

11/15/2022

# EDC Snowflake Tags - Case Study

*Sachin Jain    Principal Solutions Architect*

*Rashmi        Solutions Architect*

# Agenda

---



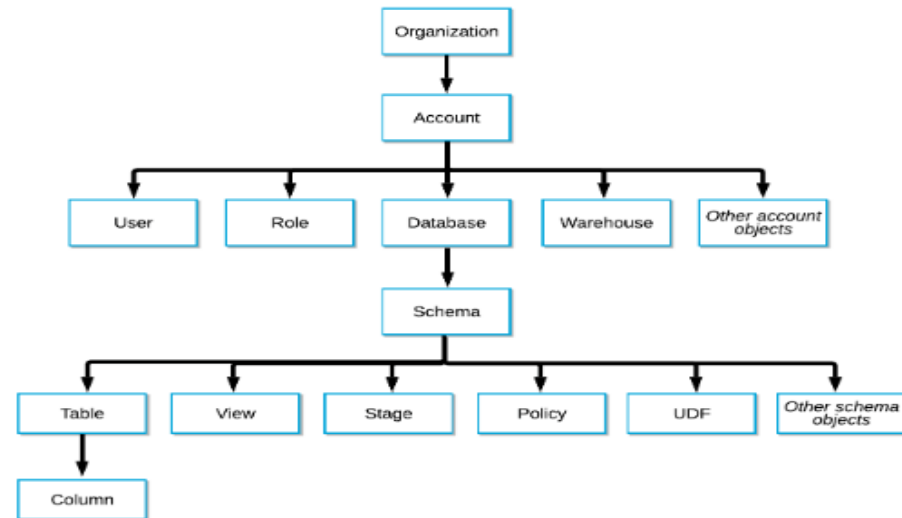
- Snowflake Tag
- Case Study
- Informatica Products - Solutions
- Demo

# Snowflake Tags

Tags enable data stewards to track sensitive data for compliance, discovery, protection, and resource usage use cases through either a centralized or decentralized data governance management approach.

Main Benefits of Tag :

Sensitive Data Tracking and Resource Usage



# Case Study: Identify Sensitive Data and Protect using Snowflake Tags

Catalog the snowflake tags in EDC and with the help of Informatica Axon, business users can effectively identify the sensitive/critical data elements and its resources. Then users can act against the sensitive elements either to mask or restrict the access based on the policies adhered to the elements.

## Steps to achieve the use case

- Catalog snowflake tags in EDC
- Onboard tagged datasets into Axon and enrich it with additional information like Policy enforcement, Stakeholder Information etc.
- Business users can search for sensitive assets using the tag values, Based on the policy information he can raise a change request to protect the actual data (Mask/Access)
- Using IICS (CDI/CDQ) user can apply Masking rules or restrict access to the data in any way
- Then user can update the masking status in axon using Custom Attribute

# Extract Snowflake Tags using EDC

## Configure it Using Metadex Scanner

### Edit configuration

Name\*

Description

Parallel Safe

Conversion path\*

#### Database object reader

- Read from (ddlReadSource)  
Select at least one of the following modes: "Datab...
- Connection registry path (repositoryPath)  
Path to the connection registry...
- Object include filters (objectFilter)  
List of include filters for database objects (view...
- Object exclude filters (objectExcludeFilter)  
List of exclude filters for database objects (view ...
- Deprecated Netezza active User name (defaultUserName)  
Unused, if set in Netezza config, please move it t...
- Skip unmodified objects (dbObjectsSkipUnmodified)  
Database objects, which did not have their source ...
- Include invalid and encrypted objects (dbIncludeInvalidObjects)  
Include invalid and encrypted objects in analysi...
- PatternReplacer configuration path (patternReplacerConfigurationPath)  
Path to a configuration of the PatternReplacer fun...

- Stored procedures
- Views
- MaterializedViews
- Triggers
- Tasks
- Pipes
- Database Assets
- Tags

false (default)

true (default)

SF\_Tags\_CSA  
SF\_Tags\_GCS

Overview Lineage and Impact Relationships Change Summary

Description

No description found. Click the ✎ icon to add a description for this asset.

Resource Contains Assets in the Resource

|              |                         |                         |          |               |
|--------------|-------------------------|-------------------------|----------|---------------|
| 30<br>Tables | 8<br>Views              | 3<br>Reference Data ... | 1<br>Tag | 1<br>Database |
| 1<br>Schema  | 1<br>Reference Data ... |                         |          |               |

#### System Attributes

CreateTime  
2022-06-15T15:09:42

Created By  
GCSTEAM

Last Modified  
Sep 23, 2022, 05:31 PM

Resource Name  
SF\_Tags\_GCS

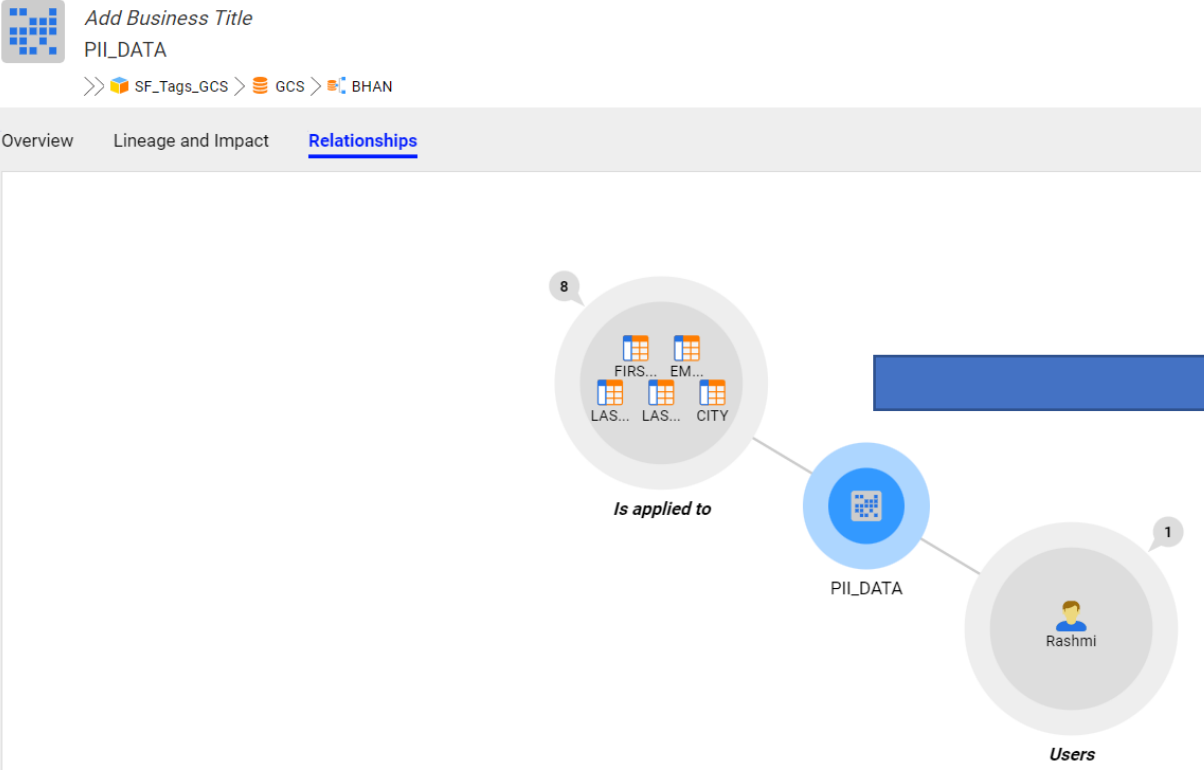
Resource Type  
Snowflake Advanced Scanner

Tag allowed values  
FirstName LastName Email CITY

Asset Type  
Tag

last\_modified\_time  
2022-06-15T15:09:42

# Tags with Objects associated



8 Is applied to

Find

| Names ^v   | Relationship |
|--|--------------|
| <b>CITY</b><br>1 Path: SF_Tags_GCS > GCS > BHAN > CUSTOMER_TAG       | Column       |
| <b>CITY</b><br>2 Path: SF_Tags_GCS > GCS > BHAN > EMPLOYEE_TAG       | Column       |
| <b>EMAIL</b><br>3 Path: SF_Tags_GCS > GCS > BHAN > CUSTOMER_TAG      | Column       |
| <b>EMAIL</b><br>4 Path: SF_Tags_GCS > GCS > BHAN > EMPLOYEE_TAG      | Column       |
| <b>FIRST_NAME</b><br>5 Path: SF_Tags_GCS > GCS > BHAN > CUSTOMER_TAG | Column       |
| <b>FIRST_NAME</b><br>6 Path: SF_Tags_GCS > GCS > BHAN > EMPLOYEE_TAG | Column       |
| <b>LAST_NAME</b><br>7 Path: SF_Tags_GCS > GCS > BHAN > CUSTOMER_TAG  | Column       |
| <b>LAST_NAME</b><br>8 Path: SF_Tags_GCS > GCS > BHAN > EMPLOYEE_TAG  | Column       |

# Columns are associated with tag values – using Script

```
EDC_Snowflake_Tags.py X
13 import os
14
15
16 requests.packages.urllib3.disable_warnings(
17     requests.packages.urllib3.exceptions.InsecureRequestWarning)
18
19 #API Used
20 search_api="/access/2/catalog/data/objects/"
21 attribute_api="/access/2/catalog/models/attributes"
22
23 file1 = open("Config.txt", "r")
24
25 #Declare Variable
26 search_API=""
27 count = 0
28 SearchURL=""
29 UName= " "
30 Pwd= " "
31 ObjectID= " "
32 TagID= " "
33 AttributeURL= " "
34 myList=[]
35 myListN=[]
36 #myTable = PrettyTable(['Index', 'Object ID', 'Custom Attribute', 'Custom Attribute Value'])
37 dfnew1 = pd.DataFrame()
38 #print(LOG)
39
40
41 #Set the Path for Log FileName
42 logname = f"{os.path.abspath(os.getcwd())}\\Log\\SnowflakeTag4. Log"
43 print("Log File "+ logname)
44
45
46 #Set the Logger
47 logger = logging.getLogger('spam_application')
48 formatter = logging.Formatter(fmt='%(asctime)s %(levelname)-8s %(message)s',
49                               datefmt='%Y-%m-%d %H:%M:%S')
50 logger.setLevel(logging.DEBUG)
51 # create file handler which logs even debug messages
```

The screenshot displays the Informatica Data Governance interface. At the top, it shows the relationship between the asset 'Add Business Title' (FIRST\_NAME) and the tag 'EMPLOYEE\_TAG'. The interface includes tabs for 'Overview', 'Lineage and Impact', and 'Relationships'. The 'Description' section is currently empty, with a message: 'No description found. Click the icon to add a description for this asset.' The 'People' section lists roles such as Data Owner, Data Steward, SME, Subject Matter Experts, and Technic\_owner, all marked as 'not assigned'. The 'Data Domains' and 'Custom Attributes' sections are also visible. A red box highlights the 'Custom Attributes' section, showing a table with columns for 'PIIDATA Tags' and 'FirstName'.

# Onboard Dataset to Axon and Enrich

Once the Object is Onboarded into Axon, create a Custom field and populate it with Tag value and apply Policy enforcement to attributes accordingly

The screenshot displays the 'DS-479: CUSTOMER\_TAG' dataset page in the Axon Data Catalog. The 'IMPACT' tab is active, showing a table of dataset attributes. A dropdown menu is open on the right, with 'Raise Change Request' highlighted. Below the main table, a 'Restricted Data Policy' is applied to the 'CITY' attribute.

| Project Name    | Level          | Attribute Name | Details    | Glossary   | Glossary Type |
|-----------------|----------------|----------------|------------|------------|---------------|
| Project Ricardo | Data Set       |                | Yes        | Party Data | DOMAIN        |
| Gerry1          | Data Set       |                | Yes        | Party Data | DOMAIN        |
| Project Ricardo | Data Attribute | EMAIL          | EMAIL      | Party Data | DOMAIN        |
| Gerry1          | Data Attribute | EMAIL          | EMAIL      | Party Data | DOMAIN        |
| Project Ricardo | Data Attribute | LAST_NAME      | LAST_NAME  | Party Data | DOMAIN        |
| Gerry1          | Data Attribute | LAST_NAME      | LAST_NAME  | Party Data | DOMAIN        |
| Project Ricardo | Data Attribute | CITY           | CITY       | Party Data | DOMAIN        |
| Gerry1          | Data Attribute | CITY           | CITY       | Party Data | DOMAIN        |
| Project Ricardo | Data Attribute | ID             | ID         | Party Data | DOMAIN        |
| Gerry1          | Data Attribute | ID             | ID         | Party Data | DOMAIN        |
| Project Ricardo | Data Attribute | FIRST_NAME     | FIRST_NAME | Party Data | DOMAIN        |
| Gerry1          | Data Attribute | FIRST_NAME     | FIRST_NAME | Party Data | DOMAIN        |

| Policy Name            | Conduct Standard  | Data Attribute | Attribute |
|------------------------|-------------------|----------------|-----------|
| Restricted Data Policy | Conduct Standards | Data Attribute | CITY      |



# Masked Status in Axon

```
60 select * from CUSTOMER_TAG;
```

Results Data Preview

✓ Query\_ID SQL 124ms 5 rows

Filter result... [Download] [Copy]

| Row | FIRST_NAME | LAST_NAME | EMAIL        | CITY         | ID |
|-----|------------|-----------|--------------|--------------|----|
| 1   | XXXX       | XXXXX     | X@XXXXX.XXX  | XXXXXXX      | 1  |
| 2   | XXXXXX     | XXXXXX    | X@XXXXX.XXX  | XXXXXXXXXX   | 2  |
| 3   | XXXX       | XXXXXX    | XX@XXXXX.XXX | XXXXXXXXXXXX | 3  |
| 4   | vvvvvvvv   | vvv       | v@vvvvvv vvv | vvvvvvvv     | 4  |

Demo