IBM Guardium
Discover and Classify

IGDC & ALATION INTEGRATION GUIDE

IBM GUARDIUM DISCOVER AND CLASSIFY

VERSION 4.2.2

TABLE OF CONTENTS

ALATION OVERVIEW

Alation is a cutting-edge platform for data governance and compliance. The IGDC's integration with Alation significantly boosts organizations' capabilities to efficiently discover, classify, and manage their data assets. By leveraging the IGDC's data cataloging and intelligence capabilities in conjunction with Alation's governance and compliance framework, organizations are equipped with a robust solution for managing data across their entire landscape. This collaboration is especially advantageous for organizations striving to streamline their data management processes, ensure regulatory compliance, and cultivate a data-literate culture.

The primary aim of this integration is to utilize the IGDC data discovery and classification features to enhance Alation's data governance capabilities. This approach enables organizations to gain a more nuanced and dynamic view of their data environment, facilitating the across varied sources, including databases, cloud storage, and SaaS platforms.

Table 1: Use cases for IGDC integration with Alation

USE CASE	DESCRIPTION
Enrichment of the Alation data catalog **(Data source export)**	The integration app exports data sources from IGDC data source catalog to the Alation data catalog as assets. This ensures the Alation's data governance capabilities cover all you data assets.
Enrichment of the IGDC data source catalog **(Data source import)**	The integration app imports the assets from the Alation data catalog to the IGDC data source catalog for further classification and cataloging. This facilitates identification and classification of sensitive data by IGDC.

APPLICATION FLOW

1. A user sets up and configures the integration in UI configurations, tests connection, and saves the configuration.

2. The integration app works in the background. It connects to IGDC and Alation to collect appropriate information, process, and send it.

3. The integration app acts as a worker. It subscribes to Kafka and updates assets in Collibra and data sources in IGDC based on new discoveries in both solutions.

PREREQUISITES

1. IGDC version 3.6.0+.

2. Alation Data Catalog version 2021.3.0+.

ALATION APP INSTALLATION, UPGRADE & REMOVAL

INSTALLATION

The integration package is provided on demand. It can be installed on a VM, machine or cluster node.

1. Extract the package directory using one of commands below depending on archive extension (zip, tar, tar.gz).

```
unzip alation-integration.zip
```

or

```
tar -xvf alation-integration.tar
```

or

```
tar -xvf alation-integration.tar.gz
```

2. Change directory to the extracted folder.

```
cd alation-integration-master
```

3. Start the stack.

```
docker-compose up -d --build
```

After stack is started you will be able to reach the API swagger of the integration on https://<APP_HOST>:5016/docs.

UPGRADE

To install a newer version of the application, the process is identical to installation. Removal of the docker volumes may be required if explicitly specified.

REMOVAL

1. To delete the application, stop the containers from the directory with deployment files.

```
cd alation-integration-master
docker-compose down
```

2. Remove the docker volumes created for the project. Volumes can be listed using the command below.

```
docker volume rm ${docker volume ls -q |grep alation}
```

ALATION CONFIGURATION

ALATION CONFIGURATION

Alation is a cutting-edge platform for data governance and compliance. IGDC discovers data sources, analyzes them, and classifies sensitive data in the data sources.

The primary aim of this integration is to utilize the IGDC data discovery and classification features to enhance Alation's data governance capabilities. This approach enables organizations to gain a more nuanced and dynamic view of their data environment, facilitating the identification and classification of sensitive data across varied sources, including databases, cloud storage, and SaaS platforms.

INTEGRATION CONFIGURATION

To configure the app settings, go to **Ecosystem > Alation Topology Reporting Extension**. Then select the **Configuration** tab and configure the app settings.

Table 1: Alation configuration

PARAMETER	DESCRIPTION
Inventa's base URL	Hostname of the IGDC CM instance. *(Required)*
Inventa's Base URL	Hostname of the IGDC instance. *(Required)*
Username for Inventa instance	Username of the IGDC user designated to the Alation integration app. *(Required)*
Password for user in Inventa instance	Password of the IGDC user designated to the Alation integration app. *(Required)*
Inventa's KeyCloak port	Number of the KeyCloak port used by IGDC. *(Optional)*
Alation's Base URL	Hostname of the Alation instance. *(Required)*
Username for Alation instance	Username of the designated user with admin access to Alation. *(Required)*
User's password for Alation instance	Password of the designated user with admin access to Alation. *(Required)*
Hostname of Inventa's external Kafka	Hostname of the IGDC Kafka. *(Required)*
Port for Inventa's external Kafka	Port of the IGDC Kafka for the integration app to connect to. *(Required)*
Username of Inventa's external Kafka	Username for access to the IGDC Kafka. *(Required)*
Password of Kafka's user	Password for access to the IGDC Kafka. *(Required)*
Security protocol used by Kafka	Name of the security protocol used by IGDC Kafka. *(Required)*
Import data sources from Alation to Inventa	Allow the integration app to import data sources from the Alation instance to the IGDC data source catalog. Options: True (import on), False (import off).
Export data sources from	Allow the integration app to export data sources and metadata (data elements, data

PARAMETER	DESCRIPTION
Inventa to Alation	categories, sensitivity classes, contains personal data tag) from the IGDC data source catalog to the Alation instance. Options: True (export on), False (export off). **Note:** To enable export of the data source metadata into the Alation assets, follow the configuration instructions below.

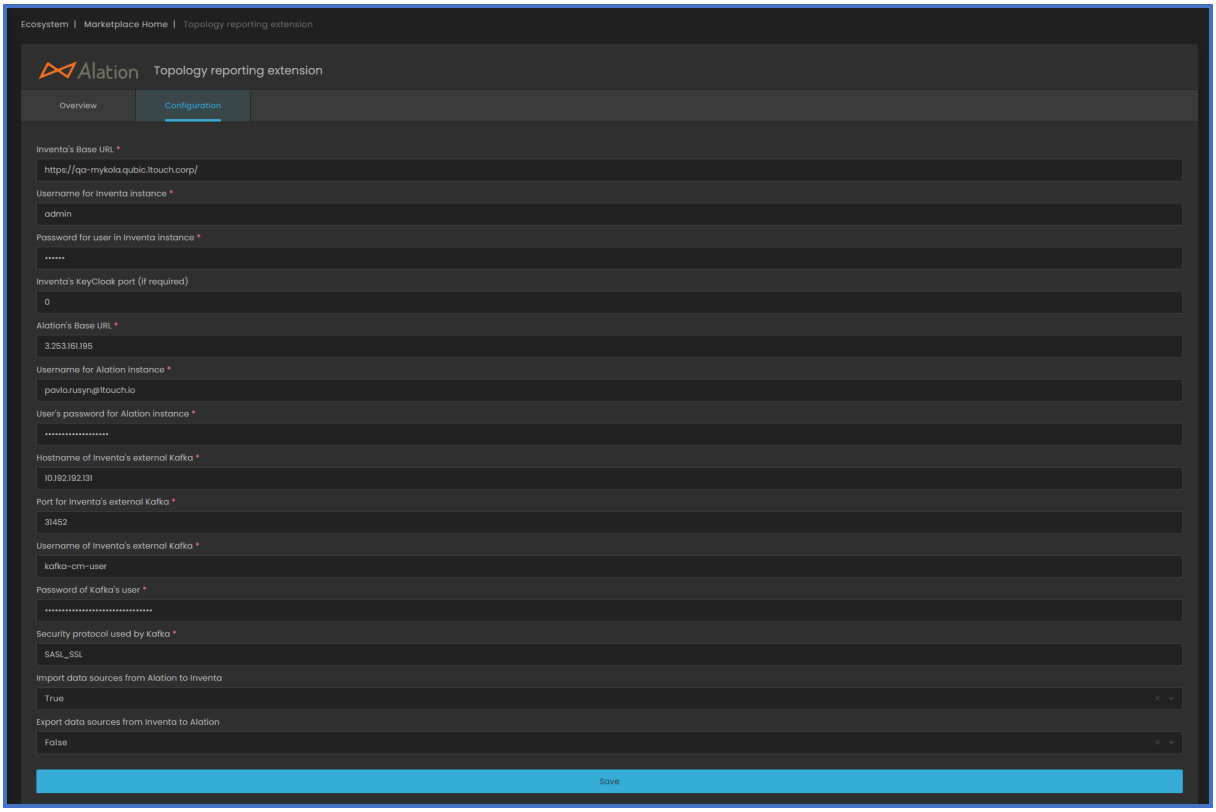

Figure 2: Alation configuration

ACTIVATION OF METADATA EXPORT TO ALATION

By default, the integration app can send data sources to Alation and create assets there, but cannot automatically add metadata (data elements, data categories, sensitivity classes, contains personal data tag) to the assets. This requires additional actions in the Alation template. To enable sending metadata, follow the instructions below.

1. In IGDC platform, go to **Ecosystem > Alation > Configuration tab** and set **Export data sources from Inventa to Alation** to **Yes**. This will trigger creation of a custom template in Alation, that needs to be configured.

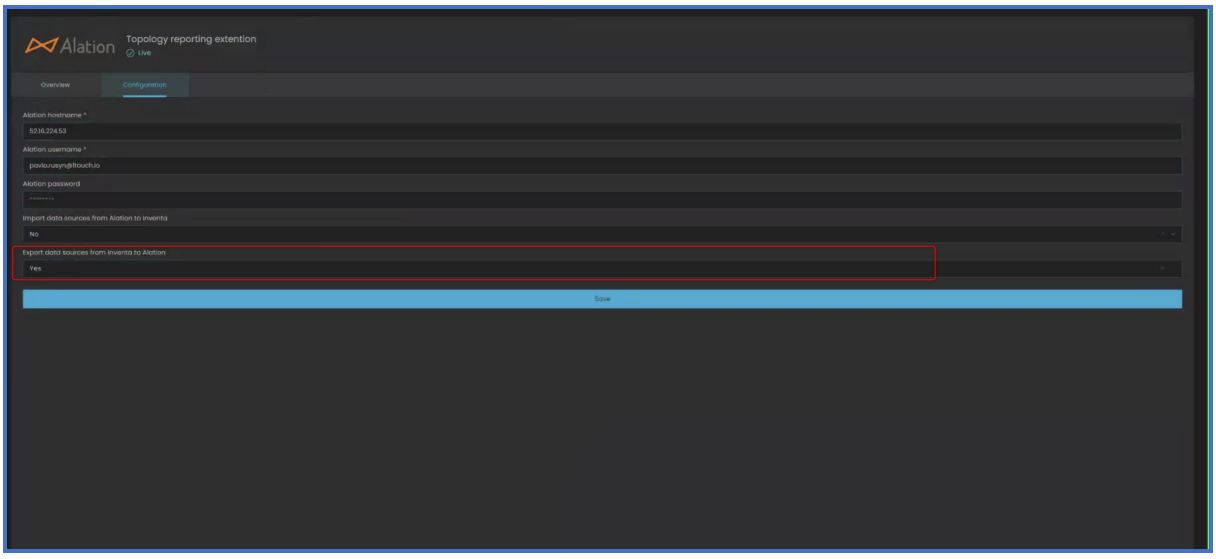

Figure 3: Activating export of data sources in IGDC

2. In Alation, go to **Admin Settings > Customize Catalog**.

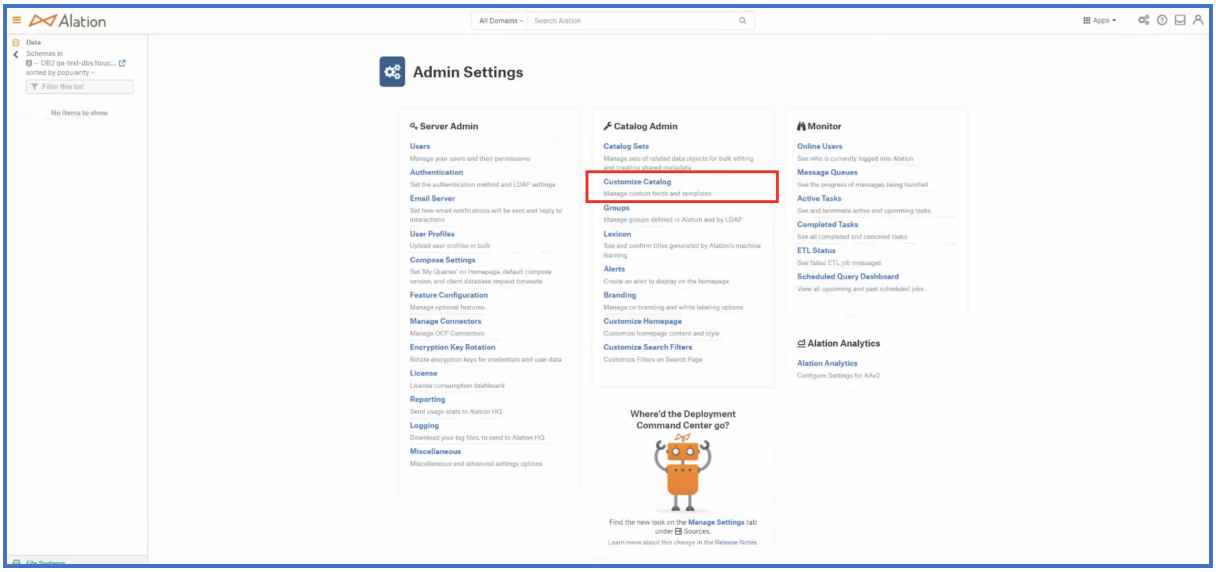

Figure 4: Customizing catalog in Alation

3. In **Customize Catalog**, select **Custom Templates**.

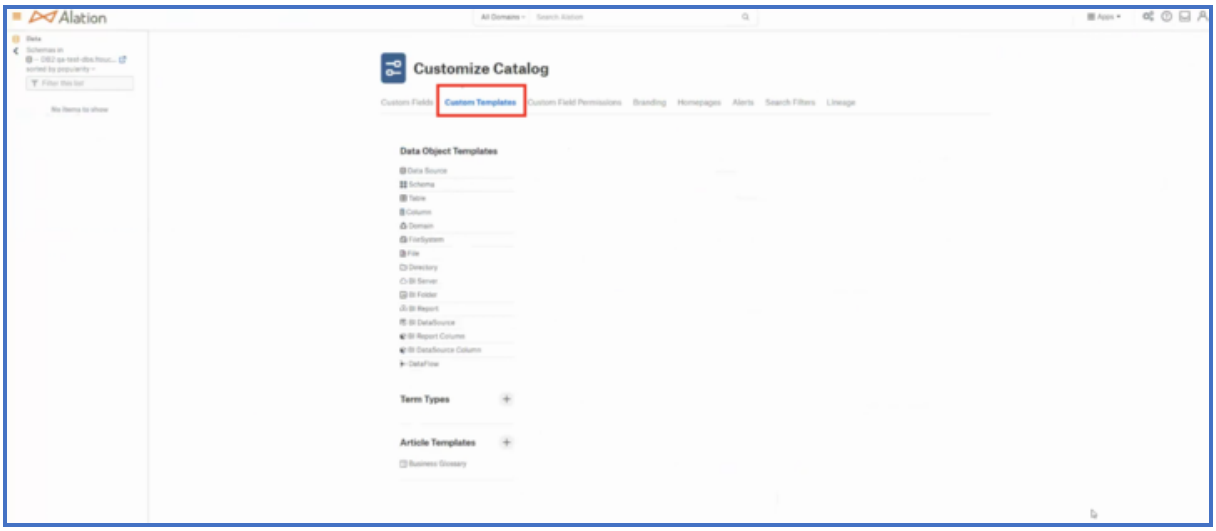

Figure 5: Custom templates in Alation

4. From **Custom Templates**, select **Data sources**. Under **Uploaded Queries**, click **Insert** and select **Custom field** from the dropdown.

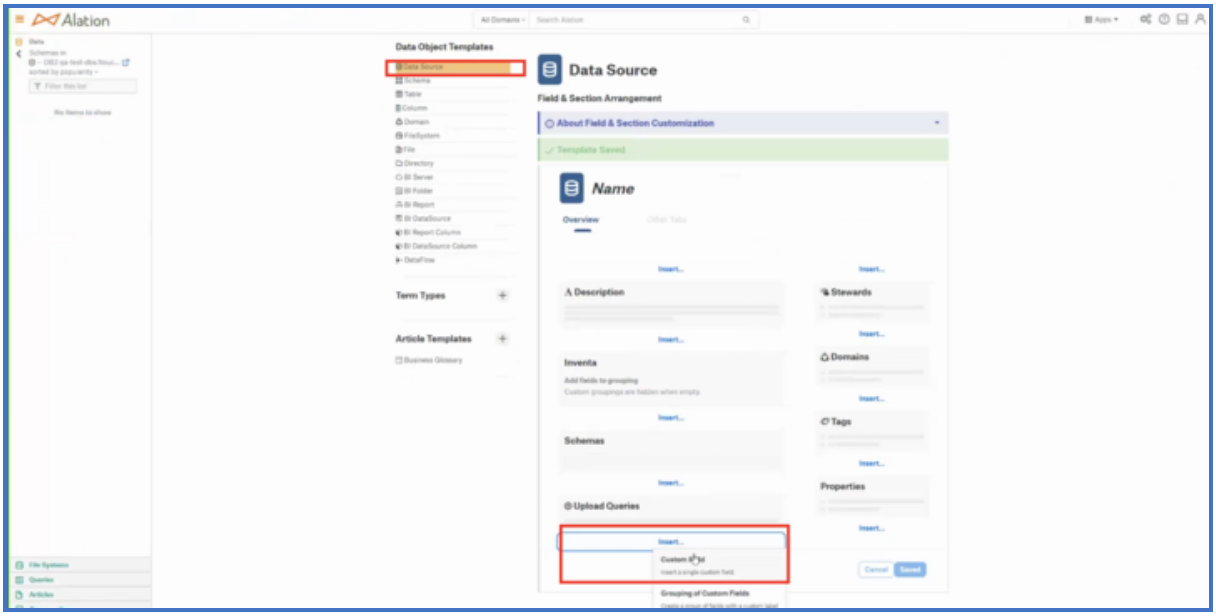

Figure 6: Adding a custom field for data source

5. For the data source, select the **ds_id** and **Data Elements** custom field.

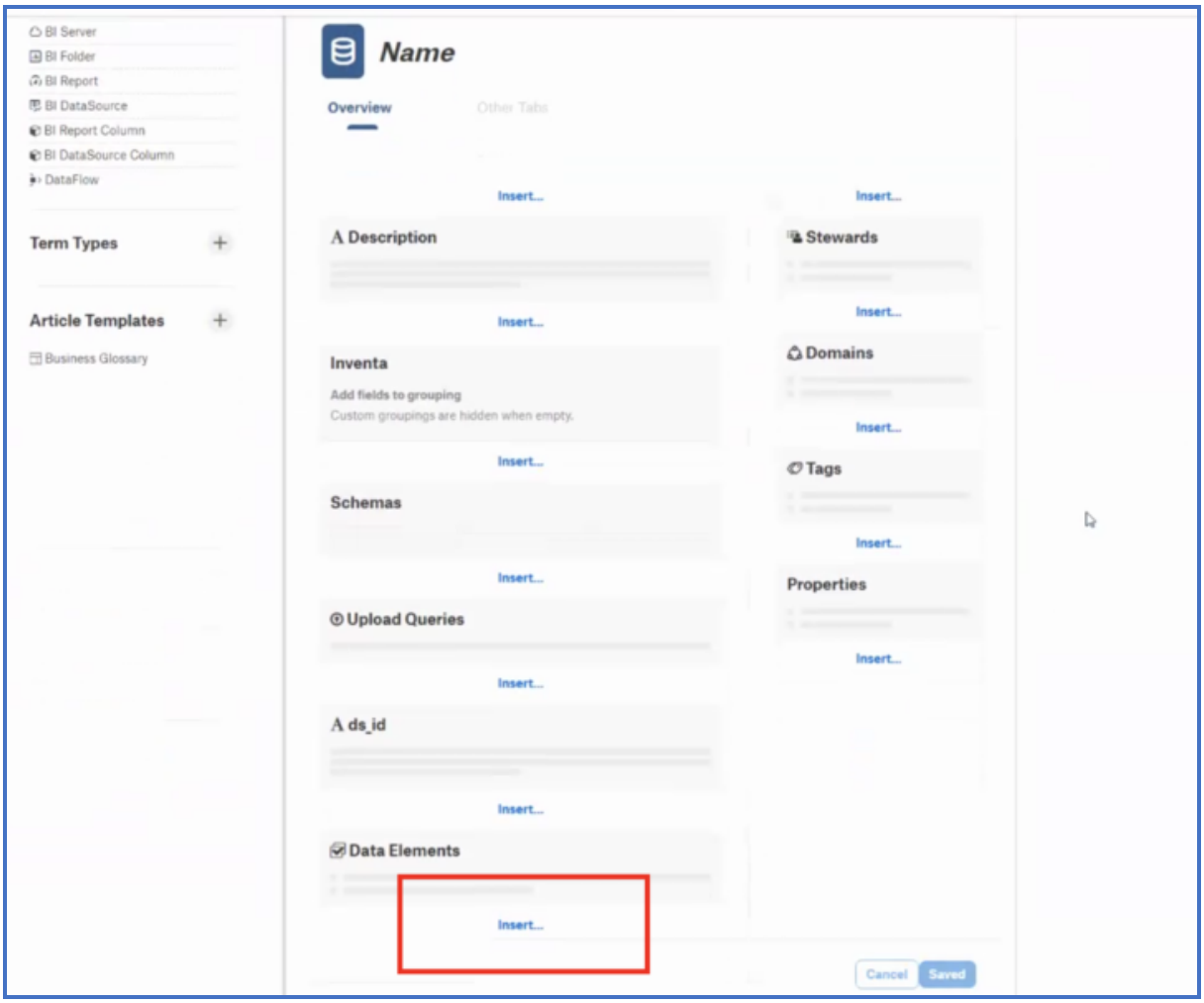

Figure 7: Ds_id and Data Elements custom fields in the data source custom template

6. Using the sidebar, go to the Schemas, Tables and Columns custom templates. For each template, add the **Data Elements** custom field as described in steps 4-5. The **ds_id** filed is not required.

7. Return to **IGDC platform > Ecosystem > Alation > Configuration tab** and restart export from IGDC to Alation: in the **Export data sources from Inventa to Alation** dropdown, select **No** and then again **Yes**.

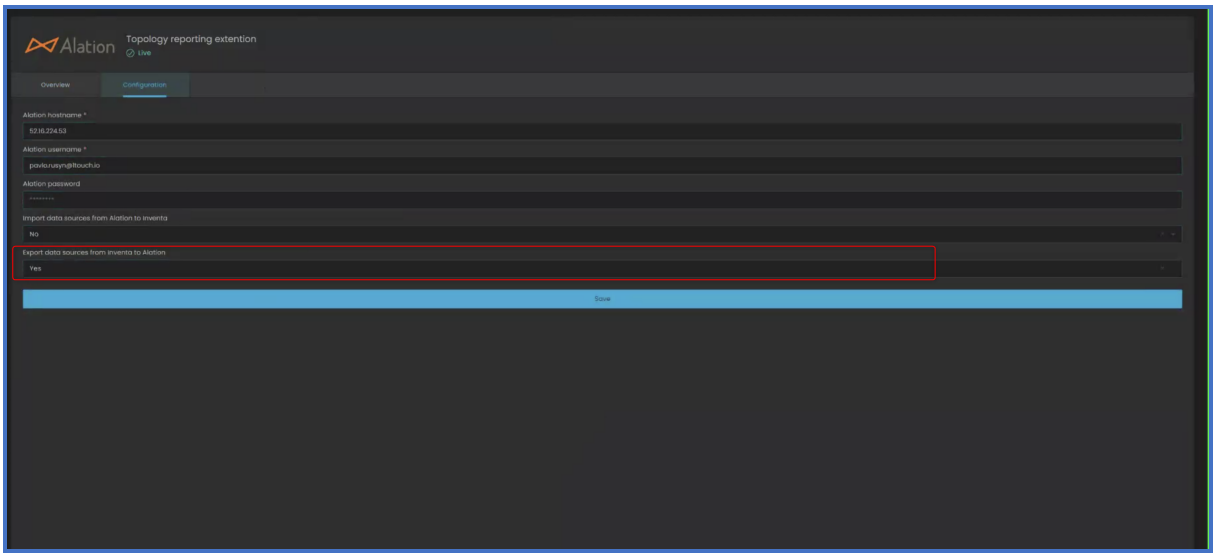

Figure 8: Activating export of data sources in IGDC

ALATION APP USAGE

USE CASES

Designed to enable seamless data exchange and communication between Alation and IGDC, this integration allows for the importation of data sources from Alation. IGDC then classifies these data sources and identifies new data resources, which are subsequently added back into Alation. This cycle of import, classification, and export enhances data governance accuracy while reducing the manual overhead, ensuring all data sources are managed in accordance with their sensitivity and compliance requirements.

ENRICHMENT OF ALATION DATA CATALOG

Automatically export data sources from the IGDC data source catalog and send them as assets to the Alation data catalog.

For example, IGDC has 100 data sources in the data source catalog (databases, data lakes, file shares, etc.). A user sets **Export data sources from Inventa to Alation=True** in **Ecosystem > Alation Topology Reporting Extension > Configuration** tab.

The integration app exports 100 data sources and their metadata (URL, id) from the IGDC data source catalog. Then they appear as assets in the Alation data catalog.

Once a user manually adds a data source to the IGDC catalog, or IGDC automatically discovers a data source, the integration app will send this data source to the Alation data catalog.

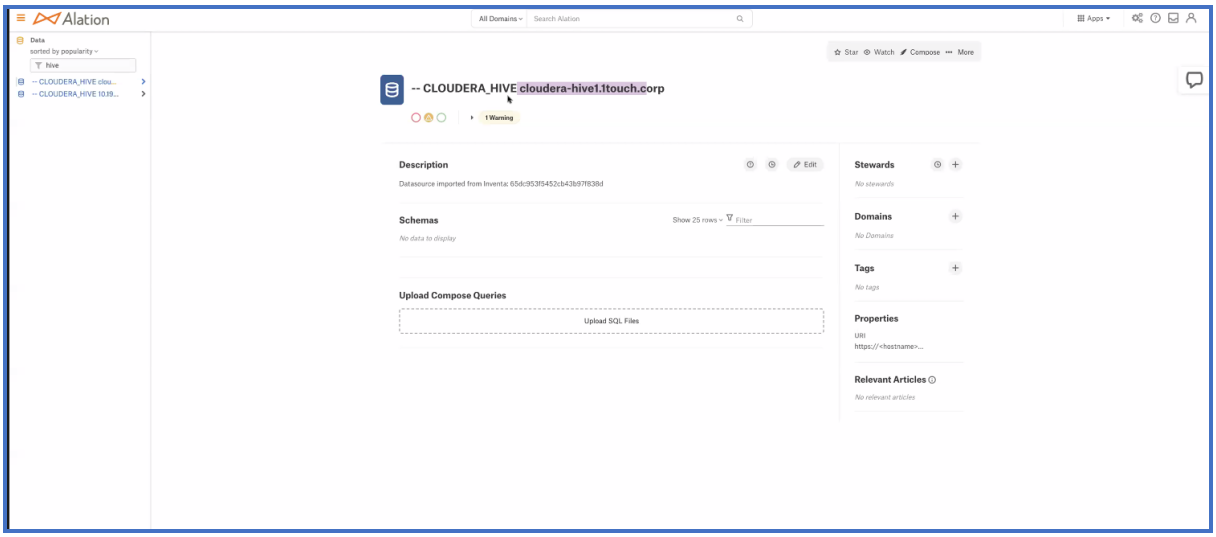

Figure 1: Asset exported from IGDC to the Alation data catalog

ENRICHMENT OF IGDC DATA SOURCE CATALOG

Automatically import data sources from the Alation data catalog to the IGDC data source catalog.

For example, Alation has 20 assets related to products supported by IGDC in the Alation data catalog. In IGDC, a user sets **Import data sources from Alation to Inventa =True** in **Ecosystem > Alation Topology Reporting Extension > Configuration** tab.

The integration app exports 20 assets from the Alation data catalog. Then they appear as data sources in the IGDC data source catalog. The **Created by** attribute for the imported data sources is **Alation - OCF (Automatically)**.

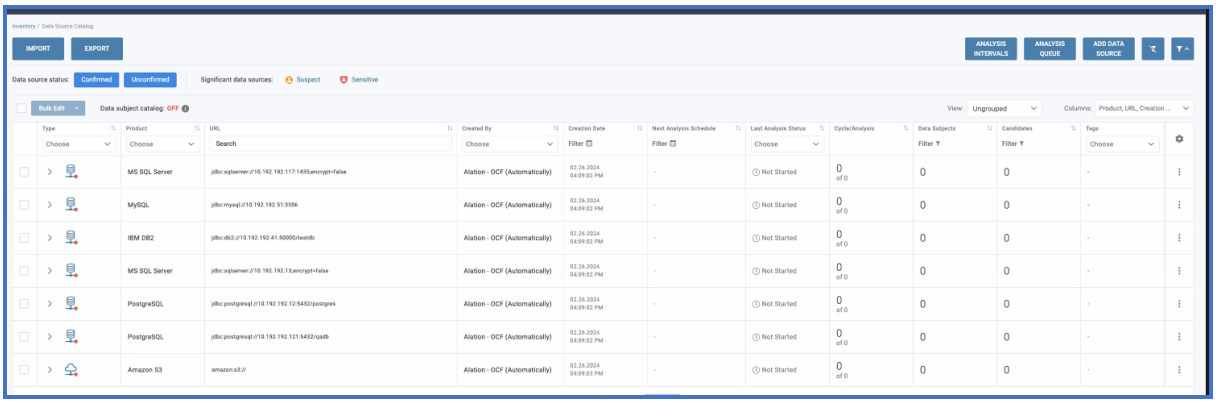

Figure 2: Data sources imported from the Alation data catalog to IGDC