
Registration of a New Dataset

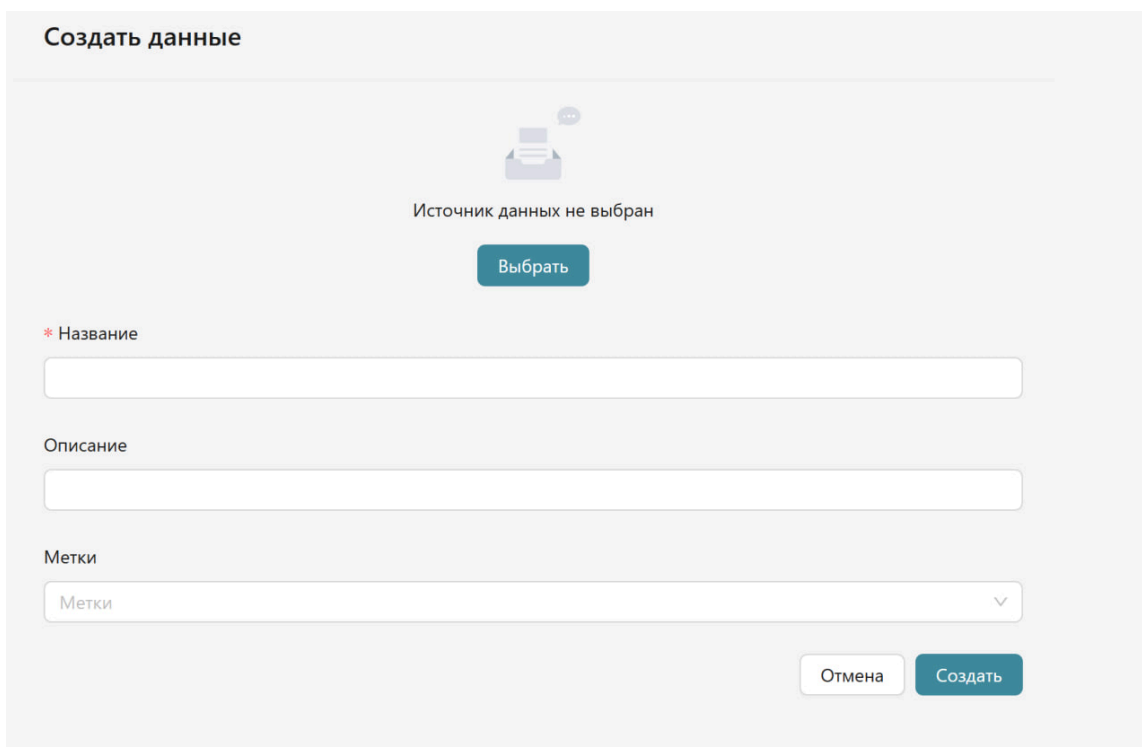
This section describes the process of registering a dataset from a source connected to the system. For more details about data sources, see the section "[Data Sources](#)".

Important

Currently, there is a **limit** on the dataset size - **1 GB**. The system has not been tested with files larger than this. The recommended file size is up to **100 MB**.

Navigation is done through the **Data** tab in the top navigation panel. In the opened **Data Catalog** window, click the "Add" button.

A form for registering a new dataset will appear:



Создать данные

Источник данных не выбран

Выбрать

* Название

Описание


Метки

Отмена Создать


First, you need to select a data source. By clicking the "Select" button, you will have the option to choose the desired source.

After selecting the source, the dataset creation form will update according to the type of the selected source. If an S3 type source is selected, the form will contain a field for uploading a file:

Создать данные


s3-prod 

Файл



Щёлкните или перетащите файл в эту область, чтобы загрузить

Поддерживает загрузку одного файла

* Скрипт для запроса данных 


Список таблиц Выполнить

Столбец, содержащий дату наблюдения

* Название

Описание

Метки

Метки 

Отмена Создать

Registration Parameters

- **Data Source** - must be selected from the list of available data sources.
- **File** - field for uploading a file to S3. Available only for S3 data sources. You can click or drag and drop the file into the specified area.
- **Data Query Script** (mandatory field) - a script in the SQL dialect that is accepted by the data source.

Tip for S3

For S3 type sources, a function to read the entire dataset will be automatically suggested, for example: `data/predictions.csv`

- **Column Containing the Observation Date** (optional field) - the name of the column that contains the reporting date for the data. This is selected after composing the script and clicking the "Execute" button.
- **Name** (mandatory field) - the name of the dataset.
- **Description** (optional field) - a description of the dataset.
- **Tags** (optional field) - tags for the dataset.








To view the list of available tables, you can use the "Table List" button.


After filling in all the necessary data, click the "Create" button. To cancel the operation, click "Cancel".

Selecting Existing Files in S3

If the data file has already been uploaded to the S3 storage, there is no need to upload it again. You can use existing files:

1. Click the "Table List" button.
2. In the appeared list, select the desired file:

Название	Скрипт для запроса данных
data/100mb30col_2024_11_21_15_53_46_...	data/100mb30col_2024_11_21_15_53_46_670.csv 
data/100mb30col_2024_11_21_15_54_38_...	data/100mb30col_2024_11_21_15_54_38_699.csv 
data/1800mb30col.csv	data/1800mb30col.csv 
data/1gb.csv	data/1gb.csv 
data/TEST_2024_9_24_11_57_59_678.csv	data/TEST_2024_9_24_11_57_59_678.csv 
data/TEST_2024_9_24_12_43_20_897.csv	data/TEST_2024_9_24_12_43_20_897.csv 
data/TEST_2024_9_24_13_16_44_767.csv	data/TEST_2024_9_24_13_16_44_767.csv 



- The **Name** column displays the file name in the S3 storage.
- The **Data Query Script** column specifies the path to the file for use in the SQL query. Copy it to later paste it into the main menu field **Data Query Script**.