



Wie sehen Ihre Testdaten aus? Anonymisierung mit Oracle Data Masking

Dirk Braunecker, CGI Deutschland
DOAG Konferenz, 21.11.2013

Referent: Dirk Braunecker



Oracle Solution Architect BI,
CGI Deutschland

Schwerpunkte

- Projektmanagement
- Technisches Design & Architektur von BI & DWH Lösungen
- Implementierung von Prozessen & Anwendungen im Bereich BI & DWH
- Oracle RDBMS, OWB, ODI, BIEE & Essbase



Inhalte

1	Data Masking - Herausforderung
2	Data Masking - Übersicht
3	Data Masking - Vorgehen
4	Data Masking mit Oracle (Übersicht / Beispiel)
5	Zusammenfassung



Data Masking Herausforderung

- Testdaten können z. B. folgende sensible Daten enthalten:
 - personenbezogene Daten (Kontodaten, Gehaltsdaten, Adressdaten; Arbeitszeiten, Geodaten etc.)
 - sensitive Unternehmensdaten (Kapitalmarkt relevante (Basis-) Daten, wettbewerbsrelevante Daten, Geschäftsgeheimnisse etc.)
- Aus gesetzlichen Datenschutzgründen müssen diese Daten vor einer Nutzung in einer Entwicklungs- oder Testumgebung anonymisiert oder entfernt werden:

§ 3, Abs. 6 Bundesdatenschutzgesetz (BDSG): „Anonymisieren ist das Ver-ändern personenbezogener Daten derart, dass die Einzelangaben über per-sön-liche oder sachliche Verhältnisse nicht mehr oder nur mit einem un-ver-hältnis-mäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder be-stimmbaren natürlichen Person zugeordnet werden können.“



Data Masking Übersicht

- Unter Anonymisierung wird eine Prozedur verstanden, die personenbezogene Daten irreversibel verändert → kein Personenbezug mehr vorhanden.
- Eine Wiederherstellung dieses Personenbezugs ist unmöglich.
- Anonymisierte Testdaten sind üblicherweise „maskierte“ Produktiv-Daten

Anonymisierung:

- Ersetzen durch Konstanten
- Verfremdung durch einen Algorithmus
- Löschen von Teilen der Daten
- Ersetzen durch Zufallswerte
- Ersetzen durch NULL-Werte (evtl. Probleme mit Fremdschlüsseln)



Data Masking Vorgehen

1

- Identifikation des Ursprungs der Testdaten

2

- Identifikation von sensiblen schützenswerten Daten

3

- Definition der Vorgehensweise zur Anonymisierung der Testdaten (Art des Data Maskings)

4

- Definition der Bereitstellung der anonymisierten Testdaten



Data Masking mit Oracle - Übersicht

Übersicht Oracle Data Masking:

- Vollständige Integration in den Oracle Enterprise Manager
- Grundsätzlich keine Quellcode-Entwicklung zur Anonymisierung notwendig
- Separate Lizenzierung der Option erforderlich

Technische Features:

- Spaltenanonymisierung in Echtzeit
- Verschiedene Arten der Anonymisierung (z. B. Ersetzen durch Konstanten)
- Einbeziehung der Fremdschlüssel
- Klonen der anonymisierten Daten in eine separate Datenbank



Data Masking mit Oracle - Übersicht

Vorgehensweise Oracle Data Masking



Vorteile:

- Zentrale Administration
- Zentrale Verantwortlichkeit



Data Masking mit Oracle - Beispiel

ORACLE Enterprise Manager 11g
Database Control

Database Instance

[Home](#) [Performance](#) [Server](#) **Schema** [Data Movement](#) [Software and Support](#)

Database Objects

[Tables](#)
[Indexes](#)
[Views](#)
[Synonyms](#)
[Sequences](#)
[Database Links](#)
[Directory Objects](#)
[Reorganize Objects](#)

Change Management

[Dictionary Baselines](#)
[Dictionary Comparisons](#)
[Dictionary Synchronizations](#)

XML Database

[Configuration](#)
[Resources](#)
[Access Control Lists](#)
[XML Schemas](#)

Programs

[Packages](#)
[Package Bodies](#)
[Procedures](#)
[Functions](#)
[Triggers](#)
[Java Classes](#)
[Java Sources](#)

Data Masking

[Definitions](#)
[Format Library](#)

Workspace Manager

[Workspaces](#)

Start / Aufruf



Data Masking mit Oracle - Beispiel

Menüoption „Definitions“

Confirmation

Masking Definition Salary has been deleted successfully.

Data Masking Definitions

Data masking is the process of making sensitive information in test or non-production databases safe. It disguises sensitive information by overwriting realistic looking but false data of a similar type. A masking definition defines the columns to be masked and the format of masked data. You can create a new masking definition or use an existing definition for a masking operation. The Format Library contains a collection of ready-to-use masking formats.

Search

Select	Masking Definition	Database	Description	Columns	Status	Most Recent Job Ended
	No definitions					

Format Library

A masking format defines the format of masked data. You can create a new masking format and reuse it later when creating a masking definition.

[Format Library](#)

[Database](#) | [Setup](#) | [Preferences](#) | [Help](#) | [Logout](#)


Copyright © 1996-2010 Oracle. All rights reserved.



Data Masking mit Oracle - Beispiel

Definition erstellen

Create Masking Definition

 Name

* Database

Description

Columns

Add columns you want to mask and define masking format for each column. Foreign key columns are automatically added to maintain referential integrity. Dependent columns are columns that do not have foreign key constraints defined, but reference a masked column due to application constraints. You can manually add dependent columns to a masked column. Removing a column from this list will remove all foreign key and dependent columns.

Select	Owner	Table	Column	Column Group	Data Type	Format	Foreign Key Columns	De C Co
	No columns added							

Foreign Key Columns

Owner	Table	Column	Parent Owner	Parent Table	Parent Column
No foreign key columns					



Data Masking mit Oracle - Beispiel

Tabelle und Spalten auswählen

ORACLE Enterprise Manager 11g Database Control

Setup

Data Masking Definitions > Create Masking Definition >

Add Columns

Database **orcl**

Logged In As **SYS**

Cancel

Add

Add one or more columns for masking. Foreign key columns will be added automatically. You can define masking format at once for all selected same data type.

Search

Schema

Table Name



Column Name

Column Comment

Enter a string in column comments.

Mask selected columns as a group

Select	Owner	Table Name	Column Name	Data Type	Comment
	No columns				



Data Masking mit Oracle - Beispiel

Anonymisierung für Spalte festlegen

[Data Masking Definitions](#) > [Create Masking Definition](#) >

Define Column Mask

Owner **HR**
Column **SALARY**

Table **EMPLOYEES**
Data Type **NUMBER(8,2)**

By default all records in the table will be masked using the specified format. You can optionally identify more than one subset of records using conditions. A subset of records can be masked using a corresponding masking format. The subsets will be masked in the order they are specified. A subset will not be masked again if it is specified in a subsequent condition.

Import Format Format Entry Fixed Number Add

[Expand All](#) | [Collapse All](#)

Select	Condition	Format Entry Properties				Sample
		Property	Value	Property	Value	
	▼ Conditions					
<input checked="" type="radio"/>	▼ Default Condition					
	Fixed Number	Fixed Number	<input type="text" value="4711"/>			

[Database](#) | [Setup](#) | [Preferences](#) | [Help](#) | [Logout](#)



Data Masking mit Oracle - Beispiel

Definition erstellt

ORACLE Enterprise Manager 11g
Database Control

Data Masking Definitions

Data masking is the process of making sensitive information in test or non-production databases safe. It disguises sensitive information but false data of a similar type. A masking definition defines the columns to be masked and the format of masked data. You can create a new masking definition or edit an existing definition for a masking operation. The Format Library contains a collection of ready-to-use masking formats.

Search Database "orcl" Go

View	Edit	Create Like	Generate Script	Schedule Job	Export	Clone Database	Delete
Select Masking Definition	Database	Description	Columns	Status			
<input checked="" type="radio"/> SALARY_DEFINITION	orcl	Anonymisierung des Gehalts	1	Script Not Generated			

Format Library

A masking format defines the format of masked data. You can create a new masking format and reuse it later when creating a new masking definition. For more information, see [Format Library](#).



Data Masking mit Oracle - Beispiel

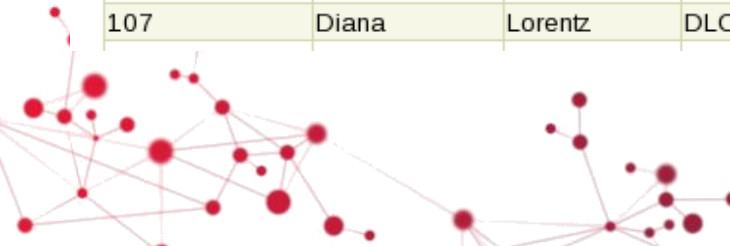
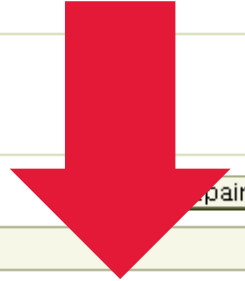
Ergebnis der SQL-Abfrage nach erfolgter Anonymisierung

Last Executed SQL

```
SELECT * FROM hr.employees
```

Last Execution Details

Results	Statistics	Plan						
Execution Time (seconds) 0.029								
EMPLOYEE_ID	FIRST_NAME	LAST_NAME	EMAIL	PHONE_NUMBER	HIRE_DATE	JOB_ID	SALARY	COMM
198	Donald	OConnell	DOCONNEL	650.507.9833	Jun 21, 1999 12:00:00 AM	SH_CLERK	4711	
199	Douglas	Grant	DGRANT	650.507.9844	Jan 13, 2000 12:00:00 AM	SH_CLERK	4711	
200	Jennifer	Whalen	JWHALEN	515.123.4444	Sep 17, 1987 12:00:00 AM	AD_ASST	4711	
201	Michael	Hartstein	MHARTSTE	515.123.5555	Feb 17, 1996 12:00:00 AM	MK_MAN	4711	
202	Pat	Fay	PFAY	603.123.6666	Aug 17, 1997 12:00:00 AM	MK_REP	4711	
203	Susan	Mavris	SMAVRIS	515.123.7777	Jun 7, 1994 12:00:00 AM	HR_REP	4711	
204	Hermann	Baer	HBAER	515.123.8888	Jun 7, 1994 12:00:00 AM	PR_REP	4711	
205	Shelley	Higgins	SHIGGINS	515.123.8080	Jun 7, 1994 12:00:00 AM	AC_MGR	4711	
206	William	Gietz	WGIETZ	515.123.8181	Jun 7, 1994 12:00:00 AM	AC_ACCOUNT	4711	
100	Steven	King	SKING	515.123.4567	Jun 17, 1987 12:00:00 AM	AD_PRES	4711	
101	Neena	Kochhar	NKOCHHAR	515.123.4568	Sep 21, 1989 12:00:00 AM	AD_VP	4711	
102	Lex	De Haan	LDEHAAN	515.123.4569	Jan 13, 1993 12:00:00 AM	AD_VP	4711	
103	Alexander	Hunold	AHUNOLD	590.423.4567	Jan 3, 1990 12:00:00 AM	IT_PROG	4711	
104	Bruce	Ernst	BERNST	590.423.4568	May 21, 1991 12:00:00 AM	IT_PROG	4711	
105	David	Austin	DAUSTIN	590.423.4569	Jun 25, 1997 12:00:00 AM	IT_PROG	4711	
106	Valli	Pataballa	VPATABAL	590.423.4560	Feb 5, 1998 12:00:00 AM	IT_PROG	4711	
107	Diana	Lorentz	DLORENTZ	590.423.5567	Feb 7, 1999 12:00:00 AM	IT_PROG	4711	



Zusammenfassung Data Masking mit Oracle

Erste Schritte in Data Masking Projekten:

- Erstellung Testdaten-Anonymisierungskonzeptes / Vorgehensmodells
- Erstellung Attributkatalog mit Zuordnung wie welche Attribute zu anonymisieren / pseudonymisieren sind
- Einarbeitung der unternehmensspezifischen Vorgaben zu Data Masking
- Abgleich dieser Vorgaben mit den gesetzlichen Regelungen
- Start mit der Umsetzung

Besonderheiten von Oracle Data Masking:

- Einfache Handhabung des Tools





Ich freue
mich auf Ihre
Kommentare &
Fragen!

Dirk Braunecker
Oracle Solution Architect BI

Mobile: +49 151 42 23 10 10
E-Mail: dirk.braunecker@cgi.com