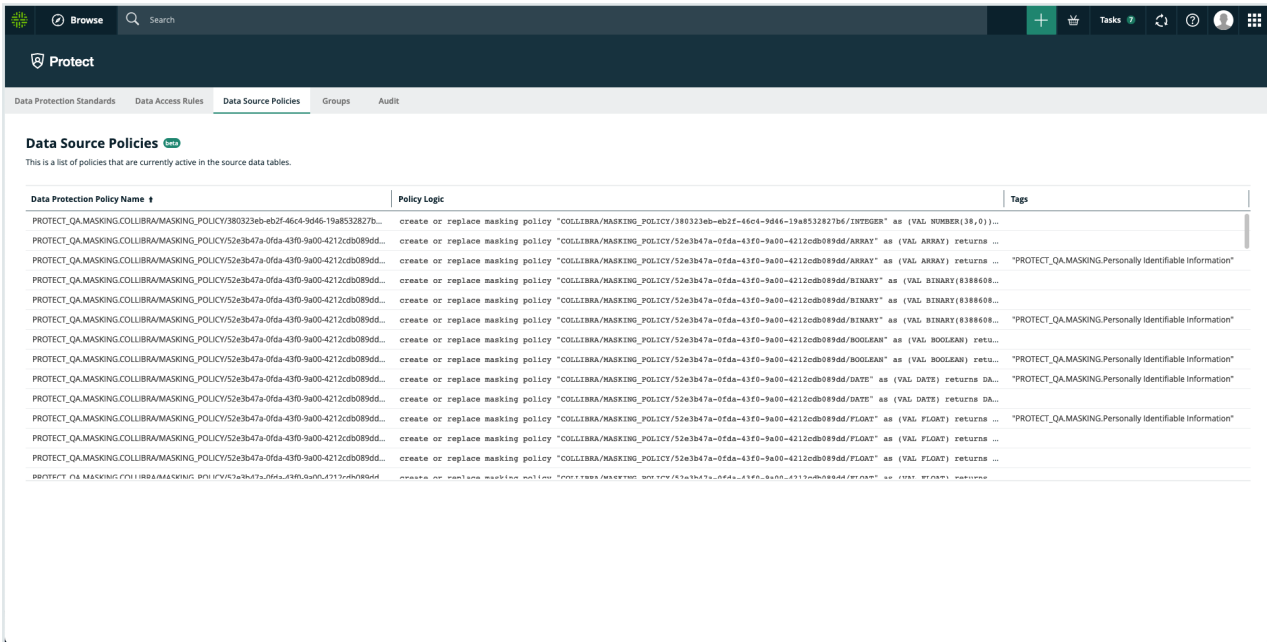


The data protection policy table displays a list of policies that are currently active in the source data tables. This includes policies that were created via Collibra Protect as well as policies that were created in the data source manually.

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 Collibra 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.
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.



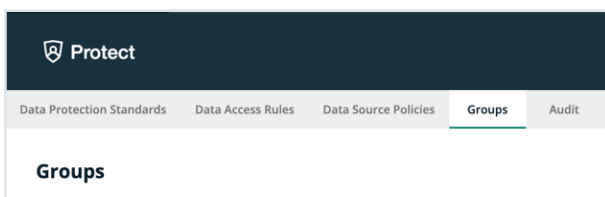
Types of policies on Snowflake

There are three types of policies on Snowflake: Column-based policies, row access policies, and tag-based policies. Each type can be created in Collibra 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 displays the 'Groups' page in Collibra Protect. It features a header with the 'Protect' logo and navigation tabs. Below the tabs, 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.' Below this is a table listing active groups.

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 Collibra Protect group
System Reference	
Created By	User who created the Collibra Protect group
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

An audit log contains information about the queries that were run to access the data and the data that was accessed.

Generate an audit log

You can generate an audit log of access records from the data source on the **Audit** page.

Note The time that it takes for the actions performed in a data source to appear in an audit log in Collibra Protect varies from several minutes to hours, depending on the data source.

Steps

1. In Collibra Protect, click the **Audit** tab.
2. Depending on your data source, click **BigQuery** or **Snowflake**.
3. Click one of the following buttons: **Today**, **Yesterday**, **A week ago**, **30 days ago**.

Tip The start date corresponding to the button that you clicked is shown in the **Start Date** field. Alternatively, you can enter or select a date in the **Start Date** field, and then click **Generate Log**.

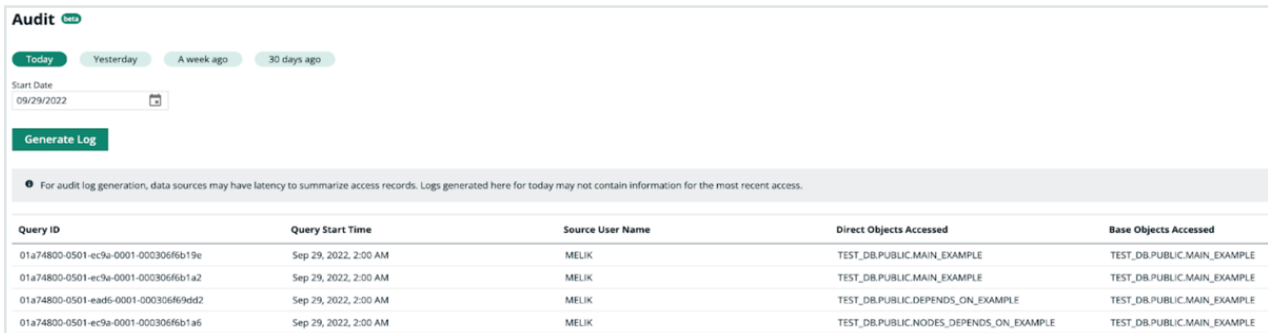
» The audit log is generated.

Important

- Generating an audit log may take up to a minute. After clicking **Generate Log**, do not navigate away from the **Audit** page because doing so cancels the audit log generation.



- The audit log contains the first 1,000 records from the selected start date. If you want to view the remaining records, contact your data source administrator.



The screenshot shows the 'Audit' interface with a 'Generate Log' button and a table of audit records. The table has five columns: Query ID, Query Start Time, Source User Name, Direct Objects Accessed, and Base Objects Accessed. The records show queries executed by user MELIK on Sep 29, 2022, at 2:00 AM, accessing various database objects.

Query ID	Query Start Time	Source User Name	Direct Objects Accessed	Base Objects Accessed
01a74800-0501-ec9a-0001-0003066b19e	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.MAIN_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ec9a-0001-0003066b1a2	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.MAIN_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ea46-0001-0003066f94d2	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.DEPENDS_ON_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE
01a74800-0501-ec9a-0001-0003066b1a6	Sep 29, 2022, 2:00 AM	MELIK	TEST_DB.PUBLIC.NODES_DEPENDS_ON_EXAMPLE	TEST_DB.PUBLIC.MAIN_EXAMPLE

Audit log data

The following table describes the columns that are shown in an audit log.

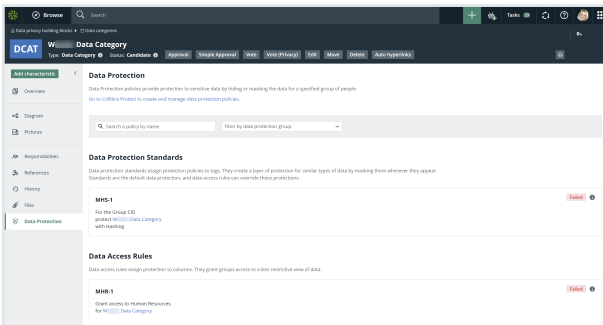
Column	Description
Query ID	The ID of the query in the source database.
Query Start Time	The date and time of the query in the source database.
Source User Name	The name of the user in the source database who ran the query to access the data.
Direct Object Accessed	The database object (a table or a view) that was used to access the data.
Base Object Accessed	The database object that was accessed.

Data protection in the asset pages

The asset pages for the following asset types contain the **Data Protection** tab to allow you to view, filter, create, and manage data protection standards and data access rules:

- [Business Process](#)
- [Data Category](#)
- [Data Set](#)
- Custom asset types such as [Column](#), [Database](#), [Schema](#), and [Table](#), derived from the aforementioned asset types via [prescriptive paths](#)

Note Data protection standards support only Data Category assets and data classifications.



View or filter standards and rules

Requirements and permissions

You have the **Protect Reader** global role.

Steps

On the asset page (for the one of the [aforementioned](#) asset types), click the **Data Protection** tab.

» Data protection standards and data access rules that are linked to the asset are shown.

Tip

- To filter the standards and rules by name, in the **Search a policy by name** field, enter the name of the standard or rule that you want to view.
- To filter the standards and rules by group, in the **Filter by data protection group** field, select the group for which you want to view the standard or rule.

Create or manage standards and rules

Requirements and permissions

You have the **Protect Author** and **Protect Admin** global roles.

Steps

1. On the asset page (for the one of the [aforementioned](#) asset types), click the **Data Protection** tab.
2. Click the following link: **Go to Collibra Protect to create and manage data protection policies.**

Tip For information about how to create and manage data protection standards and data access rules, go to [Data protection standards](#) and [Data access rules](#).

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

This topic contains examples to describe how data protection standards and data access rules behave when different types of masking affect the same column.

Note In the 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 + - for **Data Category** **Data Classification** Personal and family details

and rows where has

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

Conflicting filters affecting the same column

This topic contains examples to describe how data protection standards and data access rules behave when conflicting filters affect 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 ▼ for **Data Category** Data Classification Select a data category ▼

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

and Hide ▼ rows where State ▼ 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 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 Select a data category

for Data Category Data Classification

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

Chapter 11

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

Collibra Protect periodically synchronizes with an aggregation of all data protection standards and data access rules. These standards and rules form a data source-agnostic representation containing all databases, schemas, tables, and columns, as well as their protections and accesses. The synchronization process then triggers the [Edge capabilities](#), such as Collibra Protect for Snowflake. These Edge capabilities 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.

Protect for Snowflake

Data protection standards in Collibra Protect rely on the [tag-based masking policies](#) of Snowflake. The name of the data category or data classification selected in a standard becomes a tag with the same name. The tag is applied to all the affected columns to enforce data protection. For more information, go to [Examples](#).

Examples

This topic contains examples to describe how Snowflake behaves in relation to certain data protection standards and data access rules.

Example 1



Introduction

This example describes the behavior in Snowflake when a standard is applied to a data category and a rule is applied to a data set with categorized columns in Protect.

The example considers the following:

- A standard created for the **Everyone**, **Human Resources**, **Marketing**, and **Sales** groups, to protect the columns in the **Personally Identifiable Information** data category by default masking.

The screenshot shows the configuration for a standard in the Snowflake Protect interface. It includes the following elements:

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

- A rule created for the **Human Resources** group and the **Employee Data** asset, without any protection applied to the columns in the **Personally Identifiable Information** data category.

The screenshot shows the configuration for a rule in the Snowflake Protect interface. It includes the following elements:

- Set rule for:**
- group:** Human Resources
- asset:** Employee Data
- with:** No masking
- for:** Data Category: Personally Identifiable Information

There is also a checkbox for "Grant access to all data tables linked to these asset columns." which is checked.

Standard

When the **standard** is synchronized and active, the standard results in 14 masking policies—one policy for each **Snowflake data type**. The masking policies are created at the schema level with the following naming convention: `COLLIBRA/MASKING_POLICY/<asset ID>/<snowflake type>`.

Results Data Preview

Query ID SQL 84ms 18 rows

Filter result...

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

All the masking policies are then associated with the **Personally Identifiable Information** tag, which is created at the schema level and assigned to those columns that need to be protected. At runtime, Snowflake fetches the right masking policy based on the **column data type**.

SHOW TAGS;

Results Data Preview

Query ID SQL 48ms 2 rows

Filter result...

Row	created_on	name	database_name	schema_name	owner	comment
1	2022-09-06 03:46:10.054...	Personally Identifiable Information	PROTECT_QA	DEMO	ACCOUNTADMIN	Generated by Collibra: 28d226cc-0ab0-4d23-b912-985312fb36b1

The following image shows a masking policy for the STRING data type. The data that is shown in the policy depends on the masking type selected in the standard. In the policy, `val` indicates the value as it is stored in the table.

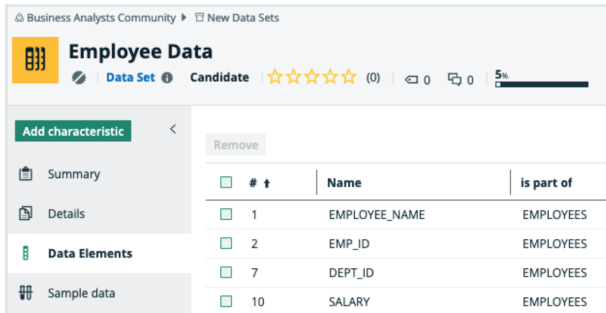
```

Details
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
    
```

Rule

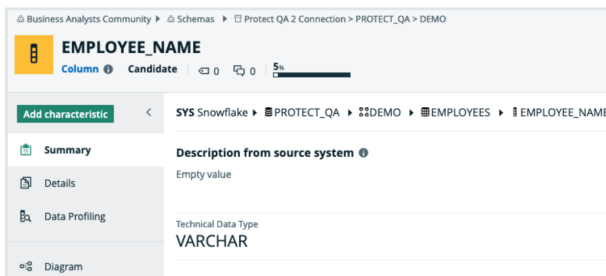
A rule results in a combination of **grant instructions**, **dynamic masking**, and **row access policies**.

Suppose that the **Employee Data** data set selected in the **rule** contains sensitive columns categorized as **Personally Identifiable Information**.



The **rule** grants access of the **Employee Data** data set to the **Human Resources** group, as indicated by the selected **Grant access...** checkbox in the rule. Then, the corresponding Snowflake role for the group can access each database, schema, and table in the data set. In addition, the column masking policy is applied to those columns that need to be protected.

Consider the **EMPLOYEE_NAME** column in the **Employee Data** data set. This column belongs to the **EMPLOYEES** table within the **DEMO** schema in the **PROTECT_QA** database.



In Snowflake, each column that is categorized as **Personally Identifiable Information** within the **Employee Data** dataset inherits the masking policy that is applied to the column in Protect. The masking policies created at the schema level use the following naming convention: **COLLIBRA/MASKING_POLICY/<asset ID>**.

Row	created_on	name	database_name	schema_name	kind	owner
18	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY7f62b08a-af5a-41ef-af64-c684-6487961	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
17	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY4667075-210f-480f-8461-46670752107	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
16	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY89806680f1-608f-808f-808f-808f80808080	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
15	2022-09-06 03:46:10.3...	COLLIBRAMASKING_POLICY8832796-8647-8844-8344-298830910ee	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN
14	2022-09-06 03:46:08.9...	COLLIBRAMASKING_POLICY082228c0-0607-4423-8073-90312836150610ANT	PROTECT_QA	DEMO	MASKING_POLICY	ACCOUNTADMIN

The following image shows the masking policy created for the **EMPLOYEE_NAME** column.

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

```

Behavior

According to the [standard](#), the **Everyone**, **Human Resources**, **Marketing**, and **Sales** groups have masked access to the data. However, according to the [rule](#), the **Human Resources** group has unmasked access to the data. As a result, the **EMPLOYEE_NAME** column has both a policy tag and a column masking policy applied to it via the standard and the rule, respectively.

In Snowflake, if both a policy tag and a column masking policy exist for a column, the column masking policy takes precedence and the policy tag is not assigned to the column. To mitigate this behavior and ensure that the protection defined in the standard is not ignored, the column masking policy also considers the conditions defined in the standard (policy tag).

Thus, when a standard is created for the **Human Resources**, **Marketing**, and **Sales** groups to mask the **Personally Identifiable Information** column by default masking, and when a rule is created for the **Human Resources** group to not mask the same column, the result is as follows:

- The column is not masked for the **Human Resources** group.
- The column is masked for the **Marketing** and **Sales** groups via default masking.

Example 2

Introduction

This example describes the behavior in Snowflake when multiple standards affect the same column without conflict.

The example considers the following:

- A standard created for the **HR** group to protect the columns in the **Personally Identifiable Information** data category by hashing.
- A standard created for the **Marketing** group to protect the columns in the **Personal Information** data category by default masking.
- The **Personally Identifiable Information** and **Personal Information** data categories share the same column named **DOB**.

Behavior

Protect creates a tag for each standard and adds a policy to each tag. The two tags are then linked to the **DOB** column. In addition, Protect creates a masking policy that is an aggregation of the policies from the two tags. This aggregated masking policy, which is then applied to the **DOB** column, thus contains the content of both the tag policies.

```
1 CASE
2     WHEN CURRENT_ROLE() = 'HR' THEN hash(val)::NUMBER
3     WHEN CURRENT_ROLE() = 'MARKETING' THEN 0
4     ELSE val
5 END
```

When a policy exists for the **DOB** column, Snowflake considers only the column masking policy, ignoring all the tag policies associated with the column. Because the column masking policy is an aggregation of all the tag policies, the protection that is defined in the two standards is not ignored.

Thus, Protect handles multiple standards with tag policies for Snowflake by creating a column masking policy, which considers the protection defined in the standards.

Masking and data types

Snowflake provides several functions to transform the data. This topic describes how Snowflake transforms the data for a given Protect masking type.

- **Default masking:** Snowflake does not support this masking type. Protect, however, uses the default masking type to apply protection to a wide range of data types. A default masking value is applied to each column according to the data type of the column.

Default masking values for data types

Column data type	Snowflake data type	Default masking value
NUMBER	NUMBER	0
DECIMAL	NUMBER	0
NUMERIC	NUMBER	0
INT	NUMBER	0
INTEGER	NUMBER	0
BIGINT	NUMBER	0
SMALLINT	NUMBER	0
TINYINT	NUMBER	0
BYTEINT	FLOAT	0
FLOAT	FLOAT	0
FLOAT4	FLOAT	0
FLOAT8	FLOAT	0
DOUBLE	FLOAT	0

Column data type	Snowflake data type	Default masking value
DOUBLE PRECISION	FLOAT	0
REAL	FLOAT	0
VARCHAR	VARCHAR	*
CHAR	VARCHAR	*
CHARACTER	VARCHAR	*
STRING	VARCHAR	*
TEXT	VARCHAR	*
BINARY	BINARY	00
VARBINARY	BINARY	00
BOOLEAN	BOOLEAN	false
DATE	DATE	1970-01-01
DATETIME	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIME	TIME	00:00:00
TIMESTAMP	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIMESTAMP_LTZ	TIMESTAMP_LTZ	1969-12-31 16:00:00.000-0800
		<p>Note This may change depending on the time zone.</p>

Column data type	Snowflake data type	Default masking value
TIMESTAMP_NTZ	TIMESTAMP_NTZ	1970-01-01 00:00:00.000
TIMESTAMP_TZ	TIMESTAMP_TZ	1969-12-31 16:00:00.000-0800 <div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> <p>Note This may change depending on the time zone.</p> </div>
VARIANT	VARIANT	0
OBJECT	OBJECT	{}
ARRAY	ARRAY	[]
GEOGRAPHY	GEOGRAPHY	{"coordinates": [0,0], "type": "Point"} (aka point(0, 0) and visualization can change based on user preferences)

- **Hashing:** Uses the following Snowflake functions:
 - *SHA2* (for strings)
 - *HASH* (for numbers)
- **Show last:** Uses the following expressions:
 - *substr(to_varchar(value), length(value) - n, n)* (for strings)
 - *mod(value, power(10,n))* (for numbers)

Tip In the expressions, *value* indicates the content and *n* indicates the number of characters to be shown.

- **No masking:** Returns the raw content.

Note

- You can apply the **Hashing** and **Show last** masking types to only the following Snowflake data types: FLOAT, NUMBER, and STRING.
- If a selected masking type cannot be applied to a certain data type—for example, when you attempt to apply the **Hashing** masking type to the DATE data type—the **Default masking** type is applied to the data type to guarantee protection.

Snowflake privileges

To perform actions in Snowflake, Collibra Protect uses an Edge connection that must be configured with a user and a role that can manage grants; create and assign masking policies, row access policies, and tags; and manage usage access on databases and schemas involved in the protection. This enforcement role requires the following Snowflake privileges.

Snowflake privilege	Description
[APPLY MASKING], [APPLY ROW ACCESS], [APPLY TAG], [MANAGE GRANTS], [IMPORTED PRIVILEGES ON DATABASE SNOWFLAKE]	Required for the role performing the actions.
[USAGE]	Required on each database and schemas where policies are applied to the role performing the actions.

Snowflake privilege	Description
<pre>[CREATE MASKING POLICY], [CREATE ROW ACCESS POLICY], [CREATE TAG]</pre>	<p>Required on each schema where policies are applied to the role performing the actions.</p>

Example

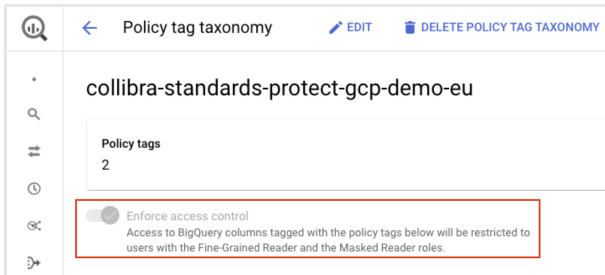
Suppose that a role named PROTECT exists in Snowflake and is responsible for managing access on all schemas within a database named DEMO. Then, the following statements can be used to enable the Snowflake PROTECT role to perform the enforcement.

```
GRANT APPLY MASKING POLICY ON ACCOUNT TO ROLE PROTECT;
GRANT APPLY ROW ACCESS POLICY ON ACCOUNT TO ROLE PROTECT;
GRANT APPLY TAG ON ACCOUNT TO ROLE PROTECT;
GRANT MANAGE GRANTS ON ACCOUNT TO ROLE PROTECT;
GRANT IMPORTED PRIVILEGES ON DATABASE SNOWFLAKE TO ROLE PROTECT;
GRANT USAGE ON DATABASE DEMO TO ROLE PROTECT;
GRANT USAGE ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE MASKING POLICY ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE ROW ACCESS POLICY ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
GRANT CREATE TAG ON ALL SCHEMAS IN DATABASE DEMO TO ROLE PROTECT;
```

Protect for BigQuery

Collibra Protect uses Google's Policy tag taxonomies to create tags and assign the tags to your BigQuery columns. Policy tag taxonomies inherently apply access control. This

means that the tags applied to your BigQuery columns will be accessible only by the Protect groups configured in your data protection standards and data access rules.



BigQuery masking rules

Each Protect masking type has an equivalent counterpart in BigQuery called a [masking rule](#). As such, masking rules in BigQuery correspond to masking types in Protect.

Note The BigQuery masking rules are not the same as the Protect data access rules.

The following table contains the equivalent [BigQuery masking rule](#) for a given Protect masking type.

Protect masking type	Equivalent BigQuery masking rule
Default masking	Default masking value
Hashing	Hash (SHA256) <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;"> <p>Note BigQuery supports the Hash (SHA256) masking rule only for certain columns depending on their data types. If Hash (SHA256) cannot be applied to a certain column due to the data type of the column, the following masking rule is applied instead: Default masking value.</p> </div>

Protect masking type	Equivalent BigQuery masking rule
Show last	Default masking value <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note BigQuery does not support the Show last masking type. The Show last masking type is supported only on MadCap:variable name="Protect.Snowflake".</p> </div>
No masking	Fine-Grained Reader <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note Each Protect group to which you assign standards has an equivalent counterpart in BigQuery called a GCP principal. BigQuery grants the Fine-Grained Reader role to the assigned GCP principal to allow the GCP principal to view the data to which no masking is applied in Protect.</p> </div>

BigQuery data types

The following table contains the BigQuery masking rule that Protect supports for a given BigQuery data type.

Summary

- Protect supports the BigQuery **Default masking value** rule for all types of columns.
- Protect does not support the BigQuery **Nullify** rule for any type of column.
- Protect supports the BigQuery **Hash (SHA256)** rule only for the following types of columns: BYTES, STRING.

BigQuery data type	BigQuery masking rule supported by Protect
ARRAY	Default masking value
BIGNUMERIC	Default masking value
BOOL	Default masking value
BYTES	<ul style="list-style-type: none"> • Default masking value • Hash (SHA256)

BigQuery data type	BigQuery masking rule supported by Protect
DATE	Default masking value
DATETIME	Default masking value
FLOAT64	Default masking value
GEOGRAPHY	Default masking value
INT64	Default masking value
INTERVAL	Default masking value
JSON	Default masking value
NUMERIC	Default masking value
STRING	<ul style="list-style-type: none"> • Default masking value • Hash (SHA256)
STRUCT	Default masking value
TIME	Default masking value
TIMESTAMP	Default masking value

BigQuery group mapping

The Collibra Protect group mapping for BigQuery must follow the syntax for principal identifiers. For example, the Protect group, **Sales**, maps to the BigQuery group email address, **sales@example.com**.

```
{
  "name": "Sales",
  "mappings":
  [
    {
      "provider": "GoogleBigQuery",
```



```
    "identity": "group:sales@example.com"  
  }  
]  
}
```

BigQuery permissions

To perform actions in BigQuery, Collibra Protect uses a GCP connection that must be configured with a service account having the following permissions:

- `bigquery.dataPolicies.create`
- `bigquery.dataPolicies.delete`
- `bigquery.dataPolicies.get`
- `bigquery.dataPolicies.getIamPolicy`
- `bigquery.dataPolicies.list`
- `bigquery.dataPolicies.setIamPolicy`
- `bigquery.dataPolicies.update`
- `bigquery.datasets.get`
- `bigquery.jobs.create`
- `bigquery.rowAccessPolicies.create`
- `bigquery.tables.get`
- `bigquery.tables.list`
- `bigquery.tables.setCategory`
- `bigquery.tables.update`
- `datacatalog.categories.getIamPolicy`
- `datacatalog.categories.setIamPolicy`
- `datacatalog.taxonomies.create`
- `datacatalog.taxonomies.get`
- `datacatalog.taxonomies.list`
- `datacatalog.taxonomies.update`
- `logging.logEntries.list`
- `resourceManager.projects.get`