Oracle Technology Netw

Replacing Sensitive Data By Using the Data Masking Pack

This tutorial contains the following sections:

Purpose

This tutorial shows you how to identify sensitive data, select appropriate mask formats for the sensitive fields, and apply the mask.

The steps in this tutorial can be performed using either **Oracle Enterprise Manager Database Control 11g Release 2**, or **Oracle Enterprise Manager Grid Control 10.2.0.5**.

Time to Complete

Approximately 1 hour.

Overview

The Data Masking Pack helps organizations share production data in compliance with privacy and confidentiality policies by replacing sensitive data with realistic but scrubbed data based on masking rules.

There are two primary use cases for the Data Masking Pack. First, DBAs who want to take a copy of production data for testing purposes and use the Data Masking Pack to replace all sensitive data with innocuous but realistic information, and then make this database available to developers. Second, organizations want to share production data with third parties while hiding sensitive or personally identifiable information.

Scenario

In this tutorial, you review the HR schema in a staging environment to identify sensitive data and then select appropriate mask formats for the sensitive data. After creating the mask, you then proceed to apply the mask and replace the sensitive fields in the production copy with realistic but scrubbed data.

Software and Hardware Requirements

The following is a list of software requirements:

Oracle Database 11g Release 2 (with Spatial option installed and supporting objects created in the database)
Oracle Enterprise Manager Grid Control 10.2.0.5

Note: Oracle Enterprise Manager **Grid Control** 10.2.0.5 is **only required** if you want to use it to perform the steps in this OBE, or if you want to apply data masking to an Oracle Database 11g Release 1 database. If you want to use Enterprise Manager Database Control, then you can only apply the data masking to an Oracle Database 11g Release 2 database.

Prerequisites

Before starting this tutorial, you should:

1. Install Oracle Database 11g Release 2.

2. Verify that the Spatial option is installed and the supporting database objects are created. Follow these steps for the verification:

    1. Use SQL*Plus and logon to the database as sys or system.
    2. Issue this select statement: SELECT name, value FROM v$option WHREE name = 'Spatial'
        You should see True in the value column
        If you see False, then you need to use the Oracle Installer and install the Spatial option.
    3. Issue this statement: DESCRIBE sdo_geometry
        You should see the object type and its methods listed
        If you see an "object does not exist" error message, then you need to install Oracle Intermedia and then create the objects for Oracle Spatial.
        To install Intermedia, as sys, run:
            cd $ORACLE_HOME/ord/admin/ordinst.sql
            Make sure there are no errors. Java in the server must be installed prior to intermeida. You will get an error if it's not. If you get the error, install Java in the server, and then run ordinst again.
            After ordinst.sql is successfully run, you can create the Spatial objects. As sys, run:
                cd $ORACLE_HOME/md/admin
                mdinst.sql

3. **Optionally** install Oracle Enterprise Manager Grid Control 10.2.0.5 if you want to use it instead of Enterprise Manager Database Control.

4. Download and unzip the datamask.zip file into your working directory. In this tutorial the working directory is datamask.

Creating Tables for the Tutorial

To create tables to use during this data masking tutorial, perform the following steps:

1. Log in to SQL*Plus as the `SYSTEM` user.

```
[oracle@host01 datamask]$ sqlplus system

SQL*Plus: Release 11.2.0.1.0 Production on Wed Apr 28 10:49:17 2010

Copyright (c) 1982, 2009, Oracle.  All rights reserved.

Enter password:

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - Production
With the Partitioning, OLAP, Data Mining and Real Application Testing opti
ons

SQL>
```

2. Execute the `OBE_DM_setup01.sql` script to prepare for this tutorial. This script deletes objects that may have been created during a previous execution of the tutorial.

```
SQL> @OBE_DM_setup01
SQL>
SQL> --- Cleanup from possible previous executions
SQL>
SQL> DROP USER HR_TEST CASCADE;
DROP USER HR_TEST CASCADE
          *
ERROR at line 1:
ORA-01918: user 'HR_TEST' does not exist


SQL>
SQL> DROP USER OE_TEST CASCADE;
DROP USER OE_TEST CASCADE
          *
ERROR at line 1:
ORA-01918: user 'OE_TEST' does not exist


SQL>
```

3. Execute the `OBE_DM_setup02.sql` script to create the `HR_TEST` and `OE_TEST` schemas. This step simulates cloning the production database to an instance that is used for masking, but much smaller in scope. Enter a password for the `HR_TEST` and `OE_TEST` users when prompted.

```
SQL> @OBE_DM_setup02
SQL> ----------------------------------------------------------
SQL> --  File created - Tuesday-April-27-2010
SQL> ----------------------------------------------------------
SQL> CREATE USER hr_test
  2  identified by &&1
  3  default tablespace users
  4  temporary tablespace temp
  5  /
Enter value for 1: oracle_1
```

4. The `HR_TEST` and `OE_TEST` users are created and objects are created in their schemas.

```
SQL> Insert into CUSTOMERS (CUSTOMER_ID,CUST_FIRST_NAME,CUST_LAST_NAME,CUS
T_ADDRESS,PHONE_NUMBERS,NLS_LANGUAGE,NLS_TERRITORY,CREDIT_LIMIT,CUST_EMAIL
,ACCOUNT_MGR_ID,CUST_GEO_LOCATION,DATE_OF_BIRTH,MARITAL_STATUS,GENDER,INCO
ME_LEVEL) values (981,'Daniel','Gueney',OE_TEST.CUST_ADDRESS_TYP('1668 Cho
ng Tao','111181','Beijing',null,'CN'),OE_TEST.PHONE_LIST_TYP('+86 10 012 3
839'),'zhs','CHINA',200,'Daniel.Gueney@REDPOLL.COM',148,null,to_date('07-O
CT-73','DD-MON-RR'),'married','M','K: 250,000 - 299,999');

1 row created.

SQL>
SQL> ------------------------------------------------
SQL> --   END DATA FOR TABLE CUSTOMERS
SQL> ------------------------------------------------
SQL>
SQL>
```

5. Execute the `OBE_DM_setup03.sql` script to add columns and data to the `HR_TEST.EMPLOYEES` table.

```
SQL> @OBE_DM_setup03
SQL> set echo on feed on
SQL> grant select on oe_test.customers to hr_test;

Grant succeeded.

SQL>
SQL> connect hr_test/&&1
Connected.
SQL> update locations set state_province = city where state_province is nu
ll;

6 rows updated.

SQL> update locations set postal_code = trunc(dbms_random.value(300000, 40
0000)) where postal_code is null;

1 row updated.
```

6. The script also creates a table named `HR_TEST.MASK_DATA` which is used to illustrate how you use a data table from a commercial provider.

```
SQL> create table mask_data as
  2  select cust_first_name first_name, cust_last_name last_name,
  3      c.cust_email email,
  4   (select p.column_value as phone_numbers from table (phone_numbers) p
 where rownum = 1) phone_number,
  5      c.CUST_ADDRESS.STREET_ADDRESS street_address,
  6      c.CUST_ADDRESS.CITY city,
  7      c.CUST_ADDRESS.STATE_PROVINCE state_province,
  8      c.CUST_ADDRESS.POSTAL_CODE postal_code,
  9      c.CUST_ADDRESS.country_id country_id
 10  from oe_test.customers c;

Table created.

SQL>
```

Identifying Sensitive Data

In this section you will view data in the `HR_TEST.EMPLOYEES` table to determine which columns should be masked.

1. Launch Enterprise Manager Database Control by entering the following URL: https://<hostname>:1158/em

   Note: If you are using Enterprise Manager **Grid Control** 10.2.0.5, specify the appropriate URL for your environment.

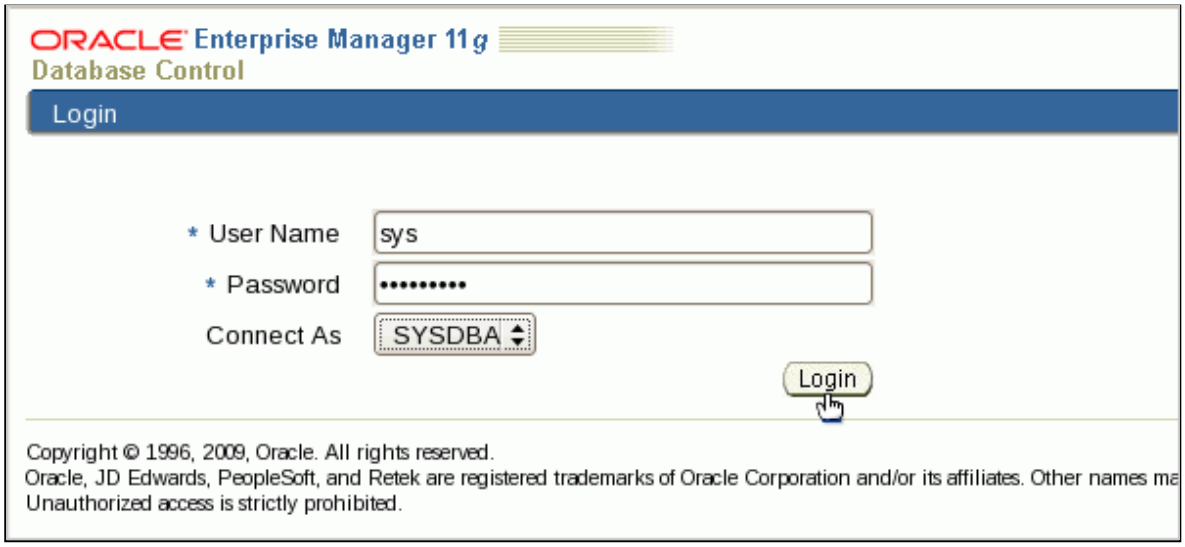
2. Enter the following information to log in to Enterprise Manager Database Control:

   Username: `SYS`

   Password: ********

   Connect As: `SYSDBA`

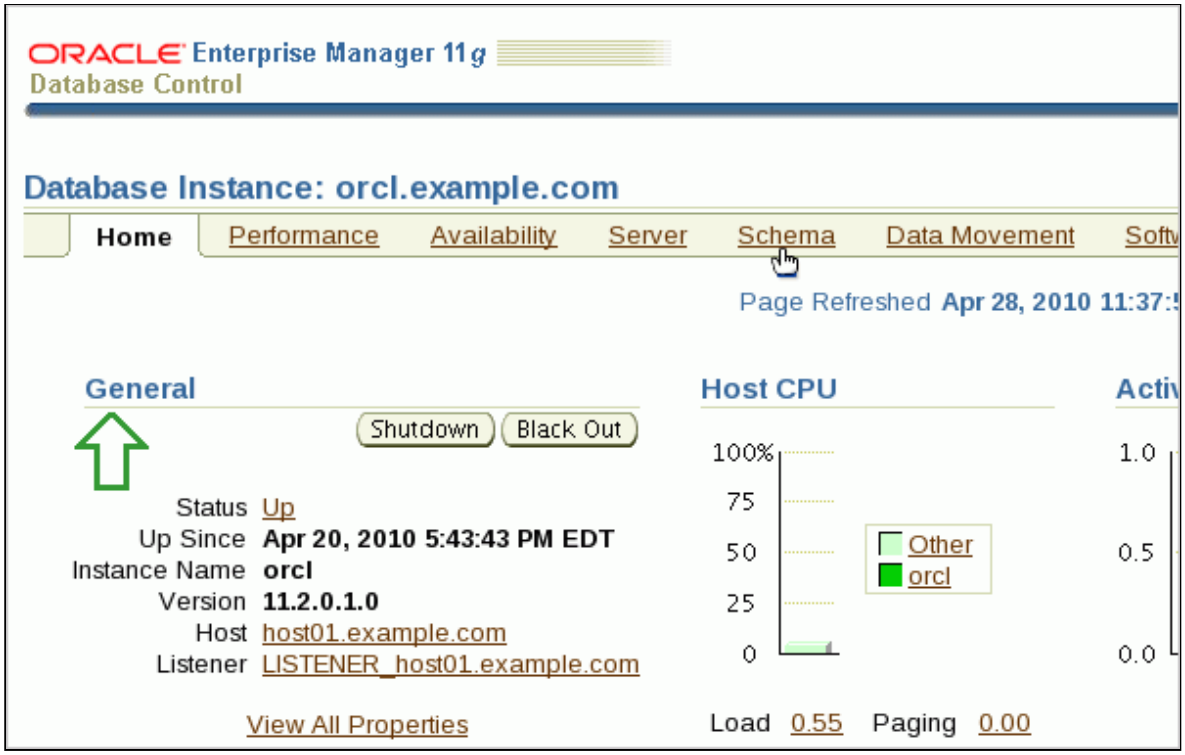
   Note: If you are using Enterprise Manager **Grid Control**, login using the `SYSMAN` database account.



3 . Click the Schema tab on the Database Instance Home page.

   Note: If you are using Enterprise Manager **Grid Control**, navigate to the Database Instance Home page by clicking

Targets tab > Databases tab > Oracle database SID link. You may be asked to set the preferred database credential. If so, use the `SYSTEM` database account.



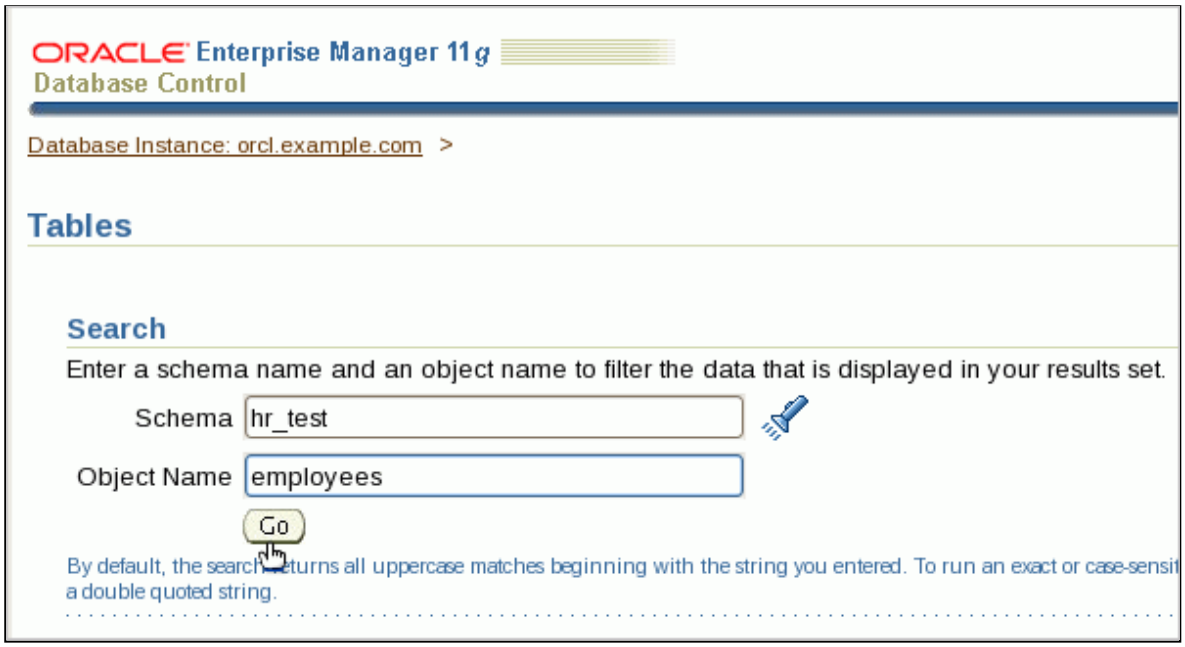
4 . Click Tables in the Database Objects region.

5 . Enter the following information:

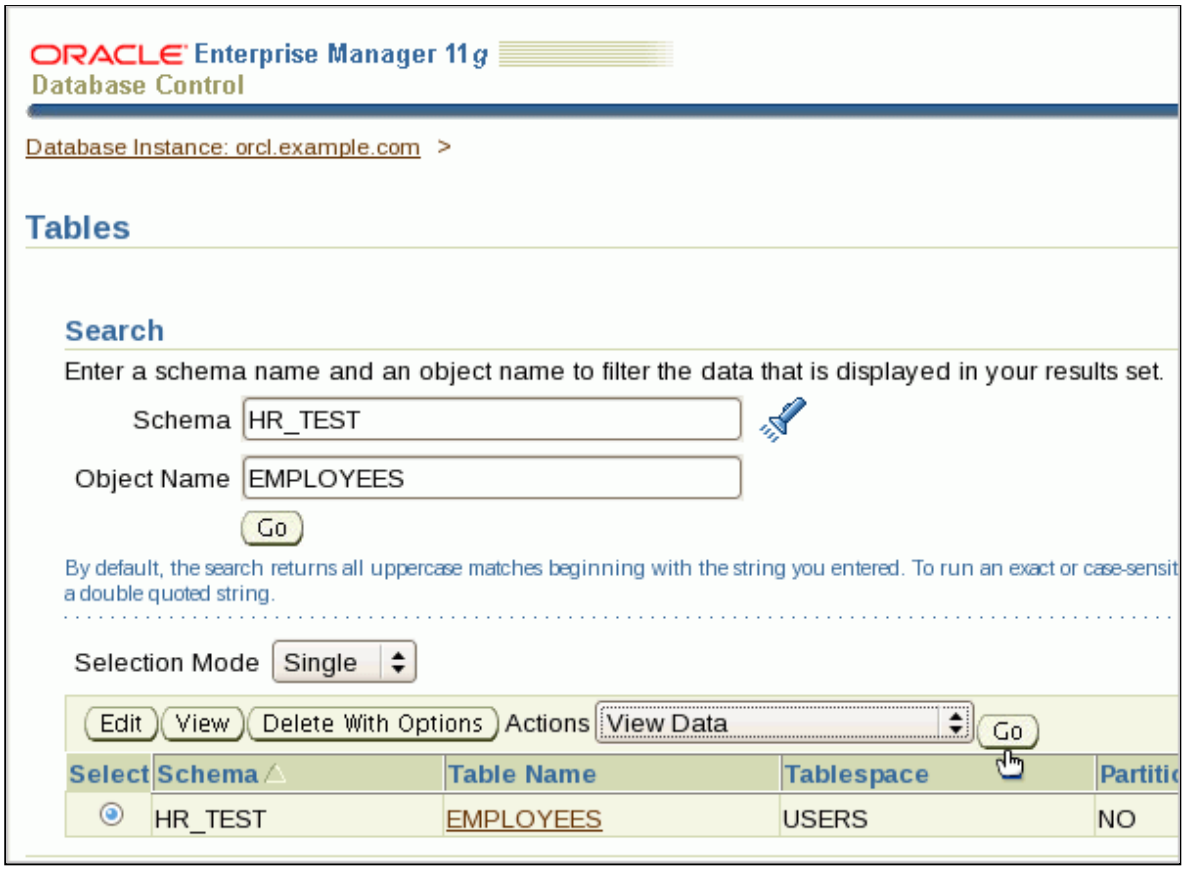
Schema: `hr_test`

Object Name: `employees`

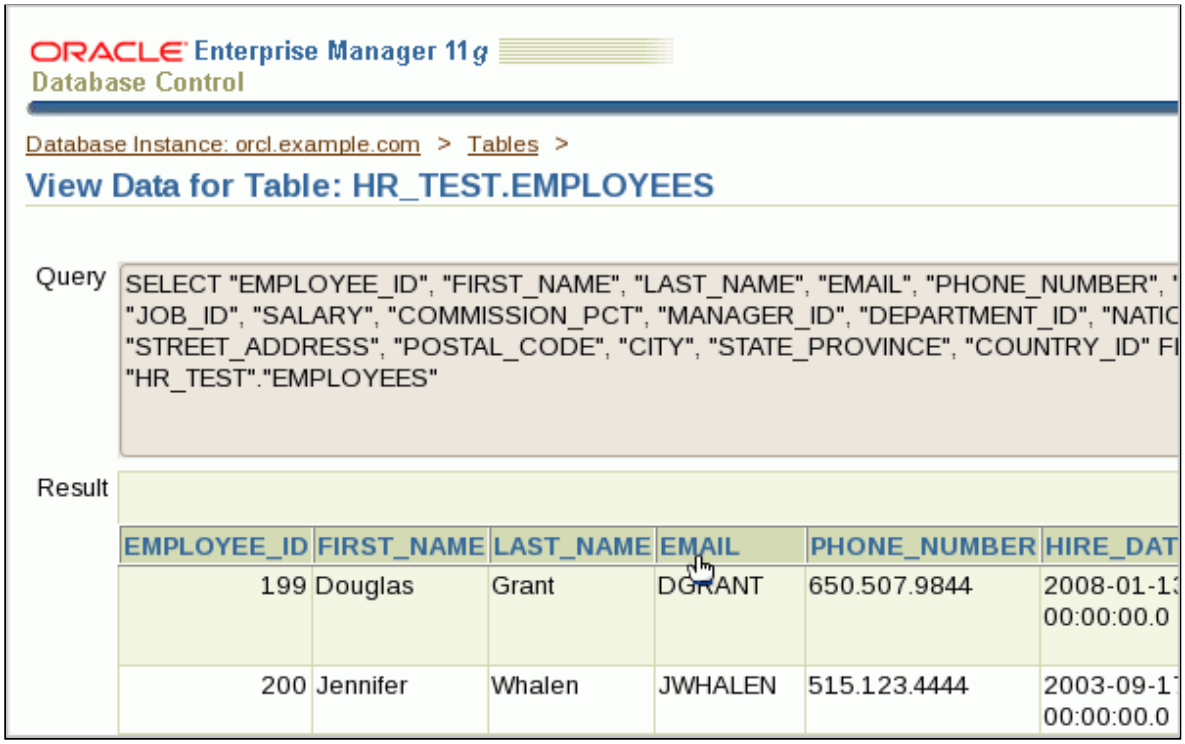
Click Go.



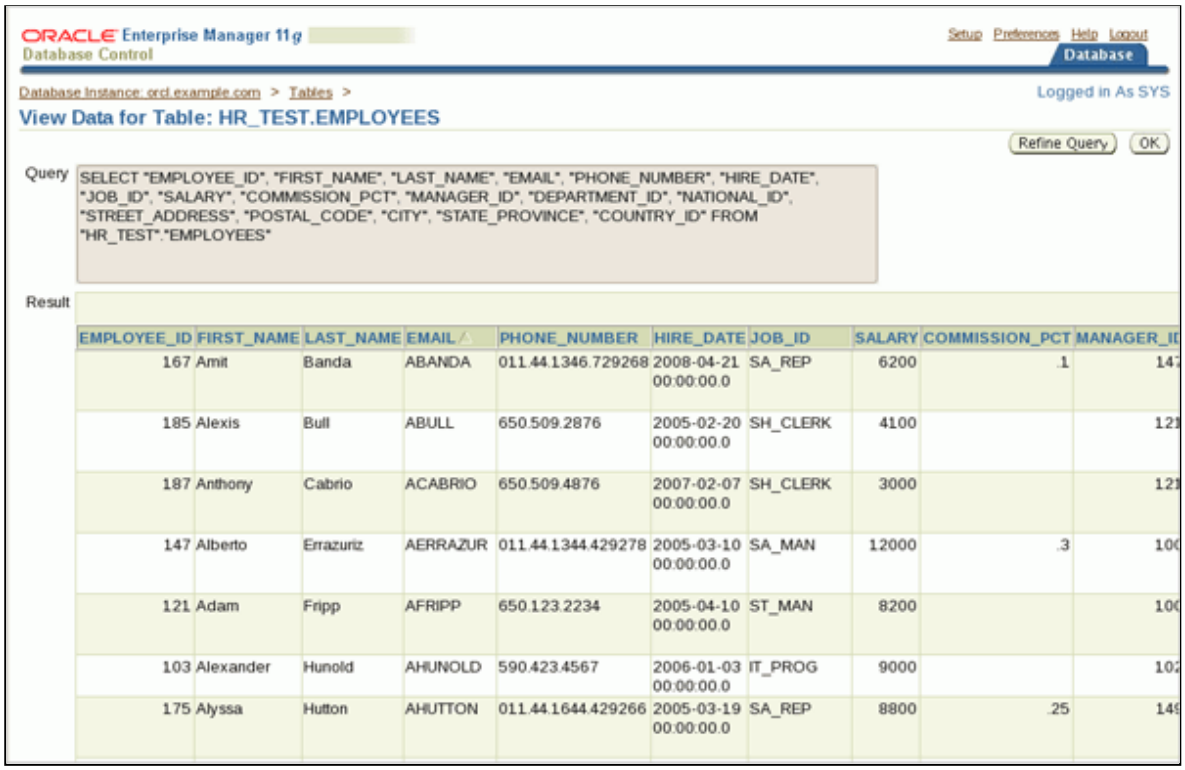
6 . Select View Data in the Actions list. Click Go.



7. View the data in the `HR_TEST.EMPLOYEES` table to determine likely candidates for data masking. For ease of comparison after perform the data masking operation, click `EMAIL` to sort the rows on the `EMAIL` column.



8. The rows are now sorted based on the `EMAIL` column. Keep this browser window open for comparison with the data after it has been masked. In the next section, you will begin masking the columns that you have identified as containing sensitive data.



Creating Masking Definitions: EMPLOYEE_ID Column

You have been informed that the `EMPLOYEES.EMPLOYEE_ID` column contains sensitive data. In this section you create a masking definition for the `EMPLOYEE_ID` column of the `HR_TEST.EMPLOYEES` table. You must also add a dependent column to the masking definition. The `MANAGERS.MGR_ID` column is not declared as a foreign key, but is dependent on `EMPLOYEES.EMPLOYEE_ID` at the application level.

1. Open a new browser window or tab, and launch Enterprise Manager Database Control by entering the following URL: https://<hostname>:1158/em

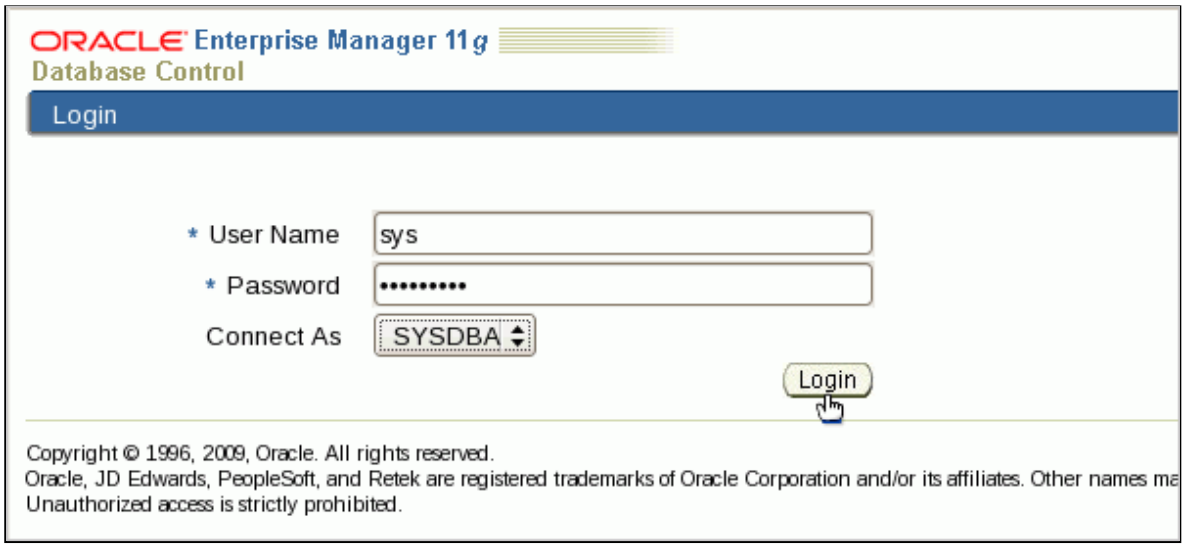
   Note: If you are using Enterprise Manager **Grid Control** 10.2.0.5, specify the appropriate URL for your environment.

2. Enter the following information to log in to Enterprise Manager Database Control:

   Username: `SYS`
   Password: *********
   Connect As: `SYSDBA`

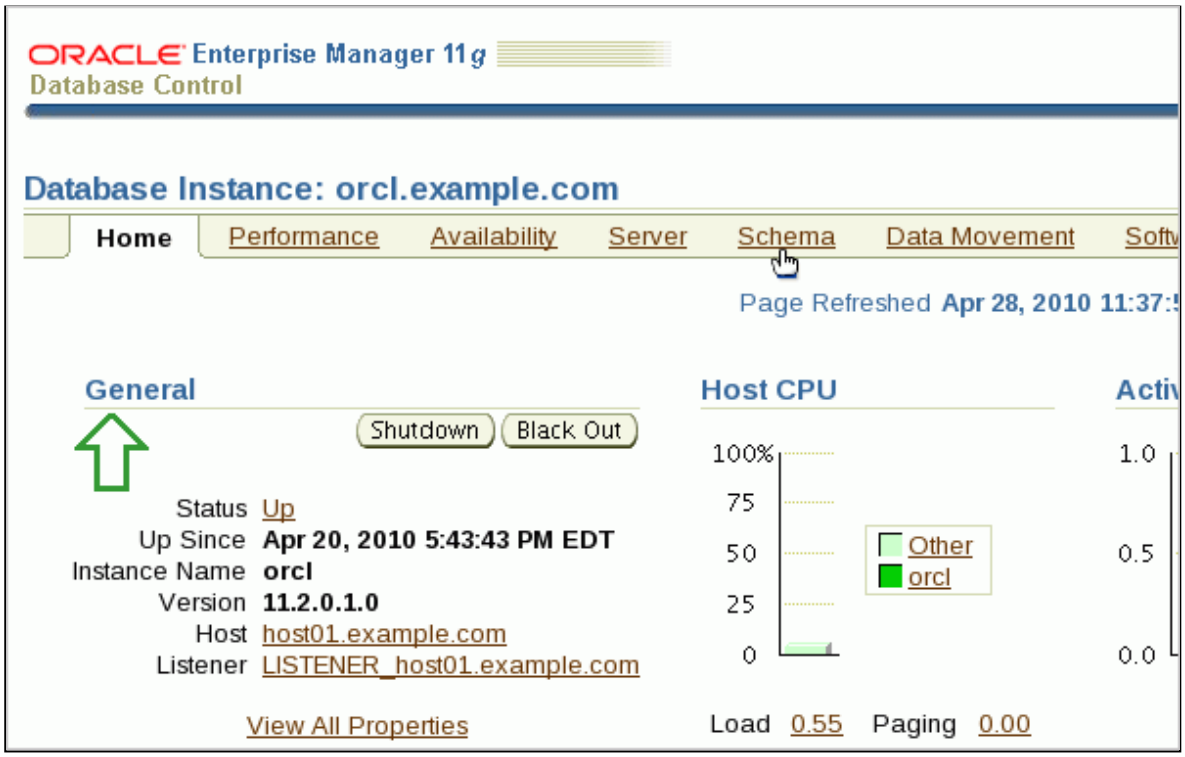
   Click Login.

   Note: If you are using Enterprise Manager **Grid Control**, login using the `SYSMAN` database account.

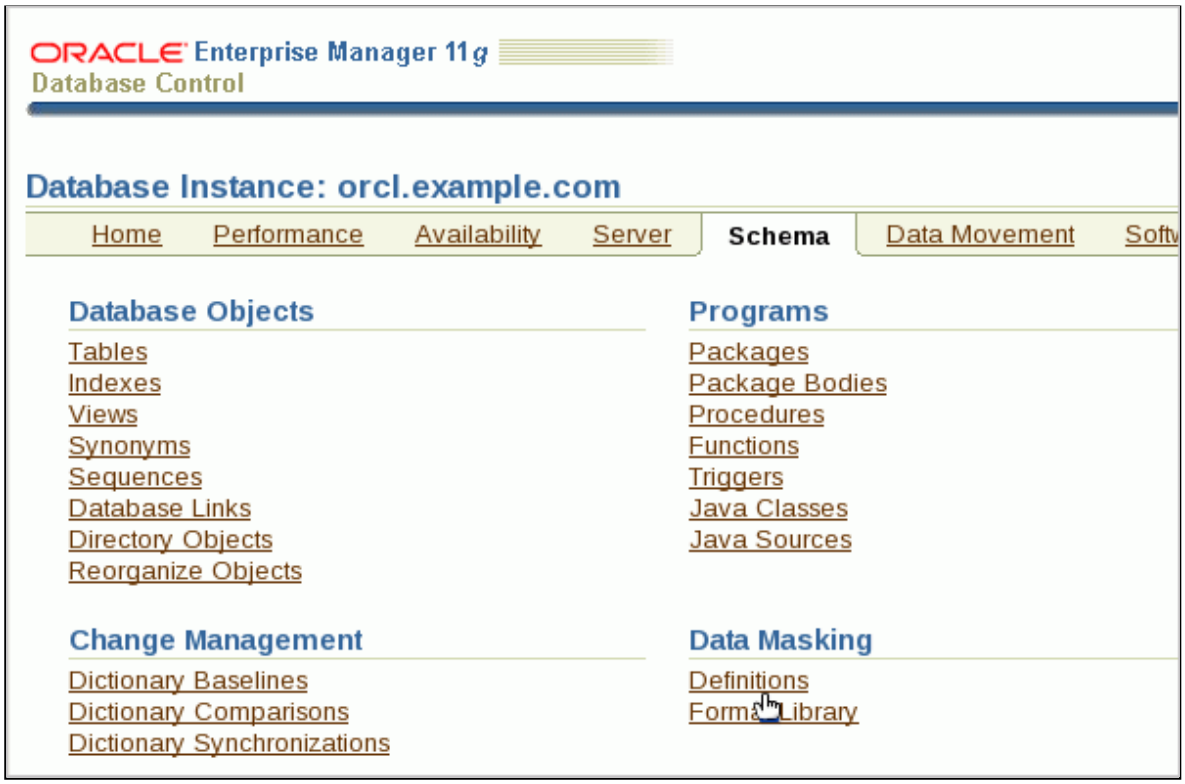


3 . On the Database Instance Home page, click the Schema tab.

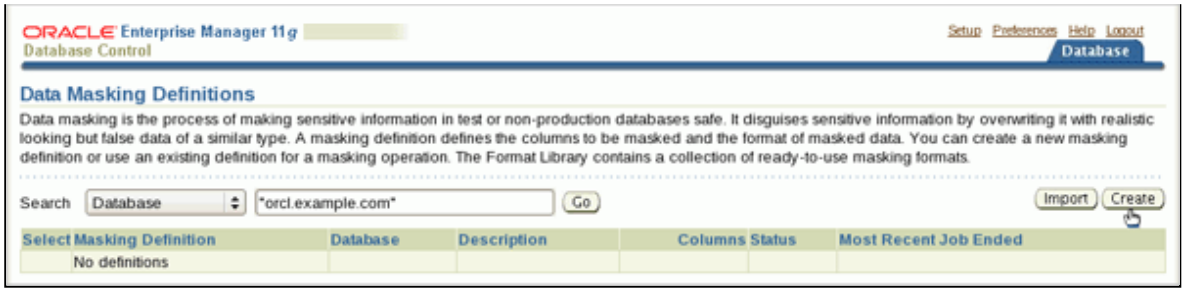
Note: If you are using Enterprise Manager **Grid Control**, navigate to the Database Instance Home page by clicking Targets tab > Databases tab > Oracle database SID link.



4 . Click Definitions in the Data Masking section.



5 . On the Data Masking Definitions page, Click Create.

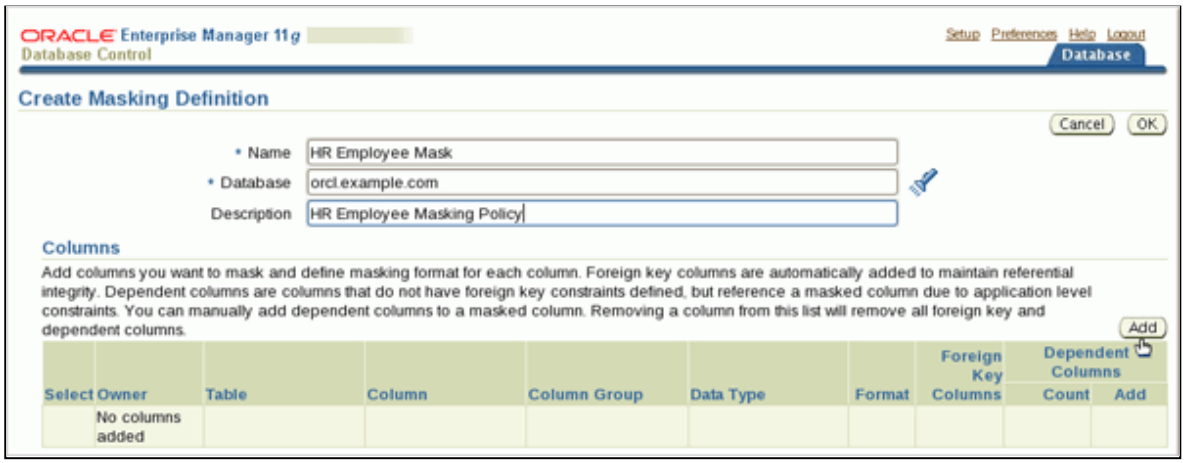


6 . On the Create Masking Definition page, enter the following:

Name: HR Employee Mask

Description: HR Employee Masking Policy

In the Columns section, click Add.

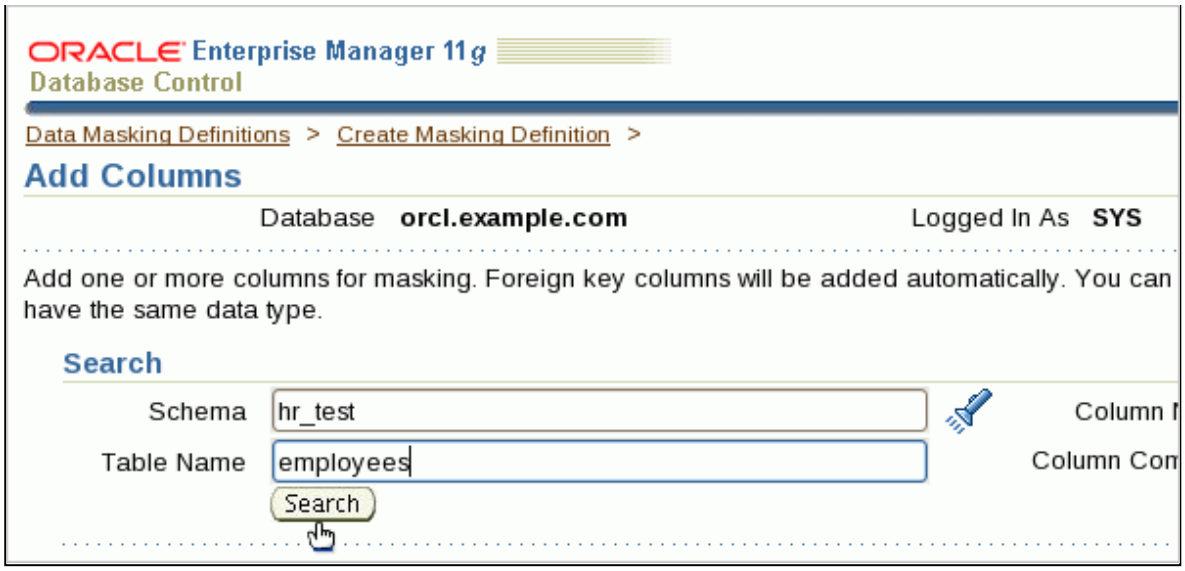


7 . On the Add Columns page, enter the following:

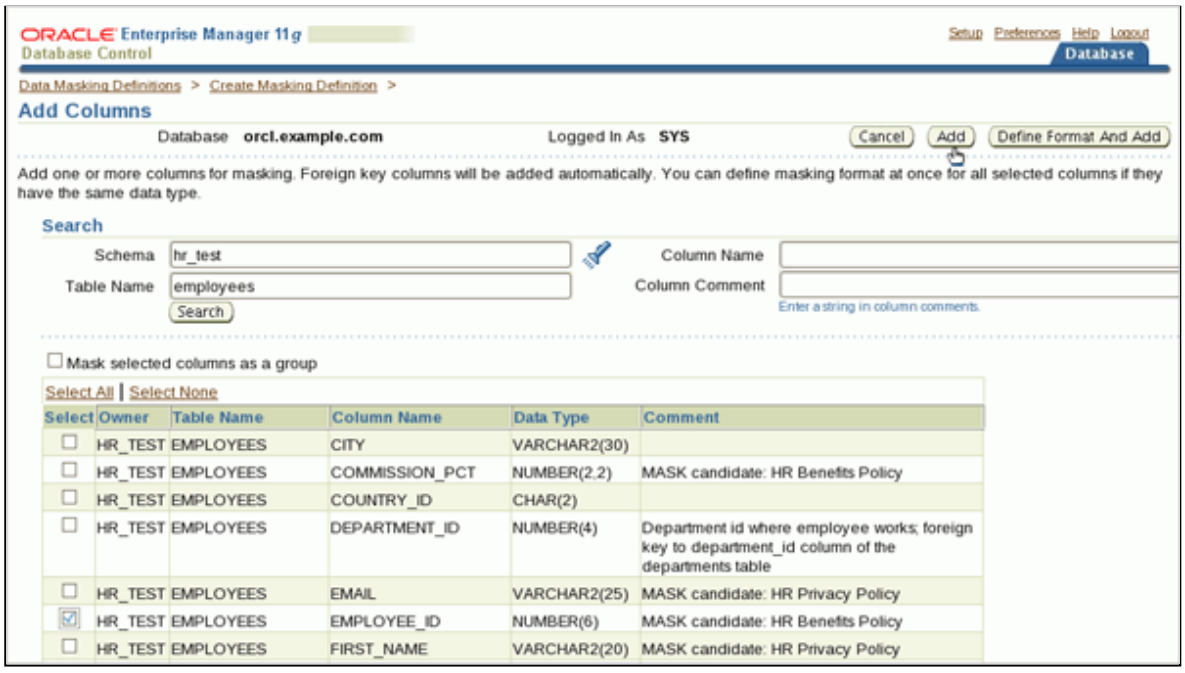
Schema: `hr_test`

Table Name: `employees`

Click Search.



8 . Notice that the Comment column contains information supplied by the application DBA indicating which columns are mask candidates. Select the `EMPLOYEE_ID` column. Click Add.



9 . Note that the foreign key columns were automatically added to the masking definition. On the Create Masking Definition page, click the + icon under the Dependent Columns heading.

10 . On the Add Dependent Columns page, enter the following:

Schema: `hr_test`

Table Name: `managers`

Click Search.

11 . Select the `MGR_ID` column. Click Add.

12 . On the Create Masking Definition page, click the icon under the Format heading.

13 . On the Define Column Mask page, select Random Numbers from the Format Entry list. Click Add.

14 . Enter the following information:

Start value: 100000

End value: 999999

Click the icon in the Sample column to view sample data.

15 . Sample data is displayed. You can click the icon again to view additional sample values. Click OK when you have finished viewing sample values.

16 . On the Create Masking Definition page, click OK.

17 . The Data Masking Definitions page is displayed showing the HR Employee Mask masking definition.

Creating Masking Definitions: FIRST_NAME and LAST_NAME Columns

Creating Masking Formats for FIRST_NAME and LAST_NAME

Create the masking formats for the `EMPLOYEES.FIRST_NAME` and `EMPLOYEES.LAST_NAME` columns by using the `HR_TEST.MASK_DATA` table as the source of masking data. These steps illustrate how you would use a data table from a commercial data provider to mask confidential data such as names.

1. On the Data Masking Definitions page, click the Format Library link.

2. On the Format Library page, click Create.

3 . On the Create Format page, enter the following information:

Name: Anglo-American First Name

Description: Masking format for first name

Select Table Column in the list and click Go.

4. On the Create Format page, enter the following information:

   Table Name: `hr_test.mask_data`

   Column Name: `first_name`

   Click OK.

5. On the Create Format page, click OK.

6. A confirmation message is displayed on the Format Library page. Click Create.

7. On the Create Format page, enter the following information:

   Name: Anglo-American Last Name

   Description: Masking format for last name

   Select Table Column in the list and click Go.

8. On the Create Format page, enter the following information:

   Table Name: `hr_test.mask_data`

   Column Name: `last_name`

   Click OK.

9. On the Create Format page, click OK.

10 . A confirmation message is displayed on the Format Library page.

Creating Masking Definitions for FIRST_NAME and LAST_NAME

Create the masking definitions for the `EMPLOYEES.FIRST_NAME` and `EMPLOYEES.LAST_NAME` columns. Use the mask formats that you defined in the previous step.

1. On the Format Library page, click the Data Masking Definitions link.

2. On the Data Masking Definitions page, select HR Employee Mask. Click Edit.

3 . On the Edit Masking Definition: HR Employee Mask page, click Add.

4 . On the Add Columns page, enter the following information:

Schema: `hr_test`

Table Name: `employees`

Click Search.

5 . On the Add Columns page, select the `FIRST_NAME` and `LAST_NAME` columns. Click Add.

6 . On the Edit Masking Definition: HR Employee Mask page, click the Format icon in the `FIRST_NAME` row.

7 . On the Define Column Mask page, click Import Format.

8 . On the Import Format page, select Anglo-American First Name. Click Import.

9 . Click the icon in the Sample column to view sample data.

10 . On the Define Column Mask page, click OK.

11 . On the Edit Masking Definition: HR Employee Mask page, click the Format icon in the `LAST_NAME` row.

12 . On the Define Column Mask page, click Import Format.

13 . On the Import Format page, select Anglo-American Last Name. Click Import.

14 . On the Define Column Mask page, click the icon in the Sample column to view sample data.

15 . On the Define Column Mask page, click OK.

16 . On the Edit Masking Definition: HR Employee Mask page, click OK.

17 . On the Data Masking Definitions page, observe that four columns are masked.

Creating Masking Definitions: SALARY Column

Add the `EMPLOYEES.SALARY` column to the HR Employee Mask masking definition and specify the Shuffle mask format. The Shuffle format is used to shuffle the values in the column amongst the rows.

1. On the Data Masking Definitions page, select HR Employee Mask. Click Edit.

2. On the Edit Masking Definition: HR Employee Mask page, click Add.

3 . On the Add Columns page, enter the following information:

Schema: `hr_test`

Table Name: `employees`

Click Search.

4 . On the Add Columns page, select the `SALARY` column. Click Define Format and Add.

5 . On the Define Column Mask page, select Shuffle from the Format Entry list. Click Add.

6 . On the Define Column Mask page, click the icon in the Sample column to view sample data.

7 . On the Define Column Mask page, click OK.

8 . On the Edit Masking Definition: HR Employee Mask page, click OK.

Creating Masking Definitions: COMMISSION_PCT Column

Add the `EMPLOYEES.COMMISSION_PCT` column to the HR Employee Mask masking definition and specify the Shuffle mask format.

1. On the Data Masking Definitions page, select HR Employee Mask. Click Edit.

2. On the Edit Masking Definition: HR Employee Mask page, click Add.

3 . On the Add Columns page, enter the following information:

Schema: `hr_test`

Table Name: `employees`

Click Search.

4 . On the Add Columns page, select the `COMMISSION_PCT` column. Click Define Format and Add.

5 . On the Define Column Mask page, select Shuffle from the Format Entry list. Click Add.

6. On the Define Column Mask page, click the icon in the Sample column to view sample data.

7. On the Define Column Mask page, click OK.

8. On the Edit Masking Definition: HR Employee Mask page, click OK.

Creating Masking Definitions: NATIONAL_ID Column

Implement condition-based masking for the `NATIONAL_ID` column. Configure the masking so that the `NATIONAL_ID` column is masked with the National Insurance Number Formatted format for UK employees and Social Security Number Formatted for US employees. The `NATIONAL_ID` column for employees from other countries does not need to be masked.

1. On the Data Masking Definitions page, select HR Employee Mask. Click Edit.

2. On the Edit Masking Definition: HR Employee Mask page, click Add in the Columns section.

3. On the Add Columns page, enter the following information:

   Schema: `hr_test`

   Table Name: `employees`

   Click Search.

4. Select the `NATIONAL_ID` column. Click Define Format and Add.

5. On the Define Column Mask page, click Add Condition.

6 . On the Define Column Mask page, enter the following SQL query in the Condition field:

```
national_id in
(select national_id from hr_test.employees where country_id = 'UK')
```

Click Import Format.

7 . On the Import Format page, select National Insurance Number Formatted. Click Import.

8 . On the Define Column Mask page, click the icon in the Sample column to view sample data.

9 . On the Define Column Mask page, click Add Condition.

10 . On the Define Column Mask page, enter the following SQL query in the Condition field:

```
national_id in
(select national_id from hr_test.employees where country_id = 'US')
```

Click Import Format.

11 . On the Import Format page, select Social Security Number Formatted. Click Import.

12 . On the Define Column Mask page, click the icon in the Sample column to view sample data.

13 . On the Define Column Mask page, select Default Condition. Select Preserve Original Data in the Format Entry list. Click Add.

14 . On the Define Column Mask page, click OK.

15 . On the Edit Masking Definition: HR Employee Mask page, click OK.

16 . On the Data Masking Definitions page, observe that seven columns are defined for masking.

Performing the Data Masking Operation

Use Enterprise Manager Database Control to generate the data-masking script and schedule the data masking job.

1 . On the Data Masking Definitions page, select HR Employee Mask. Click Generate Script.

2 . The Processing: Generating Data Masking Script page is displayed.

3 . A message is displayed indicating that the script has been generated.

4 . Scroll down the Script Generation Results: HR Employee Mask page. Expand Impact Report.

5 . View the Impact Report and verify that there are no errors. Click Schedule Job.

6. On the Schedule Data Masking Job: HR Employee Mask page, enter the host credentials. Select Immediately in the Start section. Click Submit.

7. On the Data Masking Definitions page, a message is displayed indicating that the job has been submitted. Click View Job Details.

8. On the Job Run: MASKING_JOB_NNN page, verify that the Status is Succeeded. Click Database to return to the Database Home page.

Querying Masked Data

Now view the data that was masked and compare the results with the unmasked data.

1. Click the Schema tab on the Database Instance Home page.

   Note: If you are using Enterprise Manager **Grid Control**, navigate to the Database Instance Home page by clicking Targets tab > Databases tab > Oracle database SID link.

2. Click Tables in the Database Objects region.

3. Enter the following information:

   Schema: `hr_test`

   Object Name: `employees`

   Click Go.

4. Select View Data in the Actions list. Click Go.

5. View the masked data in the `HR_TEST.EMPLOYEES` table. Click `EMAIL` to sort the rows on the `EMAIL` column.

6. Compare the values in the columns with the values you viewed prior to masking. Refer to your other Enterprise Manager Database Control window for the comparison.

7. Click OK to return to the Tables page.

8. Click the Database tab to return to the Database Instance Home page.

Removing Objects Created During this Tutorial

To clean up following this tutorial, perform the following steps.

1. Return to your SQL*Plus window. Logged in to SQL*Plus as the `SYSTEM` user, execute the `OBE_DM_cleanup.sql` script.

2. Return to Enterprise Manager Database Control. Click the Schema tab. Click Definitions in the Data Masking section. On the Data Masking Definitions page, select the HR Employee Mask and click Delete.

3. The HR Employee Mask masking definition is deleted from the Enterprise Manager Database Control repository.

Summary

In this tutorial, you have learned how to:

Use Oracle-supplied masking formats from the Format Library
Create masking formats
Create masking definitions
Generate the masking script
Schedule the data masking job

Resources

Oracle Enterprise Manager Concepts 10g Release 5 (10.2.0.5) B31949-10
Oracle Database 11g: Security course (D50323GC20)

Credits

Lead Curriculum Developer: Donna Keesling
Other Contributors: Jagan Athreya, James Spiller, Jenny Tsai-Smith

---

User Comments

Title:

Please log in or register first Post anonymously

By submitting a comment, you confirm that you have read and agreed to the terms and conditions.

This feedback form is for tutorial corrections and suggestions. Because of the volume of suggestions, we cannot reply to every comment. In particular:

If you have general questions about this Oracle software product, consult the OTN forums instead.
If you are encountering a software problem and need to request support or consulting, file a service request on My Oracle Support.
If you want to order hardcopy documentation, go to the Oracle Store.

Submit a new comment (Comments are moderated and will not appear immediately.)