

# SIEMENS

Insights Hub

Insights Hub Data Contextualization

System Manual


04/2024


<u>Introduction</u>	<b>1</b>
<u>User Interface and User Rights</u>	<b>2</b>
<u>Data Registry Management</u>	<b>3</b>
<u>Query Management</u>	<b>4</b>
<u>Monitoring Jobs</u>	<b>5</b>
<u>Using Ontologies</u>	<b>6</b>


## Legal information

### Warning notice system

This manual contains notices you have to observe in order to ensure your personal safety, as well as to prevent damage to property. The notices referring to your personal safety are highlighted in the manual by a safety alert symbol, notices referring only to property damage have no safety alert symbol. These notices shown below are graded according to the degree of danger.

 <b>DANGER</b>
indicates that death or severe personal injury <b>will</b> result if proper precautions are not taken.

 <b>WARNING</b>
indicates that death or severe personal injury <b>may</b> result if proper precautions are not taken.

 <b>CAUTION</b>
indicates that minor personal injury can result if proper precautions are not taken.

<b>NOTICE</b>
indicates that property damage can result if proper precautions are not taken.


If more than one degree of danger is present, the warning notice representing the highest degree of danger will be used. A notice warning of injury to persons with a safety alert symbol may also include a warning relating to property damage.

### Qualified Personnel

The product/system described in this documentation may be operated only by **personnel qualified** for the specific task in accordance with the relevant documentation, in particular its warning notices and safety instructions. Qualified personnel are those who, based on their training and experience, are capable of identifying risks and avoiding potential hazards when working with these products/systems.

### Proper use of Siemens products

Note the following:

 <b>WARNING</b>
Siemens products may only be used for the applications described in the catalog and in the relevant technical documentation. If products and components from other manufacturers are used, these must be recommended or approved by Siemens. Proper transport, storage, installation, assembly, commissioning, operation and maintenance are required to ensure that the products operate safely and without any problems. The permissible ambient conditions must be complied with. The information in the relevant documentation must be observed.

### Trademarks

All names identified by ® are registered trademarks of Siemens AG. The remaining trademarks in this publication may be trademarks whose use by third parties for their own purposes could violate the rights of the owner.

### Disclaimer of Liability

We have reviewed the contents of this publication to ensure consistency with the hardware and software described. Since variance cannot be precluded entirely, we cannot guarantee full consistency. However, the information in this publication is reviewed regularly and any necessary corrections are included in subsequent editions.

# Table of contents

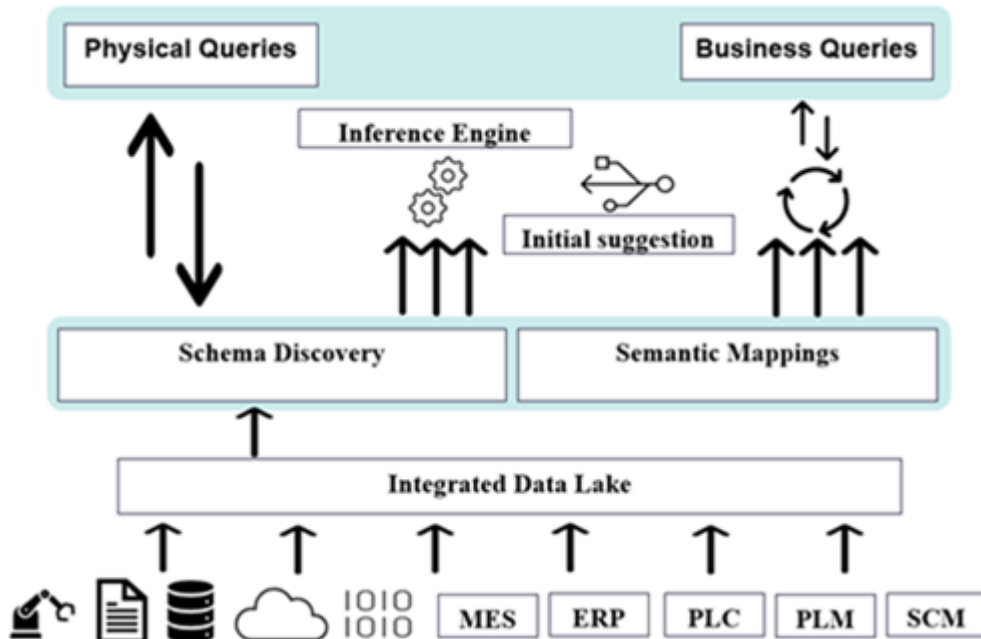
<b>1. Introduction</b> .....	4
1.1. Introduction.....	4
<b>2. User Interface and User Rights</b> .....	6
2.1. User Interface.....	6
<b>3. Data Registry Management</b> .....	10
3.1. Overview of Data Registry.....	10
3.2. Creating the Data Registry.....	10
3.3. Uploading the files to Data Registry.....	13
3.4. Deleting the Data Registry.....	16
<b>4. Query Management</b> .....	17
4.1. Creating a Query.....	17
4.2. Executing the Query.....	20
4.3. Deleting the Query.....	21
<b>5. Monitoring Jobs</b> .....	23
5.1. File Ingestion Jobs.....	23
5.2. Query Execution Job.....	24
<b>6. Using Ontologies</b> .....	25
6.1. Inferring the Ontology.....	25
6.2. Creating a new Ontology.....	28

# Introduction

## 1.1 Introduction

Data Contextualization allows the user to correlate data from multiple sources and generate powerful insights that would not be possible if the data were stored in separate apps. The workflow of Data Contextualization consists of the following steps:

1. In the Data Contextualization application, create one or more data sources by constructing Data Registries.
2. Add files to Integrated Data Lake, which leads to schema generation for the registry.
3. Create SQL query to correlate data from one or more data sources.
4. Execute the query and review the query results. The [Business Intelligence](#) application in Insights Hub is used to produce dashboards based on query results in Data Contextualization.



Data Contextualization also allows the user to create a semantic model as an abstraction layer over one or more physical schemas.

A domain expert often creates the semantic model. The user can select multiple configurations to combine the physical tables based on their shared properties as part of the semantic model definition which provides an inference feature using which the user can infer one or more registries or data sources. This also provides an initial suggestion which can be modified while building a semantic model.

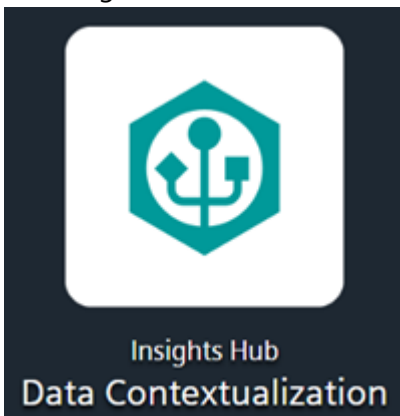
A domain expert creates the semantic model. The user can select multiple configurations to combine the physical tables based on their shared properties. In addition, the user can define domain-specific business attributes and relate them to physical schema attributes. The semantic model once generated can be used to write semantic queries using SQL statement.

# User Interface and User Rights

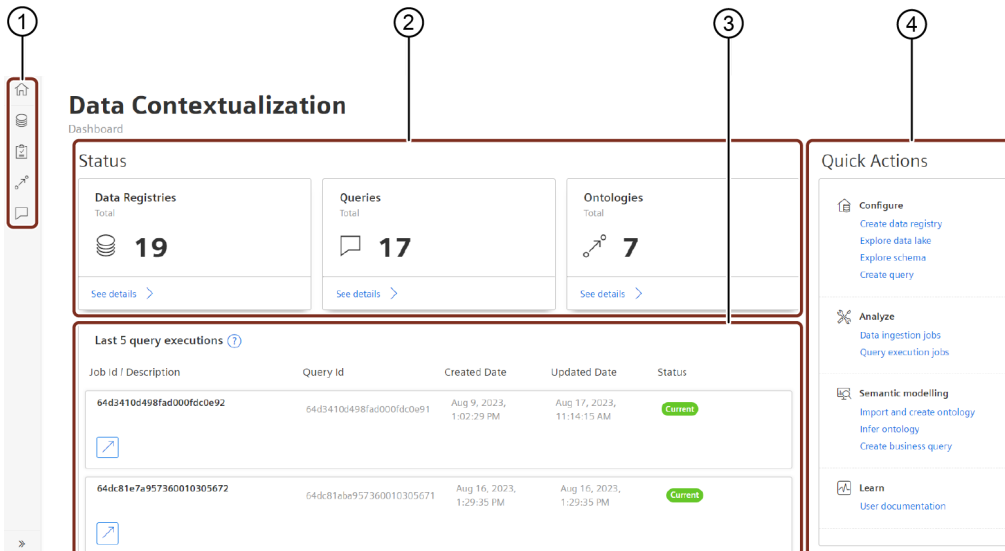
# 2

## 2.1 User Interface

Data Contextualization application can be accessed from Insights Hub Launchpad by clicking the following icon:



The Home screen of Insights Hub Data Contextualization is as shown below:



- ① Navigation pane
- ② Data objects
- ③ Latest query executions
- ④ Quick Actions

## Navigation tabs

Navigation tabs are available on the left side of the Home screen, as shown below:



- ① Home page
- ② Data Registries
- ③ Jobs
- ④ Ontologies
- ⑤ Queries

## Data Registries

From the left navigation, click "Data Registries". The following screen is displayed:

Type	Source Name	Data Log	File Type
dc	ktts	PARQUET	
dc	kt	PARQUET	
dc	pcb	PARQUET	
misp	pcb2	CSV	
misp	kt2	CSV	
tp	kt	CSV	
tp	pcb	CSV	
edi	zcopy3	PARQUET	
edi	zcopy	PARQUET	
misp	tpc	CSV	

- ① Navigation pane
- ② Search the files with object name
- ③ Displays the object view details
- ④ Creates a registry
- ⑤ Refresh button
- ⑥ Sorts the display results.

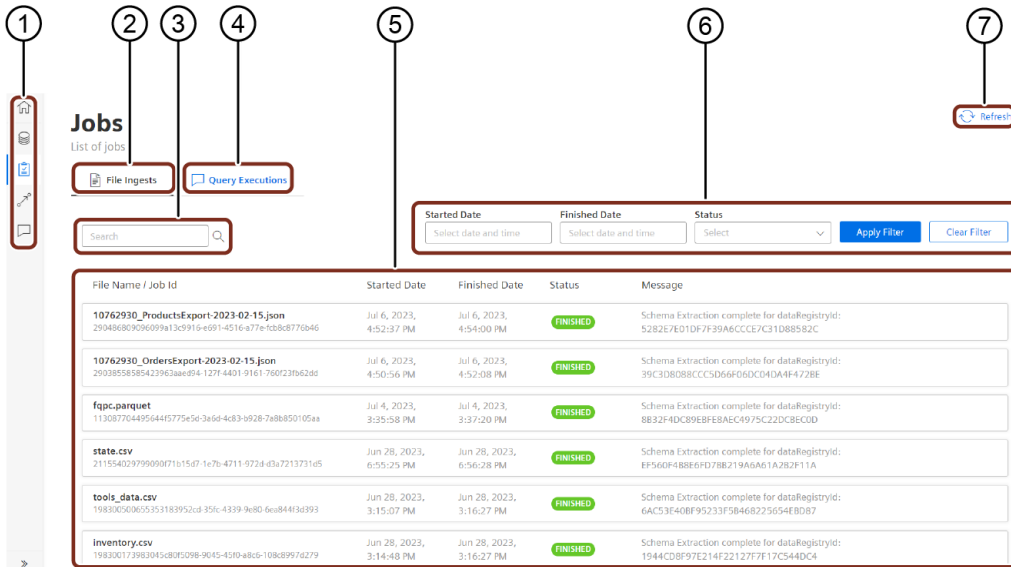
The following options are available: Created date,

- ⑦ General description
- ⑧ Schema details

2.1 User Interface

### Jobs

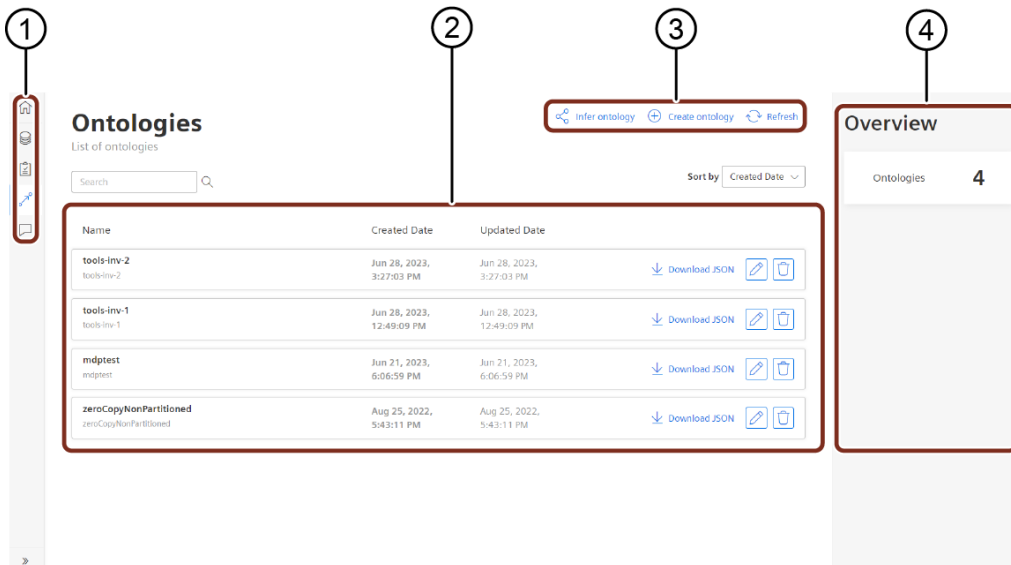
From the left navigation, click "Jobs". The following screen is displayed:



- ① Navigation pane
- ② File Ingests
- ③ Search bar
- ④ Query Executions
- ⑤ Displays the list of Jobs ingested
- ⑥ Filter bar
- ⑦ Refresh button

### Ontologies

From the left navigation, click "Ontologies". The following screen is displayed:

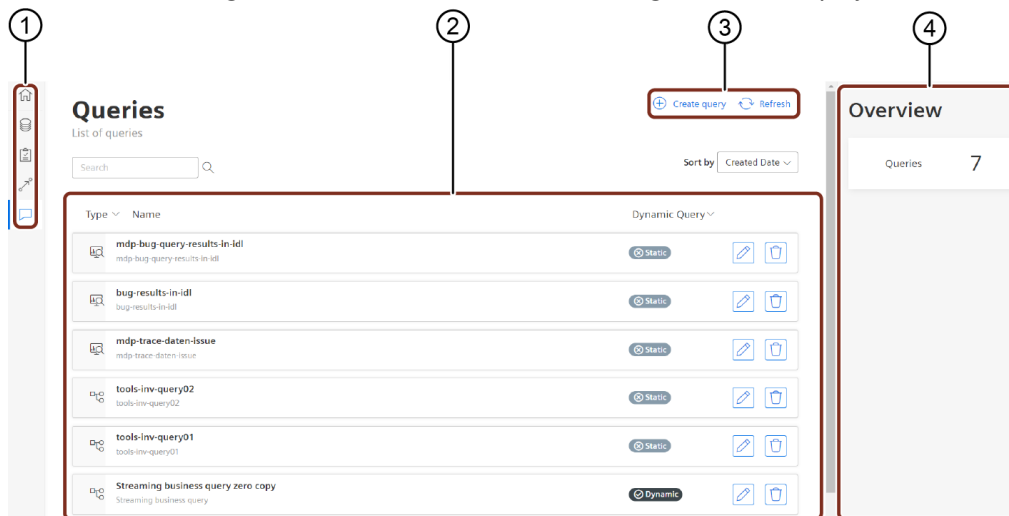


- ① Navigation pane
- ② Work area
- ③ Buttons to infer ontology, create ontology and refresh the results
- ④ Overview screen



## Queries

From the left navigation, click "Queries". The following screen is displayed:



- ① Navigation pane
- ② Work area
- ③ Buttons to create a query and refresh the results
- ④ Overview screen

## User Rights

The Data Contextualization application, the user rights are assigned from the Settings application, by adding the following user roles:

- **sdiui(mdsp:core) admin:** This role entitles all read or write access and administrative access to Data Contextualization user interface.
- **sdiui(mdsp:core) enduser:** This role entitles end-user query access to Data Contextualization user interface.

A user having any one of these roles can access all the features of the Data Contextualization user interface.

# Data Registry Management

# 3

## 3.1 Overview of Data Registry

Data Contextualization application in Insights Hub makes it possible to contextualize data across different data sources and derive meaningful insights from the data.

Data Registry feature enables the user to register a data source. For more information on how to create a Data Registry, refer [Creating a Data Registry](#).

Data Registries are classified into two types:

**Enterprise Data Registry:** Enterprise Data Registry is used to represent the unstructured enterprise data. Enterprise Data Registry enables a user to create a registry and initiate the schema discovery process by uploading files in the Integrated Data Lake. The user can tag the files in Integrated Data Lake to indicate the associated files to a specific Data Registry. For more information on creating Enterprise Data Registry, refer the chapter

- [Creating Enterprise Data Registry](#). The process involved in enterprise Data Registries are:
  - [Enterprise Data Registry Creation](#)
  - [Enterprise Data Registry File upload](#)

**IoT Data Registry:** IoT Data Registry is used to represent an Industrial IoT data source. IoT Data Registry enables a user to create a registry and initiate the process by importing data into Integrated Data Lake. For more information on creating IoT Data Registry, refer the chapter [Creating IoT Data Registry](#). The process involved in enterprise Data Registries

- are:
  - [IoT Data Registry Creation](#)
  - [IoT Data Registry File upload](#)

## 3.2 Creating the Data Registry

Creation of a Data Registry is the first step in the Data Contextualization workflow. For types of Data Registries supported in Data Contextualization, refer [Overview of Data Registry](#).


The below section explains the steps to create a Data Registry ([#overview-of-data-registry](#)) and [IoT Data Registry](#)) from Data Contextualization application.

## Creating Data Registry

### Creating Enterprise Data Registry

To create an Enterprise Data Registry, follow these steps:

1. In "Data Contextualization", click "Data Registries" from the left navigation.
2. In Data Registries screen, click "Create Registry".
3. Select "Enterprise data" option and enter the "Source name" and "Data tag".
4. Select file "Upload strategy", "Allowed file type" and "File pattern".

Enable "Use registry source path" toggle button to enter the path for creating a new folder in Integrated Data Lake. For example, /dc/jsontest. By using the , it is possible to

5. navigate to the created folder directly and upload the necessary files.

6. Click "Confirm".

### Create Data Registry ✕

**Category**

**Enterprise data**

Source name \*  Data tag \*

**IOT data**

**Files**

Define how new files will be handled (upload strategy can only be defined upon registry creation)

**Upload strategy \***

Append  
 Replace

**Allowed file type \***

CSV  
 XML  
 JSON  
 Parquet

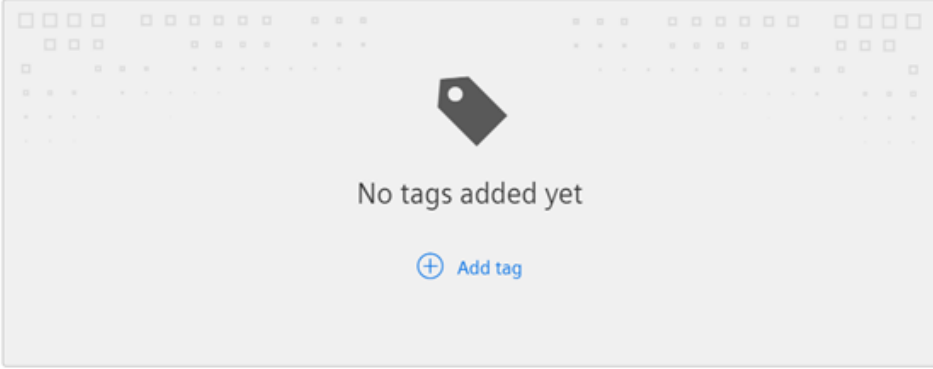
**File pattern \***

**Use registry source path**

**Location of the source files \***

**Partition Key**

**Registry metadata tags**

  
No tags added yet  
[+ Add tag](#)

It is also possible to create a registry without enabling the "Use registry source path" toggle button. In such cases, the user needs to upload files to the 'sdi' folder in Integrated Data Lake.

For more information, refer [Uploading the files to Data Registry](#).

## Creating IoT Data Registry

To create an IoT Data Registry, follow these steps:

1. In "Data Contextualization", click "Data Registries" from the left navigation.
2. Click "IoT data" checkbox on registry creation UI.
3. Select the required Asset and Aspect from the list displayed.
4. Click "Confirm".

The screenshot shows the "Create Data Registry" dialog box. It has a title bar with "Create Data Registry" and a close button. Below the title bar, there are two radio buttons: "Enterprise data" (unselected) and "IOT data" (selected). Under the "IOT data" section, there are two search fields. The "Asset" field has a dropdown menu showing "> mn-test". The "Aspect" field has a dropdown menu showing "contactPoint". At the bottom, there is a blue information box with a downward arrow icon and text: "Select an asset and an aspect to create an IOT registry. The schema discovery for the IOT registry can be started by importing the timeseries data in the TSI/sdi folder of the data lake." At the very bottom, there are "Cancel" and "Confirm" buttons.

## Result

The Data Registry is successfully created.

When an Enterprise Data Registry is created, the user can upload files in the Integrated Data Lake with appropriate tags to initiate the schema discovery process for the registry.

When an IoT Data Registry is created, the user can import timeseries data in the folder "TSI/sdi" of the Integrated Data Lake application to initiate the schema discovery process.

For more information, refer [Uploading Files to Data Registry](#).

## 3.3 Uploading the files to Data Registry

The files are uploaded to the Data Registry from the Data Explorer in Integrated Data Lake application. The files uploaded could be in CSV, JSON, PARQUET, or XML format.

When the file is uploaded in Integrated Data Lake application, a notification is triggered in Data Contextualization. Data Contextualization uses the folder metadata tag to determine the registry for the file and the application then proceeds to ingest the file into the designated registry and initiates the schema discovery process.



## Uploading the files for Enterprise Data Registry

In Enterprise Data Registry, the user can initiate the schema discovery process by uploading files in Integrated Data Lake after creating the registry. The user needs to tag the files in the Integrated Data Lake to indicate association of the files to a specific Data Registry. The user can use following procedure to tag and upload the files in the Integrated Data Lake.




For Enterprise Data Registry, the folder must be created in the "sdi" folder of the Integrated Data Lake. User can create different folders for different registries.


## Uploading a file in Integrated Data Lake

To upload the file into Integrated Data Lake, follow these steps:

In the Integrated Data Lake, navigate to the 'Data Explorer' window and click "Create 1. Folder".

2. Create a folder inside the "sdi" folder.

3. Enter the folder name and click  to save the folder.

4. Click  to add the metadata tags for the created folder.

5. Select "Metadata Tags", click "Add Tag".

Enter metadata tag to the folder in the following format, `registryId_ABCD` where 6. "ABCD" is the "registryId".

7. Click "Save".

**Data Explorer** Copy Path Back

Folder View

[sdi](#) / [test](#) / test3

**Overview** **Metadata Tags**

**Folder tags** Refresh View Add Tag

Add/Edit/Delete metadata tags here for easy searching and referring objects later.

- Maximum 8 tags are allowed.
- Maximum 128 characters are allowed within a single metadata tag.

Name	Action
registryid_F11D8F3D53EF41FDD449C68545312AD3	<span>✓</span> <span>✕</span> <span>🗑️</span>

**Save** **Cancel**

8. In the Upload Objects, click "Add Files" and select the file to be uploaded.

**Upload Objects** ✕

Upload path:  
sdi/ 📄

+ **Add Files**

Object	Type	Size
inventory.csv	CSV	174.00 bytes <span>⊖</span>

**Cancel** **Upload**

9. Click "Upload".




- If a user tags a folder with a registryId, all files copied in that folder will be ingested in the same registry.
- A user can optionally add a metadata tag on a file. The metadata must be in the same format as the folder metadata. When both file metadata and (parent) folder metadata are present, Data Contextualization gives preference to the file metadata. So, file metadata can be used to override the registryId specified on the folder metadata.

## Result

Uploading the file is successful and the upload status changes to successfully uploaded.

Object	Progress	Size
sdi/inventory.csv	Successfully uploaded	174.00 bytes

In Data Explorer window, click  "Refresh View" button to view the uploaded files in the table.

## Uploading a file for IoT Data Registry

After creating a registry in the IoT Data Registry, user needs to import the IoT data to Integrated Data Lake. To import the IoT data into Integrated Data Lake, refer [Importing data into Integrated Data Lake](#).




For IoT Data Registry, the folder must be created in the "sdi" folder of the Integrated Data Lake. User can create different folders for different registries.

## 3.4 Deleting the Data Registry

Data Contextualization application enables the user to delete a Data Registry.

### Procedure

To delete the data registry, follow these steps:

1. In Data Registry, select the registry which needs to be deleted.
2. Click .
3. Click "Yes, delete registry".



Delete registry

Are you sure you want to delete the registry?

Yes, delete registry

No, cancel

### Result

The data registry is successfully deleted.



## 4.1 Creating a Query

A Query is a request for specific data, executed through commands such as SQL statements, which enables the manipulation of data with Integrated Data Lake. Each time a Query is created, an associated job is automatically generated. For more information on jobs, refer [File Ingestion Jobs](#) or [Query execution jobs](#).



To perform SQL-like operations, Data Contextualization makes use of the spark.sql module.

### Query details

The following types of Queries are supported:

**Physical Query:** The Query created using schemas and its properties present in registry is called a physical Query. The user is allowed to form a valid SQL statement using schema

- and its properties.

**Business Query:** The Query created using classes and its properties present in ontology is called a business Query. The user is allowed to form a valid SQL statement using class or classes and its properties. Business Queries use the underlying mappings provided in the

- semantic models to Query data from multiple sources of data.

Queries (both physical and business) in Data Contextualization can be Static or Dynamic in type.

**Static Query:** A Static Query contains parameters of which the value of parameters are defined at the time of creation. The value of any parameter cannot be modified at the

- time of execution.

**Dynamic Query:** A Dynamic Query contains parameters of which value can provided at the time of execution for column names in the selected clause. Dynamic queries help


- users to create multiple execution requests on a single Query.

For more information on Queries, refer [Creating Native SQL based query](#).

### Procedure

To create a Query, follow these steps:

1. From the left navigation, click "Queries" and then click "Create Query".
2. Select the Query type.
3. Enter the general details and SQL Statement.

Enable the toggle option  to select Dynamic Query, otherwise Static Query is used by default.

5. Click "Confirm".

### Create Query >

---

#### General

**Query type \***

Physical query  
Uses physical schema

Business query  
Uses semantic model

**Name \***

**Description**

#### SQL Statement

Dynamic Query

**Statement \***

Save results as a file

---

## Result

A Query is successfully created.



## 4.2 Executing the Query

Data Contextualization Query Service creates [Job executions](#) for Static Queries and Dynamic Queries. The Query execution process runs a Query plan and returns the results to the Query.

### Executing a Static Query

A Static Query statement includes all of the information required to run the Query. When a user creates a static query in the Data Contextualization application, the query is executed automatically. To view the status of Static Query Execution, refer [Query Execution Jobs](#).


### Executing a Dynamic Query

A Dynamic Query has one or more parameters, the values of which must be provided before the Query Execution. After executing a Dynamic Query, the user can see the status of the execution in the "Jobs" user interface. The below section explains the steps to execute a Dynamic Query.

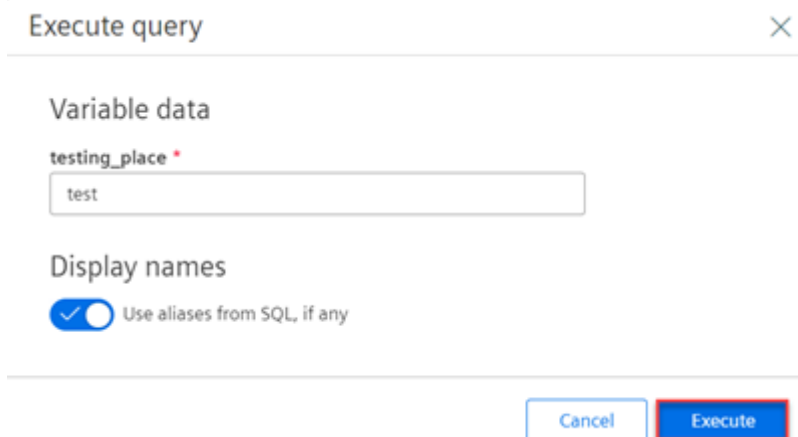
#### Procedure

To execute a Dynamic Query, follow these steps:

1. In Data Contextualization, click "Queries" from the left navigation.

Select the required Dynamic Query, click  under "Execution(s)" to view the execution page.

3. Click "Execute" and enter the Variable data.



Execute query ×

Variable data

testing\_place \*

Display names

Use aliases from SQL, if any

Cancel Execute

4. To rename the component number, enable the toggle button .

5. Click "Execute".

Execute query
✕

---

**Variable data**

testing\_place \*

**Display names**

Use aliases from SQL, if any

component\_number

Cancel
Execute

## Result

A Query is successfully executed.

**dynamic query test**

Execution(s)

- Aug 25, 2023, 12:40:14 PM ✓
- Aug 16, 2023, 1:29:35 PM ✓

SQL statement

```
select component_number from sdi_testB
where sdi_testB.testing_place like
?'testing_place'
```

**Execution from Aug 25, 2023, 12:40:14 PM**

component_number
C00
C01
C02
C03
C04
C10
C11
C12
C13
C14
C20
C21
C22
C23
C24

## 4.3 Deleting the Query

Data Contextualization enables the user to delete the folder or object in Queries.

### Procedure

To delete the folder or object, follow these steps:

1. In Queries, select the folder or object that needs to be deleted.
2. Click .

3. Click "Yes, delete query".



### Delete Query

Are you sure you want to delete the query?

---

Yes, delete query

No, cancel

## Result

The Query is successfully deleted.

# Monitoring Jobs

## 5.1 File Ingestion Jobs

A file ingestion job gets created when a file is uploaded to Integrated Data Lake with an appropriate metadata tag. This job can be used to check the status of all the file ingestion jobs for the tenant.

A user can upload one or more files in Integrated Data Lake with appropriate registry tags after creating a Data Registry in Data Contextualization. This initiates the schema discovery process in Data Contextualization. For more information, refer the chapter [Uploading files to data Registries](#).

To track the progress of the schema discovery process, a single file ingestion job is generated in Data Contextualization for each uploaded file in the Integrated Data Lake (with the relevant registry tag).

A file can be filtered by entering the file name through search bar or by providing the input details in Started Date, Finished Date or Status fields.

**Jobs**  
List of jobs

File Ingests Query Executions

Search

Started Date: Select date and time Finished Date: Select date and time Status: Select

Apply Filter Clear Filter

File Name / Job Id	Started Date	Finished Date	Status	Message
10762930_ProductsExport-2023-02-15.json 29088880905099a13c916e691-4516a77e6c8c8c8776e46	Jul 6, 2023, 4:52:37 PM	Jul 6, 2023, 4:54:00 PM	FINISHED	Schema Extraction complete for dataRegistryId: 5282E7E01DF7F39A6CCE7C31D88582C
state.csv 312867070466579a168801-22574ae4-af54-698c2adc25f	Jun 19, 2023, 4:50:03 PM	Jun 19, 2023, 4:51:54 PM	FINISHED	Schema Extraction complete for dataRegistryId: 876D0663D858B78310FA24ACEA3A598
state.csv 30862858390236f950c22-7d33-459c-bbba-33920cf0b1a4	Jun 19, 2023, 4:42:59 PM	Jun 19, 2023, 4:44:46 PM	FINISHED	Schema Extraction complete for dataRegistryId: 3706DA54502DFAC46BB4F506588C432
state.csv 303194684164495493dbfb-r110-4979-9fd6-2662d8bc200	Jun 19, 2023, 4:33:56 PM	Jun 19, 2023, 4:35:03 PM	FINISHED	Schema Extraction complete for dataRegistryId: CEA73C5C44AD5A6A6092DF98426A29
state.csv 300982905590927ec1b4ae-00b4-4d6b-9a93-b2135863a5f7	Jun 19, 2023, 4:30:15 PM	Jun 19, 2023, 4:31:25 PM	FINISHED	Schema Extraction complete for dataRegistryId: 3185BC087501EA681B4E1B8128F11F38
state.csv 290739941028820826a8ed-2642-4c11-a5c2-119bc7cd1b4	Jun 19, 2023, 4:13:10 PM	Jun 19, 2023, 4:14:31 PM	FINISHED	Schema Extraction complete for dataRegistryId: CFE9A218CDFAE8286A500C2CB40F185D
state.csv 28601754475909937a7236-b330-4022-ec57-5dfcef6c6793	Jun 19, 2023, 4:05:18 PM	Jun 19, 2023, 4:06:32 PM	FINISHED	Schema Extraction complete for dataRegistryId: A6A03FA2AA625283F42EB7C76FAE4D24
state.csv 27625213692269899f8c0eae1-4501-98b4-ef697612f506	Jun 19, 2023, 3:49:03 PM	Jun 19, 2023, 3:51:08 PM	FINISHED	Schema Extraction complete for dataRegistryId: 36459EA027EF1FE0273901CCB17ECD3

Load more

- ① Search Bar
- ② File Name/ Job ID
- ③ Started Date for the job
- ④ Finished Date for the job
- ⑤ Status of the jobs

- ⑥ Message description
- ⑦ Filter bar
- ⑧ Refresh button

## 5.2 Query Execution Job

A Query Execution Job gets created when a Query is created. Whenever a Query is executed in Data Contextualization, the application internally creates a query execution job.

The progress of the Query can be tracked by checking the status of the Query Execution Job in the "Jobs" section of the user interface. The "Jobs" user interface is also useful for reviewing all tenant queries.

A Query can be filtered by entering the Query name through the search bar or by filtering based on Query id, Status.

The screenshot shows the 'Jobs' section of the user interface. It includes a search bar (1), a list of jobs with columns for Job id / Description (2), Query Id (3), Created Date (4), Updated Date (5), and Status (6). A filter bar (7) is located above the table, and a refresh button (8) is in the top right corner.

Job id / Description	Query Id	Created Date	Updated Date	Status
6307696cf14f64000f1ebe54 Running query with parameters	6307690c988ec8000fe7a0e4	Aug 25, 2022, 5:52:04 PM	Jul 12, 2023, 11:55:02 PM	Current
64a6b50ae436b2000f8af267	64a6b50ae436b2000f8af266	Jul 6, 2023, 6:05:22 PM	Jul 6, 2023, 6:05:22 PM	Failed
64a6b0bfe436b2000f8af265	64a6b0bfe436b2000f8af264	Jul 6, 2023, 5:47:03 PM	Jul 6, 2023, 5:47:03 PM	Failed
64a6a548e436b2000f8af263	64a6a548e436b2000f8af262	Jul 6, 2023, 4:58:08 PM	Jul 6, 2023, 4:58:08 PM	Current
649c0562d57f8d000f77772d	649c0562d57f8d000f77772c	Jun 28, 2023, 3:33:14 PM	Jun 28, 2023, 3:33:14 PM	Current
649bdfaad57f8d000f77772a	649bdfaad57f8d000f777729	Jun 28, 2023, 12:52:18 PM	Jun 28, 2023, 12:52:18 PM	Current
63076ab3f14f64000f1ebe65 Running query with parameters	6307690c988ec8000fe7a0e4	Aug 25, 2022, 5:57:31 PM	Aug 25, 2022, 5:57:31 PM	Outdated
630768ba988ec8000fe7a0e3	630768ba988ec8000fe7a0e2	Aug 25, 2022, 5:49:06 PM	Aug 25, 2022, 5:49:06 PM	Failed

- ① Search Bar
- ② Job id/ Description
- ③ Query id
- ④ Created Date of the query
- ⑤ Updated Date of the query
- ⑥ Status description
- ⑦ Filter use cases
- ⑧ Refresh button



## 6.1 Inferring the Ontology

An ontology or a semantic model in Data Contextualization represents a layer of abstraction on top of one or more physical schemas (each schema being a data source). The ontology or semantic model is created by a domain expert who understands the domain and is familiar with the domain terminology. After the semantic model is developed, other users can use it to write [Business Queries](#).

The expert user can specify a business class, one or more business attributes in the business class, and map each business attribute to one or more physical schema attributes while generating the semantic model or Ontology. While building the semantic model, the user can optionally choose a configuration that describes how the physical tables should be coupled. As a result, the user can utilize the semantic model as a layer of abstraction to specify domain or business-related features and join conditions.

To infer an ontology using the available schema includes the following four steps:

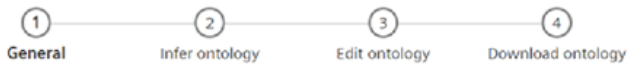
1. [General ontology](#)
2. [Infer ontology](#)
3. [Edit ontology](#)
4. [Download ontology](#)

### General ontology

In General ontology, follow these steps:

1. In Ontologies, click "Infer ontology".
2. Enter the required general details such as "Ontology Name" and "Description".

3. Click "Next".




Ontology Name \*


Description

Next >

## Infer ontology

In Infer ontology, follow these steps:

1. List the required Schemas to "Selected Schemas" from "Available Schemas" and select . Click "Infer Ontology".



Available Schemas

- dc\_kcts  
Created on October 5, 2023; Updated on October 5, 2023
- dc\_ict  
Created on October 5, 2023; Updated on October 5, 2023
- idi\_xcopy3  
Created on September 21, 2023; Updated on September 21, 2023
- idi\_xcopy  
Created on September 1, 2023; Updated on September 1, 2023
- mdp\_pcb  
Created on August 7, 2023; Updated on August 7, 2023
- mdp\_ict  
Created on August 7, 2023; Updated on August 7, 2023
- mdp\_fgpc  
Created on August 7, 2023; Updated on August 7, 2023
- trace\_data\_nsis\_products  
Created on July 6, 2023; Updated on July 6, 2023

Selected Schemas

- dc\_pcb  
Created on October 5, 2023
- mdp\_pcb2  
Created on September 26, 2023
- mdp\_ict2  
Created on September 26, 2023

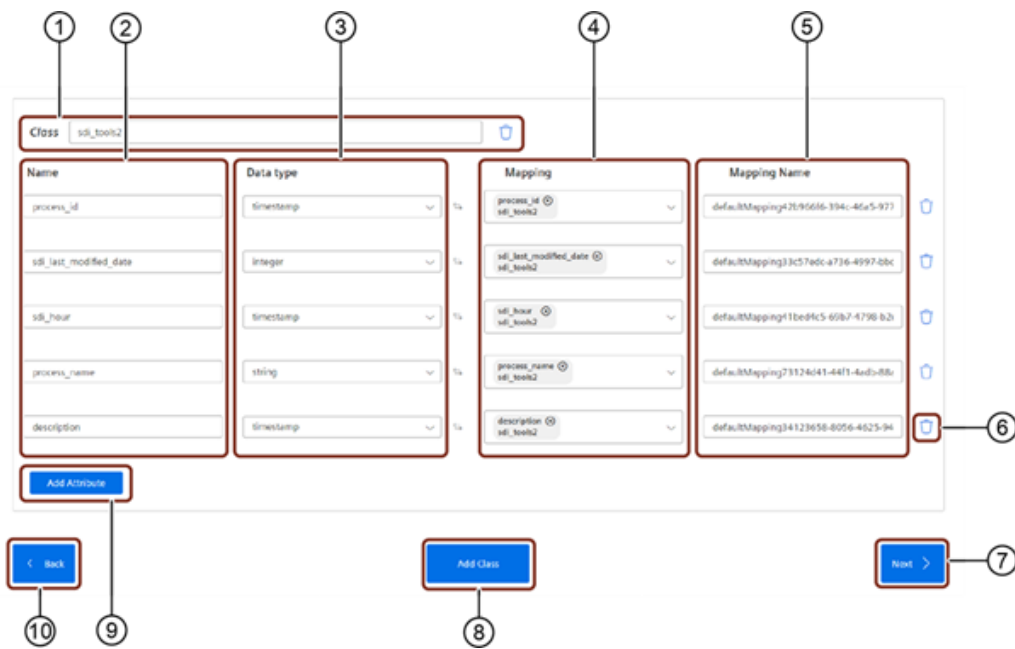
Infer Ontology

2.

A model can be created with the Class and required attributes which can be edited or deleted as per the requirement.

3. Click "Add Class" to add multiple classes.

4. Click "Next".



- ① Class
- ② Names of attributes
- ③ Data type
- ④ Mapping
- ⑤ Mapping Name
- ⑥ Delete button
- ⑦ Next button
- ⑧ Add Class
- ⑨ Add Attribute
- ⑩ Back button

## Edit ontology

In Edit ontology, a JSON file is generated. Click "Back" button to edit the required details.



- ① Back button
- ② Ontology Schema
- ③ Download ontology
- ④ Json view

## Download ontology

To Download ontology, click "Download ontology". A JSON file gets downloaded, thereby an Ontology is successfully downloaded.

## 6.2 Creating a new Ontology

A new ontology can be created to build a semantic model for Query Execution. An Ontology is created by choosing either of the key mapping type or uploading the [downloaded JSON ontology file](#).

### Key Mapping Configuration

Data Contextualization provisions a configuration to add the type of correlation between the mapped schema properties. This configuration will enable the users to retrieve commonality (Intersection) as well as aggregate data (Union) from the mapped datasets. Currently, it supports two key mapping datasets:

**Inner Join:** Inner Join is the work to get the commonality from the mapped schemas. If user do not provide keyMappingType at ontology & class level, then INNER JOIN will be

- considered by default to get data from the datasets.

**Full Outer Join:** Outer Join is the work to get all the aggregated data from the mapped schemas. This configuration can be provided at two levels named Ontology level and Class level. The key mapping type provided at class level will override the one provided at

- ontology level.

### Procedure

To create a new ontology, follow these steps:

1. In ontologies, click "Create ontology".
2. Enter the details for "Name", "Description" and "Key Mapping Type".
3. Select the [downloaded JSON file](#) from infer Ontology in "Ontology File".

4. Click "Create Ontology".

### Create Ontology ✕

**Name \***

**Description**

**Key Mapping Type \***  
 Inner Join  
 Full Outer Join

**Ontology File \***  
 No file chosen

## Result

A new Ontology is successfully created.