Data Security Center

Best Practices

Issue	01
Date	2023-11-30

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 How Do I Prevent Personal Sensitive Data From Being Disclosed During Development and Testing?

Sensitive data refers to the data that may bring serious harm to the society or individuals after being leaked.

NOTE

For individuals, privacy information, such as ID card numbers, home addresses, workplace information, and bank card numbers, is sensitive data. For enterprises or organizations, core information, such as customer information, financial information, technical information, and major decisions, is sensitive data.

Huawei Cloud Data Security Center (DSC) can perform **static data masking** on a large amount of data in one operation based on anonymization rules. Static anonymization is usually used when sensitive data in the production environment needs to be transferred to the development, test, or outside environment. It is applicable to scenarios such as development and test, data sharing, and data research.

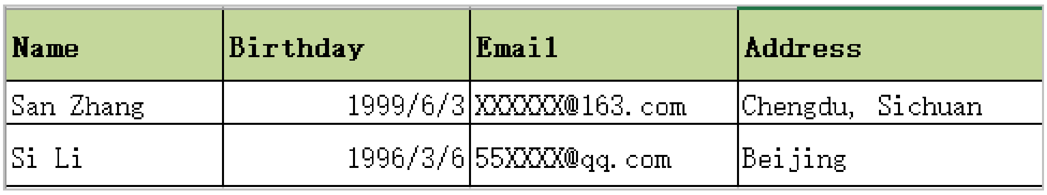
Common Causes of Data Breaches

- Insider leakage
  - Laptops or mobile devices are lost or stolen.
  - Sensitive data or storage is accessed by unauthorized personal
  - Data is stolen by employees.
  - Sensitive data is sent, printed, and copied by employees.
  - Sensitive data is accidentally transmitted out.
- Leakage caused by external attacks
  - Data access is uncontrollable, or there are security vulnerabilities in the data storage system.
  - Improper configurations allow external attacks.
  - Sensitive data or storage is accessed by unauthorized personal

Scenario

Assume that the **dsc_yunxiaoke** table in the **rsd-dsc-test** database stores the information of the following bank employees:

Figure 1-1 Bank employee information



Name	Birthday	Email	Address
San Zhang	1999/6/3	XXXXXX@163.com	Chengdu, Sichuan
Si Li	1996/3/6	55XXXX@qq.com	Beijing

To identify and mask sensitive data in the table, you can identify sensitive data and generate the identification result, and then mask the identified sensitive data using the SHA256 algorithm in **Hash**.

Step 1 Identifying Sensitive Data

Step 1 Buy DSC.

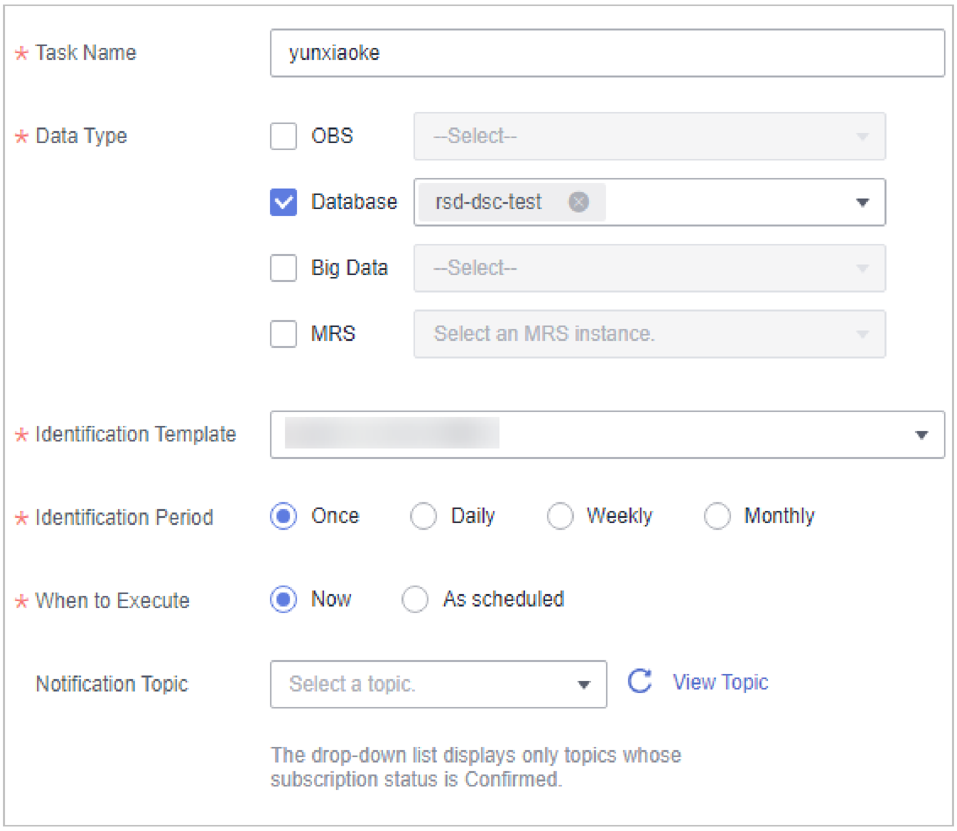
Step 2 Log in to the management console.

Step 3 In the left navigation page, click ☰, and choose **Security** > **Data Security Center**.

Step 4 In the left navigation pane, choose **Sensitive Data Identification** > **Identification Task**.

Step 5 Click **Create Task**. In the displayed dialog box, configure the basic parameters.

Figure 1-2 Creating a sensitive data identification task



* Task Name: yunxiaoke

* Data Type:
- □ OBS --Select--
- ☑ Database rsd-dsc-test
- □ Big Data --Select--
- □ MRS Select an MRS instance.

* Identification Template: [redacted]

* Identification Period: ● Once ○ Daily ○ Weekly ○ Monthly

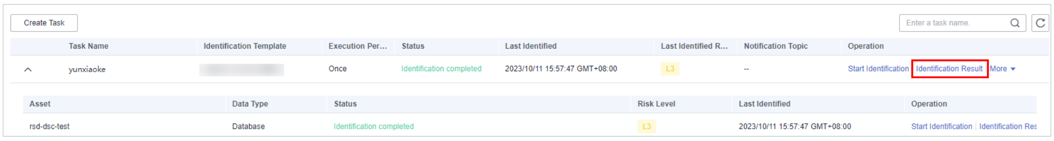
* When to Execute: ● Now ○ As scheduled

Notification Topic: Select a topic.

The drop-down list displays only topics whose subscription status is Confirmed.

Step 6 Click **OK**. The sensitive data identification task list is displayed.

Figure 1-3 Sensitive data identification task list

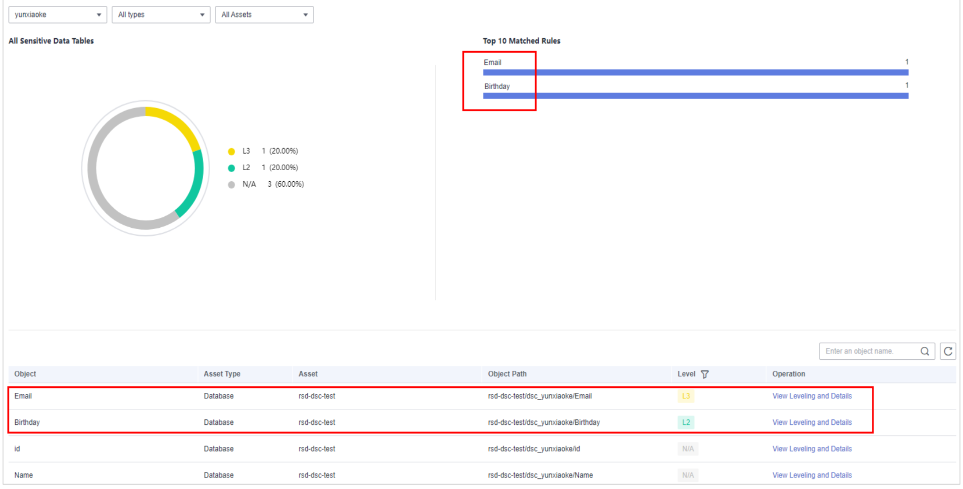


	Task Name	Identification Template	Execution Per...	Status	Last Identified	Last Identified R...	Notification Topic	Operation
^	yunxiaoke	[redacted]	Once	Identification completed	2023/10/11 15:57:47 GMT+08:00	L3	--	Start Identification Identification Result More ▾

Asset	Data Type	Status	Risk Level	Last Identified	Operation
rsd-dsc-test	Database	Identification completed	L3	2023/10/11 15:57:47 GMT+08:00	Start Identification | Identification Res

Step 7 When the status of the identification task changes to **Identification completed**. Click **View Result** in the **Operation** column to go to the result details page.

Figure 1-4 Identification result details



yunxiaoke ▾ | All types ▾ | All Assets ▾

All Sensitive Data Tables

Top 10 Matched Rules

Enter an object name.

Object	Asset Type	Asset	Object Path	Level	Operation
Email	Database	rsd-dsc-test	rsd-dsc-test/dsc_yunxiaoke/Email	L3	View Leveling and Details
Birthday	Database	rsd-dsc-test	rsd-dsc-test/dsc_yunxiaoke/Birthday	L2	View Leveling and Details
id	Database	rsd-dsc-test	rsd-dsc-test/dsc_yunxiaoke/id	N/A	View Leveling and Details
Name	Database	rsd-dsc-test	rsd-dsc-test/dsc_yunxiaoke/Name	N/A	View Leveling and Details

The birthday dates and email addresses are identified as sensitive data, as shown in **Figure 1-4**.

Step 8 Perform operations described in **Step 2. Masking Sensitive Data** to mask the sensitive data in the **Birthday** and **Email** columns of the **dsc_yunxiaoke** table in the **rds-dsc-test** database.

----End

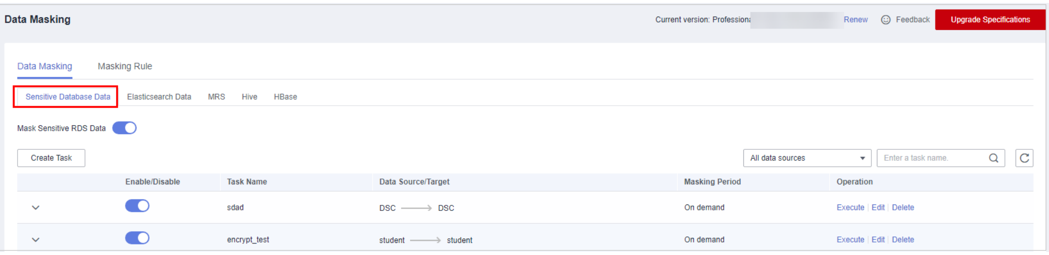
Step 2. Masking Sensitive Data

DSC supports database masking, ES masking, MRS masking, Hive masking, and HBase masking tasks. The masking methods are similar. This section uses creating a database static masking task as an example. For details about other masking methods, see:

- **Creating and Running an Elasticsearch Data Masking Task**
- **Creating and Running an MRS Data Masking Task**.
- **Creating and Running a Hive Masking Task**.
- **Creating and Running an HBase Masking Task**.

Step 1 In the left navigation pane, choose **Data Masking**. The **Data Masking** > **Sensitive Database Data Masking** page is displayed by default.

Figure 1-5 Accessing the Database Data Masking tab page



Data Masking

Mask Sensitive RDS Data ●

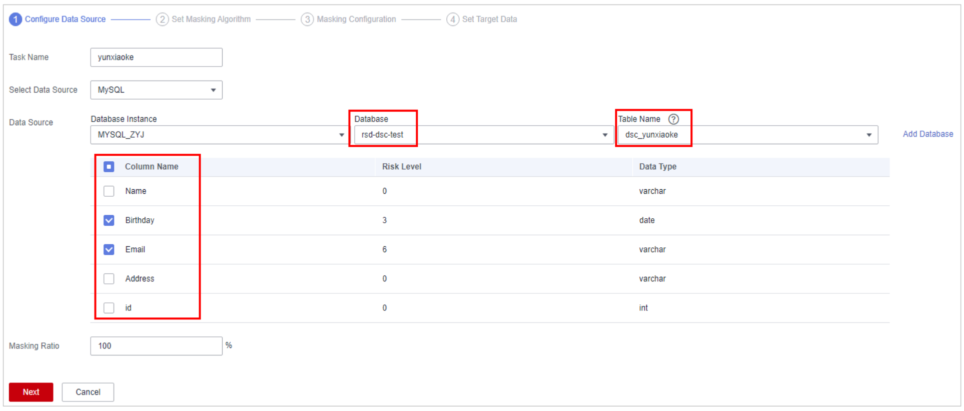
	Enable/Disable	Task Name	Data Source/Target	Masking Period	Operation
⌄	●	sdad	DSC → DSC	On demand	Execute | Edit | Delete
⌄	●	encrypt_test	student → student	On demand	Execute | Edit | Delete

Step 2 Set **Mask Sensitive RDS Data** to ●.

Step 3 Click **Create Task** to configure the data source.

Select all data types if you want a complete table that contains all types of data after the data masking is completed.

Figure 1-6 Data source configuration



Task Name: yunxiaoke

Select Data Source: MySQL

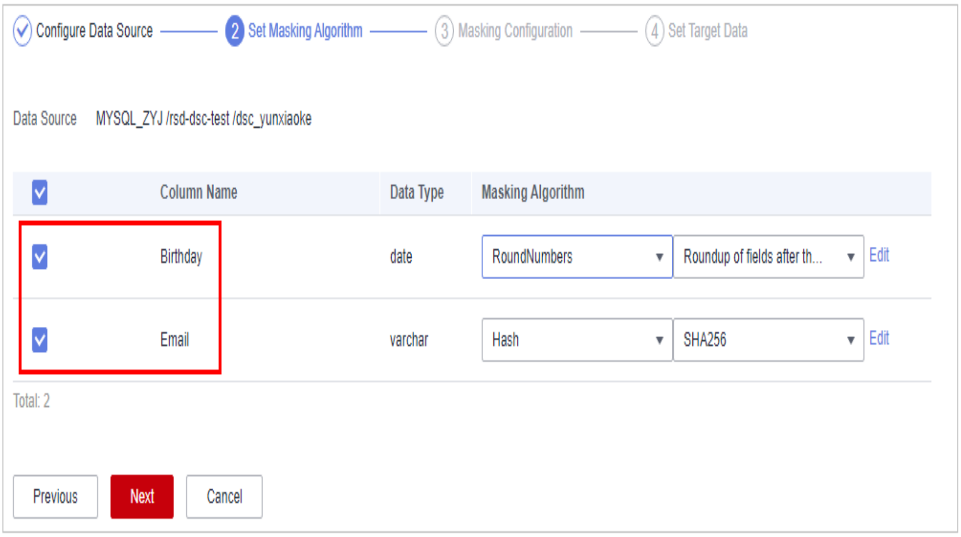
Data Source: Database Instance: MYSQL_ZYJ; Database: rsd-dsc-test; Table Name: dsc_yunxiaoke

■ Column Name	Risk Level	Data Type
□ Name	0	varchar
☑ Birthday	3	date
☑ Email	6	varchar
□ Address	0	varchar
□ id	0	int

Masking Ratio: 100 %

Step 4 Click **Next** to switch to **Set Masking Algorithm**.

Figure 1-7 Configuring the data masking algorithm



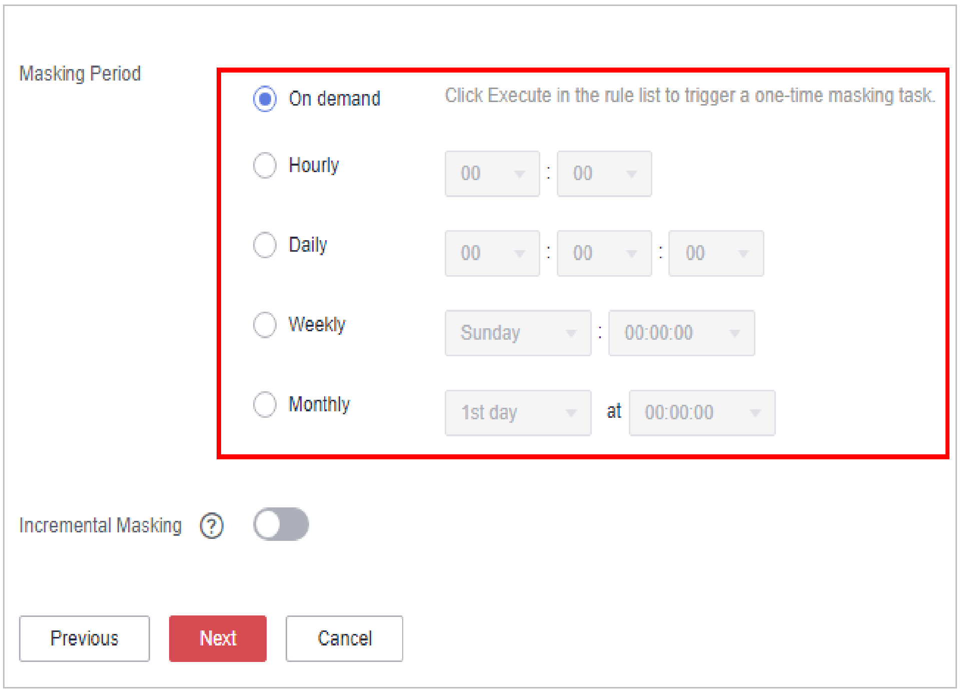
Data Source MYSQL_ZYJ /rsd-dsc-test /dsc_yunxiaoke

☑	Column Name	Data Type	Masking Algorithm		
☑	Birthday	date	RoundNumbers	Roundup of fields after th...	Edit
☑	Email	varchar	Hash	SHA256	Edit

Total: 2

Step 5 Click **Next** to switch to the **Configure Data Masking Period** page and configure the data masking period.

Figure 1-8 Configuring data masking period



Masking Period

● On demand — Click Execute in the rule list to trigger a one-time masking task.

○ Hourly — 00 : 00

○ Daily — 00 : 00 : 00

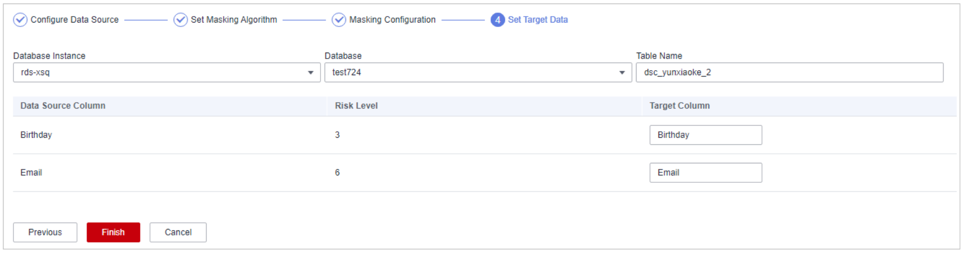
○ Weekly — Sunday : 00:00:00

○ Monthly — 1st day at 00:00:00

Incremental Masking ⓘ

Step 6 Click **Next** to the **Set Target Data** page and configure the storage location of the table generated after data masking.

Figure 1-9 Configuring the storage location of the table generated after data masking



Database Instance: rds-xsq

Database: test724

Table Name: dsc_yunxiaoke_2

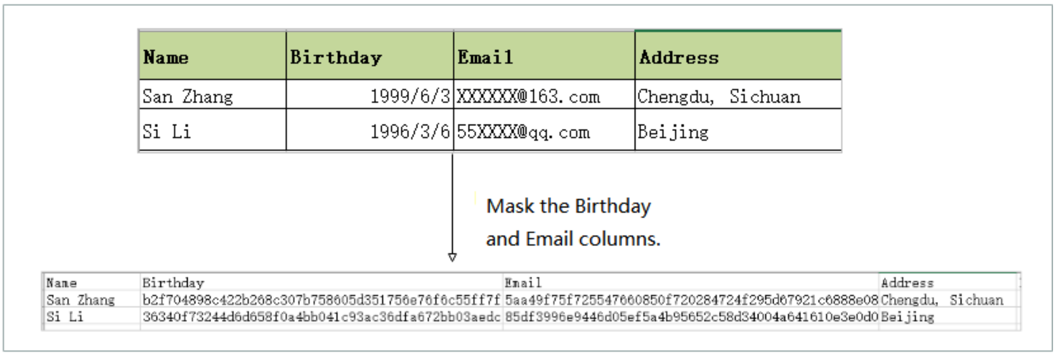
Data Source Column	Risk Level	Target Column
Birthday	3	Birthday
Email	6	Email

Step 7 Click **Finish** to return to the database data masking task list. Click to enable the masking task and then **Execute** in the **Operation** column to execute the task.

If the status changes to **Completed**, the data masking task has been successfully executed.

----End

Verifying the Result



Name	Birthday	Email	Address
San Zhang	1999/6/3	XXXXXX@163. com	Chengdu, Sichuan
Si Li	1996/3/6	55XXXX@qq. com	Beijing

Mask the Birthday
and Email columns.

Name	Birthday	Email	Address
San Zhang	b2f704898c422b268c307b758605d351756e76f6c55ff7f	5aa49f75f725547660850f720284724f295d67921c6888e08	Chengdu, Sichuan
Si Li	36340f73244d6d658f0a4bb041c93ac36dfa672bb03aedc	85df3996e9446d05ef5a4b95652c58d34004a641610e3e0d0	Beijing

2 Best Practices of OBS Data Security Protection

This document describes how to use the Data Security Center (DSC) to identify, classify, and protect sensitive data stored in OBS.

Overview

Sensitive data includes personal privacy information, passwords, keys, sensitive images, and other high-value data. Such data is usually stored in your OBS bucket in different formats. Once the data is leaked, enterprises will suffer significant economic and reputation losses.

After you authorize DSC to perform identification on the data source, DSC quickly identifies sensitive data from your massive data stored in OBS, classify the sensitive data and display it. DSC also traces the usage of sensitive data, and protects and audits data based on predefined security policies. In this way, DSC allows you to learn about the security status of your OBS data assets at any time.

Application Scenario

- Sensitive data identification

  OBS stores a large amount of data and files. However, it is difficult to have a clear knowledge of the sensitive information contained in OBS.

  You can use the built-in algorithm rules of DSC or customize industry rules to scan, classify, and grade data stored in OBS, and take further security protection measures based on the scanning results. For example, you can use the access control and encryption functions of OBS.

- Anomaly detection and audit

  The DSC can detect access, operation, and management anomalies related to sensitive data and send alarms to you for you to confirm and handle the anomalies. The following behaviors are regarded as anomalies:

  - Unauthorized users access and download sensitive data.
  - Authorized users access, download, and modify sensitive data, as well as change and delete permissions.
  - Authorized users change or delete permissions granted for buckets that contain sensitive data.

- Users who accessed sensitive files fail to log in to the device.

Procedure

Step 1 Buy DSC.

Step 2 Log in to the management console.

Step 3 In the left navigation page, click ☰, and choose **Security** > **Data Security Center**.

Step 4 In the navigation pane, choose **Assets**, and click **Allow Access to Cloud Assets** in the upper right corner of the page.

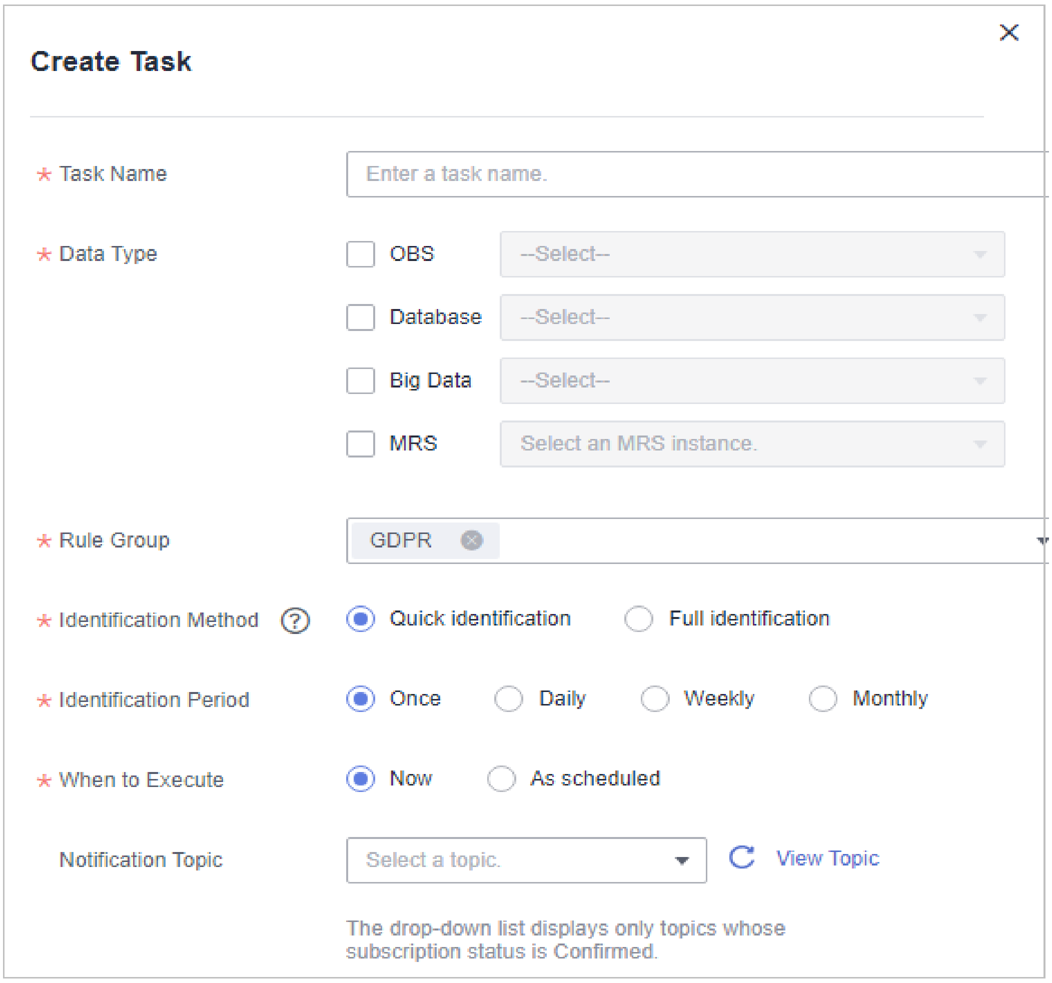
Step 5 Locate the row that contains the OBS asset, click ○ in the **Operation** column to enable authorization.

Step 6 For details about how to add OBS assets, see **Adding OBS Assets**.

Step 7 In the navigation tree on the left, choose **Sensitive Data Identification** > **Identification Task**. Click **Create Task** to configure a sensitive data scanning task.

Select **OBS** for **Data Type** and select the OBS asset added in section **Step 6**. For details about other configurations, see section **Creating a Task**.

Figure 2-1 Creating an identification task



Create Task

* Task Name: Enter a task name.

* Data Type:
- □ OBS --Select--
- □ Database --Select--
- □ Big Data --Select--
- □ MRS Select an MRS instance.

* Rule Group: GDPR

* Identification Method: ● Quick identification ○ Full identification

* Identification Period: ● Once ○ Daily ○ Weekly ○ Monthly

* When to Execute: ● Now ○ As scheduled

Notification Topic: Select a topic. View Topic

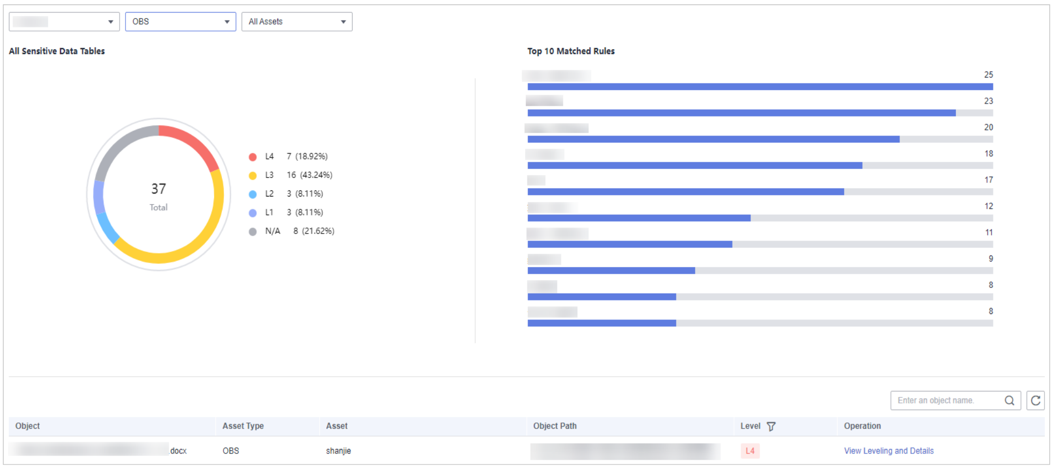
The drop-down list displays only topics whose subscription status is Confirmed.

Step 8 In the navigation pane, choose **Sensitive Data Identification** > **Identification Task**.

Step 9 Click **Identification Result** in the **Operation** column to view the Identification result.

In the upper left corner of the page, set **Task Name** to **dsctest**, **Data Type** to **OBS**, and **Asset types** to **All Assets** to filter the OBS sensitive data identification result, as shown in **Figure 2-2**.

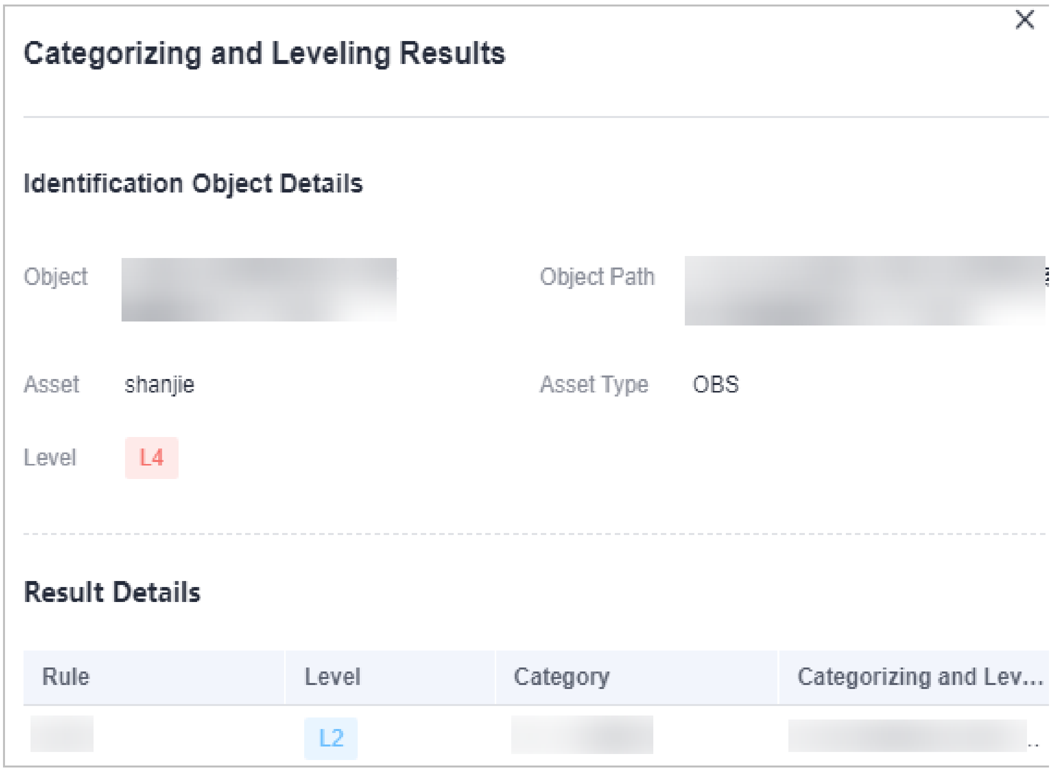
Figure 2-2 Identification result details



Object	Asset Type	Asset	Object Path	Level	Operation
docx	OBS	shanjie		L4	View Leveling and Details

Step 10 In the row containing the desired scan object, click **View Categorizing and Leveling Result Details** in the **Operation** column. The **Categorizing and Leveling Result Details** dialog box is displayed.

Figure 2-3 Categorizing and leveling results



Categorizing and Leveling Results

Identification Object Details

Object

Object Path

Asset shanjie

Asset Type OBS

Level L4

Result Details

Rule	Level	Category	Categorizing and Lev...
	L2		..

1. In the **Data Storage Security** area on the **Overview** page, check whether there are unencrypted object buckets in the risky databases. If yes, click the bucket name to go to the OBS page and encrypt the unencrypted object buckets. For details, see **Configuring Bucket Default Encryption**.
2. In the alarm list, view anomalies based on the risk level and check whether there are high-risk events. For detailed about operations, see **Viewing**

Abnormal Behaviors Through Data Usage Audit and **Viewing Details About Access Key Leakage Events**.

3. On OBS Console, modify the read and write permissions of the risky buckets or files.

----End

A Change History

Date	Description
2023-11-30	This issue is the first official release.