



BERENBERG

PRIVATBANKIERS SEIT 1590

Data Masking in der Praxis

Frank Hilgendorf, Berenberg

Agenda

- Vorstellung Berenberg
- Was ist Data Masking?
- Oracle Data Masking Pack
- Data Masking bei Berenberg
- Fazit



Berenberg – At a Glance



BERENBERG

PRIVATBANKIERS SEIT 1590

Berenberg wurde 1590 gegründet und gehört heute mit den Geschäftsbereichen Private Banking, Investment Banking, Asset Management und Corporate Banking zu den führenden europäischen Privatbanken.

- **Sitz**
Hamburg
- **Standorte**
Hamburg, London, Frankfurt, Zürich, New York, Bielefeld, Braunschweig, Bremen, Düsseldorf, München, Stuttgart, Boston, Chicago, Genf, Luxemburg, Paris, San Francisco, Shanghai, Wien
- **Verwaltetes Vermögen**
> 30 Mrd. Euro
- **Mitarbeiter**
Insgesamt: ~ 1200
IT: ~ 160
- **Persönlich haftende Gesellschafter**
Dr. Hans-Walter Peters (Sprecher), Andreas Brodtmann, Hendrik Riehmer





Was ist Data Masking?

Ziel

- Anonymisieren von (Test) Daten unter Beibehaltung realistischer Eigenschaften
- Testdatenmanagement

Grund

- Einrichtung / Verbesserung präventiver Schutzmaßnahmen
- Schutz realer Daten vor nicht autorisierten Zugriffen
- ungenügend gesicherte Test- und Entwicklungssysteme

Methoden & Algorithmen

- Erzeugen neuer, eindeutiger Werte
- Mischen / Shuffle bestehender Werte
- Einsatz von Konstanten

- Zufallszahlen
- Zufällige Zeichenketten / Datumswerte
- Benutzerdefinierte Funktionen (PL/SQL)

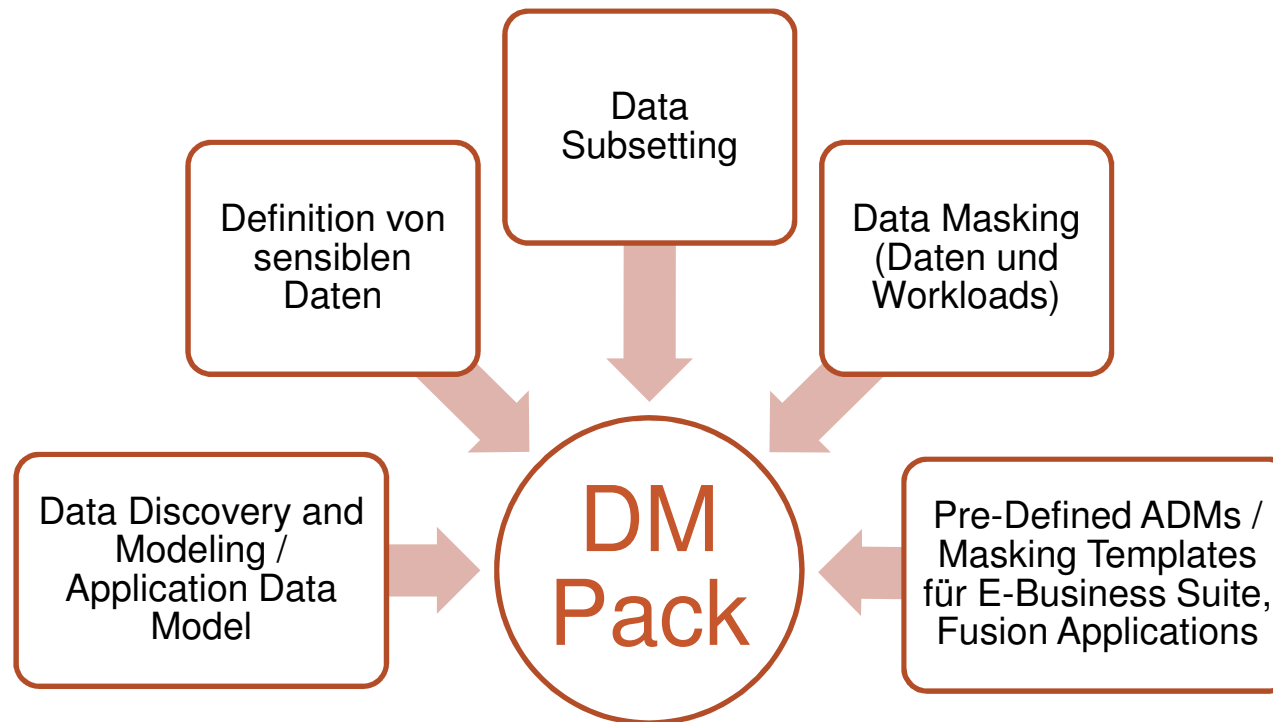




Oracle Data Masking Pack

Oracle Data Masking Pack for Oracle and non-Oracle Databases

- ist im Bundle mit dem Oracle Enterprise Manager erhältlich
- Lizenzpflichtig
- enthält Komponenten für das Testdatenmanagement

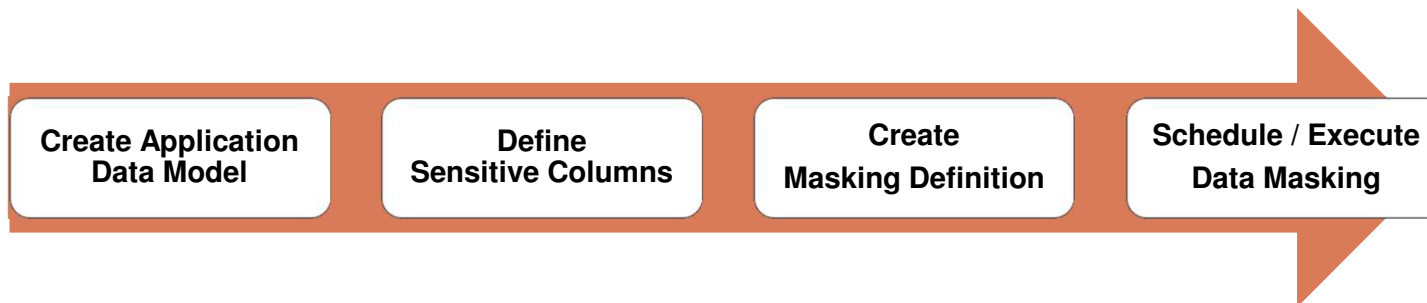




Data Masking Prozess

Nicht alle Komponenten des DM Packs sind für einen Data Masking Prozess notwendig.

- Data Subsetting
- Maskieren von Workloads
- Pre-Defined ADMs





BERENBERG

PRIVATBANKIERS SEIT 1590

Application Data Model

- ADM stores the list of applications, tables and relationships between table columns
- ADM maintains sensitive data types and their associated columns
- ADM is used by Secure Test Data Management features such as Data Masking and Data Subsetting, to produce test data securely
- Speichert die Informationen in einem zentralen Repository
- Spezielle Unterstützung für Fusion Applications und E-Business Suite





Application Data Model

- Create ADM

Create Application Data Model

Application Data Model Properties: General

* Name

Description

* Source Database

Application Suite

Type Custom Application Suite

Create One Application For Each Schema

Oracle Application Suite

Oracle E-Business Suite

Oracle Fusion Applications

- Define / Add Sensitive Columns

Add Sensitive Columns

Application

Table

Column

| Application | Table | Column | Column Comment |
|-------------|-----------|------------|---|
| HR | EMPLOYEES | FIRST_NAME | First name of the employee. A not null column. |
| HR | EMPLOYEES | EMAIL | Email id of the employee |
| HR | EMPLOYEES | HIRE_DATE | Date when the employee started on this job. A not nul |





Application Data Model

ORACLE Enterprise Manager Cloud Control 12c

The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface. The 'Data Discovery and Modeling' menu is open, highlighting 'Data Discovery and Modeling'. Below it, the 'Open Application Data Model: DOAG' window is displayed, showing a table of data models.

| Name | Short Name | Schema | Table Type | Source |
|-------------|------------|--------|------------------|------------|
| HR | HR | HR | | Dictionary |
| COUNTRIES | | HR | Transaction Data | Dictionary |
| DEPARTMENTS | | HR | Transaction Data | Dictionary |
| EMPLOYEES | | HR | Transaction Data | Dictionary |
| JOBS | | HR | Transaction Data | Dictionary |
| JOB_HISTORY | | HR | Transaction Data | Dictionary |
| LOCATIONS | | HR | Transaction Data | Dictionary |
| REGIONS | | HR | Transaction Data | Dictionary |
| VBES | | HR | Transaction Data | Dictionary |





Data Masking Definition

- Erstellung beginnt mit Selektion eines ADM
 - Nur im ADM enthaltene Spalten können maskiert werden
 - Die referenzielle Integrität wird durch das automatische Hinzufügen von Foreign Key Columns gewährleistet
- Definition Masking Format
 - Seit EM12c ist auch reversibles Masking möglich
- Auswahl div. Optionen für Performance und Security

Data Masking Options

Disable redo log generation during masking

Refresh statistics after masking

Drop temporary tables created during masking

Decrypt encrypted columns

Use parallel execution when possible

Parallel Degree Default Value

Recompile invalid dependent objects after masking

Serial

Parallel

Degree Default Value

Random Number Generation

Data Masking can use strong security techniques for random number generation, but masking will run slower. Do you want masking to favor speed or security?

Favor Speed

Favor Security

Array List

Array List

Delete

Encrypt

Fixed Number

Fixed String

Null Value

Post-Processing Function

Preserve Original Data

Random Decimal Numbers

Random Digits

Random Numbers

Random Strings

Shuffle

SQL Expression

Substitute

Substring

Table Column

Truncate

User Defined Function



Data Masking Prozess

Maskieren von Testdaten

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 it with 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

Actions

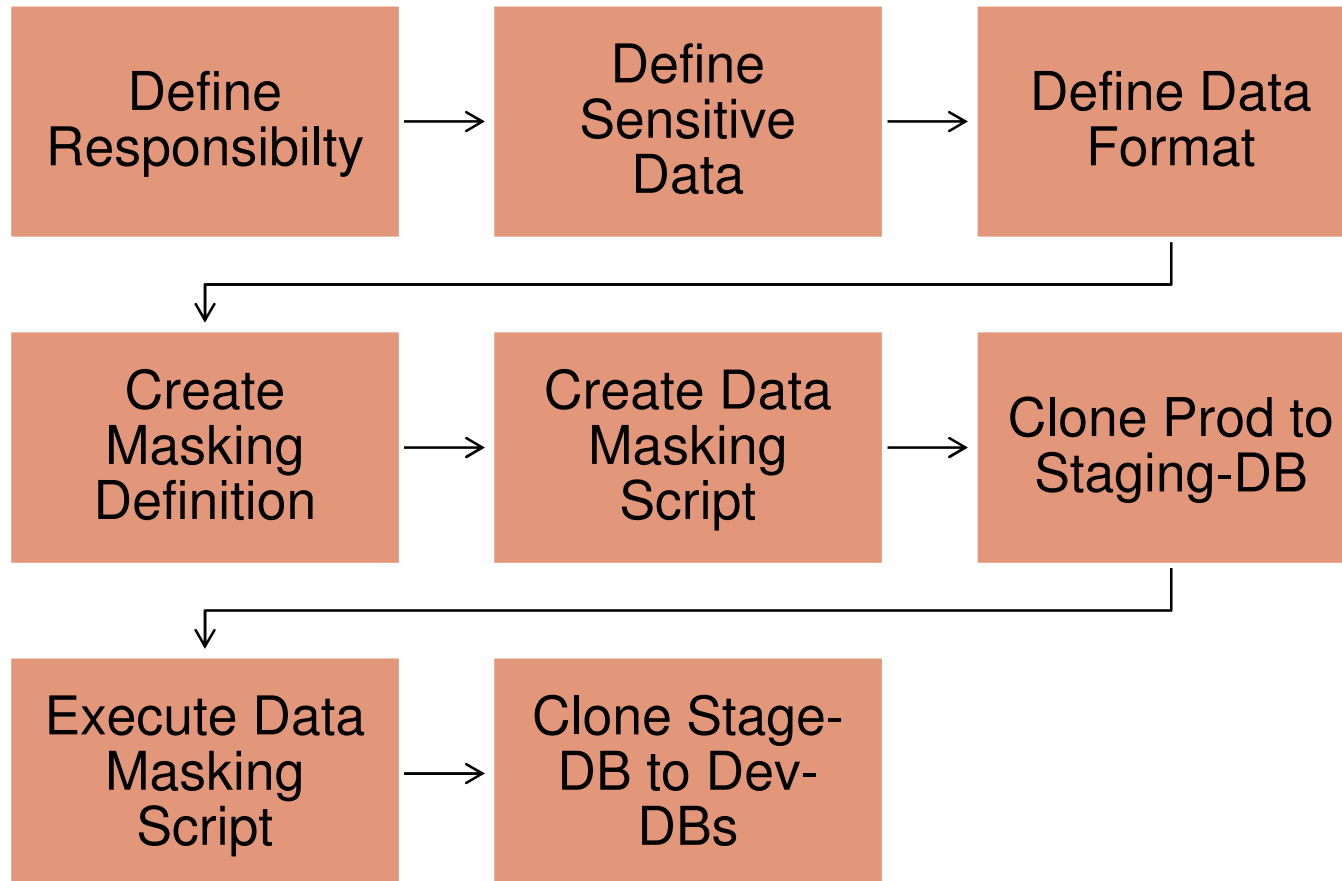
| Select | Masking Definition ▲ | Application Data Model | Description | Columns | Status | Most Recent Job Ended |
|----------------------------------|------------------------|------------------------|-------------|---------|----------------------|-----------------------|
| <input checked="" type="radio"/> | AAA_MASKING_DEF_104374 | DOAG | DOAG Demo | 1 | Script Not Generated | |

- EM GUI
 - Schedule Job
- SQLPLUS
- EM CLI
 - Enterprise Manager Command Line Interface





Data Masking Prozess





Data Masking in der Praxis

- bei Berenberg im Einsatz seit Enterprise Manager 10G
- aktuelle Version EM CC 12.1.0.2.0, DB Plugin 12.1.0.3.0
- tägliche Maskierung von 4 Staging-DBs als Basis für 30+ Dev-DBs
- 445 Spalten in 90 Tabellen, > 12 GB Daten
Kernbanksysteme, Portfoliomanagement, CRM, DWH, etc.
- vollständige Automatisierung der täglichen Data Masking Jobs über
Job Scheduler und EM CLI
- DM Administration beschränkt sich auf die EM GUI
 - Anpassen des ADM inkl. sensitive Columns
(automatisches Discovery funktioniert nur für wenige vordefinierte
Datenformate)
 - Ändern der Data Masking Definition(s)

Create Sensitive Column Discovery Job

Parameters

Select All/ Select None

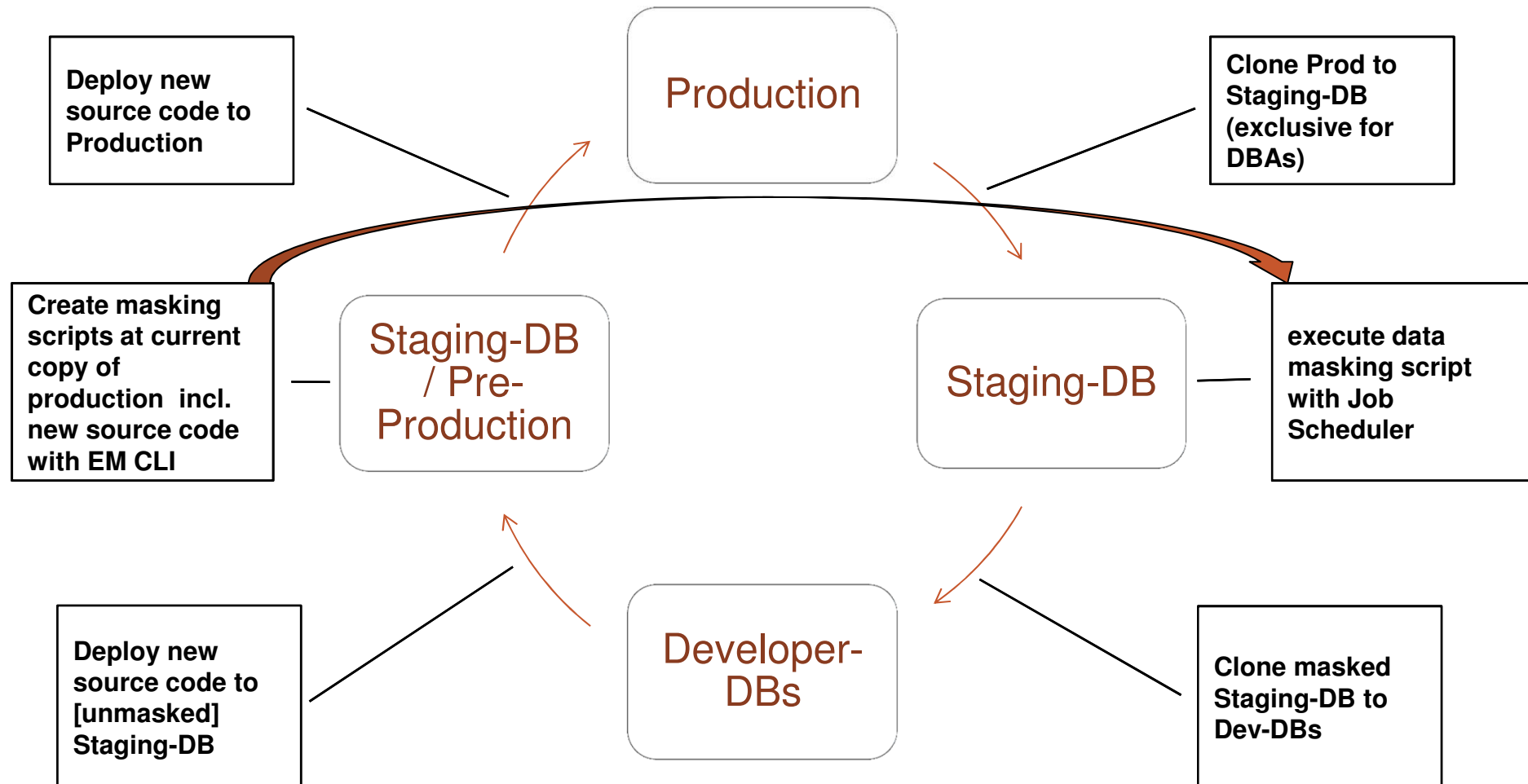
| Application |
|-----------------------------|
| <input type="checkbox"/> HR |

Select All/ Select None

| Sensitive Column Type |
|--|
| <input type="checkbox"/> <input checked="" type="checkbox"/> CREDIT_CARD_NUMBER |
| <input type="checkbox"/> <input checked="" type="checkbox"/> EMAIL_ID |
| <input type="checkbox"/> <input checked="" type="checkbox"/> IP_ADDRESS |
| <input type="checkbox"/> <input checked="" type="checkbox"/> ISBN_10 |
| <input type="checkbox"/> <input checked="" type="checkbox"/> ISBN_13 |
| <input type="checkbox"/> <input checked="" type="checkbox"/> NATIONAL_INSURANCE_NUMBER |
| <input type="checkbox"/> <input checked="" type="checkbox"/> PHONE_NUMBER |
| <input type="checkbox"/> <input checked="" type="checkbox"/> SOCIAL_INSURANCE_NUMBER |
| <input type="checkbox"/> <input checked="" type="checkbox"/> SOCIAL_SECURITY_NUMBER |
| <input type="checkbox"/> <input checked="" type="checkbox"/> UNIVERSAL_PRODUCT_CODE |



Data Masking in der Praxis





Data Masking in der Praxis - Remarks

- Mitarbeitersensibilisierung ist erforderlich
- Zuständigkeiten müssen geklärt sein
- Die erzeugten SQL Skripte sind sehr umfangreich
 - Jeder SQL Befehl wird in einer separaten Prozedur ausgeführt
Bsp.: Für 1 Tabelle mit 3 Spalten wird für 300+ Prozeduren create, execute, drop ausgeführt
 - SQL-Dateien können sehr groß werden, z.B. 40.000 Prozeduren, >60MB
 - Der Prozess erzeugt sehr viel I/O
 - je nach Umfang und I/O Subsystem kann die Laufzeit ein Problem werden



Data Masking in der Praxis - Remarks

- Keine Fehlertoleranz vorhanden
 - Ein Fehler in einer Prozedur stoppt den ganzen Prozess, z.B. „ORA-04052: error occurred when looking up remote object“ oder „ORA-01927: cannot REVOKE privileges you did not grant“
 - !! Fehlertoleranz sollte nachträglich hergestellt werden !! (SQL-Script per awk, sed, etc. bearbeiten)
 - Statt Abbruch wird ein Fehlerreport erzeugt und per E-Mail verschickt

```
CREATE OR REPLACE PROCEDURE mgmt$step_262_104074(script_id IN INTEGER, job_table IN
VARCHAR2, step_num IN OUT INTEGER)
AUTHID CURRENT_USER IS
  sqlerr_msg VARCHAR2(100);

BEGIN
  IF step_num <> 262 THEN
    return;
  END IF;
  ...
```





Data Masking in der Praxis - Remarks

- VPDs und Virtuals Columns sind nach dem Prozess nicht mehr vorhanden, müssen eventuell neu angelegt werden
- Die Verteilung der erzeugten Zufallswerte ist gut, außer bei „Random Strings“
 - 50% der Strings beginnen mit „aa%“ oder „aaa%“

Import Format Format Entry Array List Add

Expand All | Collapse All

| Select | Condition | Format Entry Properties | | | | Sample |
|----------------------------------|---------------------|-------------------------|-------|------------|-------|-------------------|
| | | Property | Value | Property | Value | |
| | ▼ Conditions | | | | | |
| <input checked="" type="radio"/> | ▼ Default Condition | | | | | aaaaaaaaadrvillee |
| | Random Strings | Start Length | 1 | End Length | 15 | |
| <input checked="" type="radio"/> | ▼ Default Condition | | | | | aaaaaaaaacfobld |
| <input checked="" type="radio"/> | ▼ Default Condition | | | | | aaajypdcnlhpxlb |
| <input checked="" type="radio"/> | ▼ Default Condition | | | | | aaaaaaaaaacgzqc |





Data Masking in der Praxis - Remarks

- Es gibt keine sicherheitsrelevante Zertifizierung für das Data Masking Pack
-> Oracle kann den richtigen Einsatz **nicht** garantieren
- Das Produkt wird von Oracle noch weiterentwickelt

Database Plug-in 12.1.0.5 Features

Masking Enhancements for Fusion Middleware, E-Business Suite, and PeopleSoft Applications

- **Enabling and Disabling Sensitive Columns in Masking Definitions**

With this new release of data masking, users can simply disable (or enable) the sensitive columns and regenerate the masking script

- **Filtering Sensitive Columns in Masking Definitions**
- **Support for Masking CLOB and BLOB Columns**

users can search for multiple sensitive data within the CLOB and BLOB columns and replace just the sensitive data with obfuscated data leaving the data structure intact.

- **Fine Grained Privilege Control of Data Masking Definitions**





Oracle Data Masking Pack for Oracle and non- Oracle Databases

- Masking Templates für E-Business Suite, Fusion Applications
- Data Discovery and Modeling Discover
- Data Masking Mask
- Export or import format
- Sensitive Column Types and Sensitive Data Discovery

Oracle Test Data Management Pack for Oracle and non- Oracle Databases

- Application Accelerators (E-Business Suite, Fusion Applications)
- Data Discovery and Modeling
- Data Subset Definitions



Data Masking in der Praxis - Fazit

- Ready to use...
- Sehr gute Performance in dynamischen Umgebungen
- Administrationsaufwand ist sehr gering, hängt nur von der Häufigkeit der Erstellung/Bearbeitung relevanter Objekte ab
- Anforderungen können schnell umgesetzt werden
- Der Prozess erzeugt nur wenige Fehler, mit zusätzlicher Fehlertoleranz ist Data Masking sehr stabil
- Neue Mitarbeiter können nach kurzer Einweisung sofort mit dem Tool arbeiten
- Der Aufwand für Entwicklung und Produktpflege einer Eigenentwicklung wäre deutlich höher





Danke für Ihre Aufmerksamkeit!

Frank Hilgendorf

BERENBERG

Joh. Berenberg, Gossler & Co. KG

Neuer Jungfernstieg 20

20354 Hamburg

Telefon +49 40 350 60-192

Telefax +49 40 350 60-980

E-Mail frank.hilgendorf@berenberg.de

www.berenberg.de