COMPUTERWOCHE
VOICE OF DIGITAL

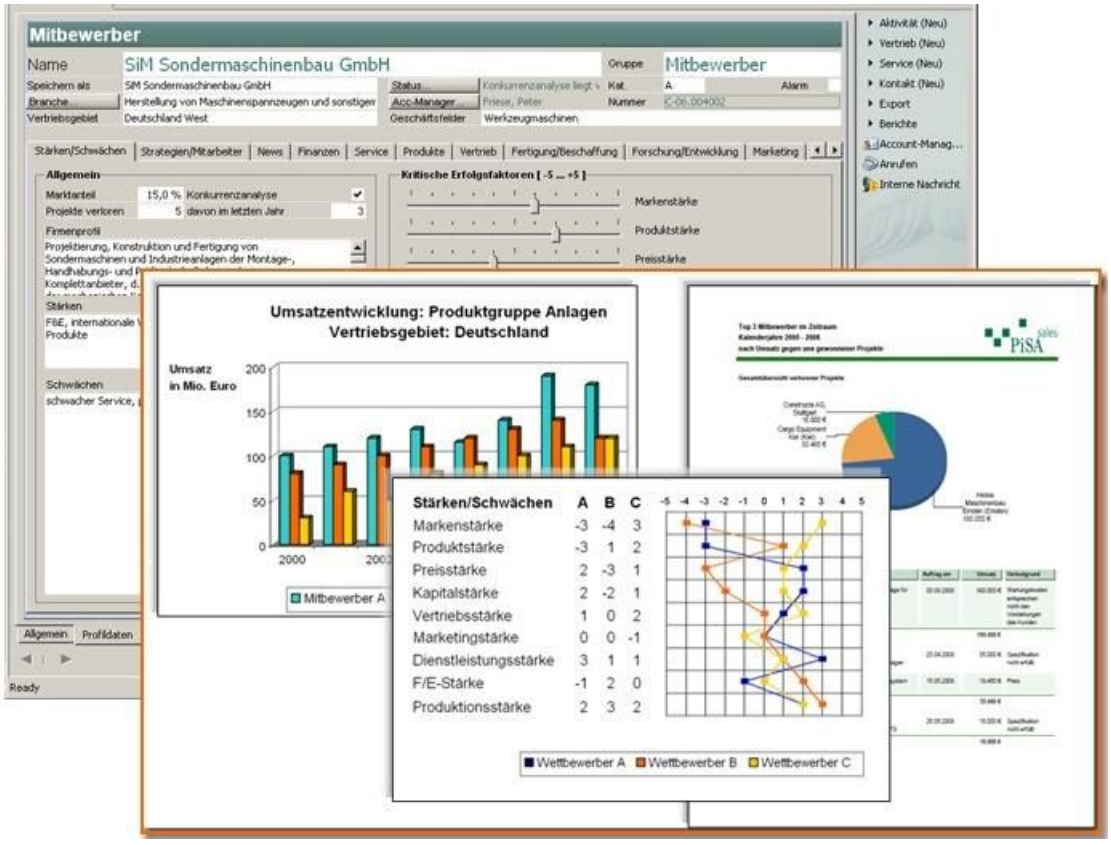
Link: https://www.computerwoche.de/a/geheime-daten-sind-fuer-alle-da,1904321

**Micro Focus-Studie zu Data Masking**

# Geheime Daten sind für alle da

**Datum:** 02.09.2009
**Autor(en):Thomas Pelkmann**

**Die meisten Unternehmen verwenden in großem Umfang vertrauliche Originaldaten für Software-Tests. Das könnte man leicht als Einladung zum Missbrauch wertvoller Firmeninformationen verstehen. Wir zeigen Ihnen, wie Sie sich davor schützen.**




Zur Migration von Datenbanken gehört auch ein sorgfältiges Maskieren der Daten, um Missbrauch vertraulicher Informationen zu verhindern.

Wer für sein Unternehmen die **Migration**[1] oder ein Upgrade von Datenbankanwendungen plant, muss ausführlich testen, ob die neue Lösung mit allen Anforderungen zurechtkommt, bevor sie den Mitarbeitern im Alltag zur Verfügung gestellt werden kann.

Das geht im Prinzip am besten mit Originaldaten, weil einzig diese Informationen optimale Wirklichkeitstreue versprechen. Klug ist so eine Operation am offenen Herzen aber nicht: Landen Kunden-, Mitarbeiter- oder Finanzdaten etwa bei IT-Servicemitarbeitern, bei externen Outsourcing-Partnern oder gar bei temporär mit den Tests beschäftigten Aushilfskräften, kann man sie auch gleich in der Zeitung veröffentlichen. Abgesehen vom möglichen wirtschaftlichen Schaden verletzt ein solches Verhalten aber nicht zuletzt auch gesetzliche Bestimmungen zum Umgang mit vertraulichen Informationen.

Trotzdem gehen Unternehmen offenbar eher sorglos mit diesen Gefahren um. Eine internationale Studie von **Micro Focus**[2] zeigt, dass für die Tests bei Migrations- und Upgrade-Projekten in großem Umfang Originaldaten verwenden werden. Für die Untersuchung hatte das Ponemon-Institute 1.350 Softwareentwickler und -tester aus Unternehmen mit einem Jahresumsatz von mehr als zehn Millionen Dollar befragt.

Dabei kam unter anderem heraus, dass 70 Prozent der IT-Fachkräfte für die Entwicklung und das Testen von Software keine maskierten Daten verwenden, sondern Originale von Kunden und Mitarbeitern. In einer anderen Studie ermittelte Forrester, dass sogar nur ganze 15 Prozent sichere Testdaten verwenden.

Grund genug zur Vorsicht gäbe es allemal: 79 Prozent der von Micro Focus befragten Unternehmen hatten innerhalb der letzten zwölf Monate mindestens einen Fall von Datensicherheitsverletzung zu verzeichnen. Zu einem sorgfältigen Umgang mit vertraulichen Informationen führten diese Probleme aber offenbar nicht: Gerade einmal sieben Prozent der Befragten vertritt die Auffassung, dass der Schutz von Daten in Entwicklungs- und Testumgebungen besonders ernst genommen wird.

"Man scheint sich überhaupt nicht darüber im Klaren zu sein, dass diese Daten besonders gefährdet sind, beispielsweise durch ehemalige Mitarbeiter oder Zulieferer", kommentiert Rainer Downar, Country Manager von Micro Focus Central Europe, dieses alarmierende Ergebnis. "Es ist höchste Zeit, effektive Datenschutzmechanismen zum Beispiel durch Datenmaskierung umzusetzen."

Das Marktforschungsunternehmen **Forrester**[3] definiert **Data Masking**[4] als einen Prozess in nicht-produktiven Umgebungen, bei dem private Daten so verborgen werden, dass Entwickler, Tester, privilegierte Anwender und Outsourcing-Partner solche Informationen nicht im Klartext lesen oder entschlüsseln können. Zu diesen Daten gehören beispielsweise Sozialversicherungs- und Kreditkartennummern, Finanzdaten und Gesundheitsinformationen.

## Data Masking gehört zu den Sicherheits-Checks

Damit Data Masking nicht als lästiges Übel bei Migrationen und Upgrades gleich hinten runter fällt, empfehlen die Marktforscher von Gartner, diese Aufgabe gleich als wichtigen Teil des Projekts zu verstehen. Dann geht es vordergründig nicht mehr bloß um das Unkenntlichmachen von Informationen, sondern substanziell um das Testen von Sicherheit und **Compliance**[5] der neuen Lösung.

Wichtige Bedingung für das Data Masking: Die falschen Daten müssen genau so funktionieren, wie die echten. Sie dürfen die Testmöglichkeiten nicht einschränken und konsistent zu den real existierenden Informationen sein. Wie weit diese Konsistenz geht, ist Definitionssache: Für die einen mag es zum Beispiel für das Maskieren von Postleitzahlen ausreichen, irgendeine fünfstellige Zahl zu bekommen, während es für andere wichtig ist, eine Nummer zu erhalten, die der tatsächlichen möglichst nahe kommt.

Dafür ist es bedeutsam, so **Gartner**[6], dass man vor der Maskierung die Logik einer Anwendung und die in ihr abgebildeten Datenbezüge versteht. Das aber sollten Sie nicht allein Ihren IT-Mitarbeitern überlassen. Den besten Überblick über die Eigenheiten des Datenbestandes haben nämlich eher die Kollegen aus dem Business oder der Rechtsabteilung.

Für den konkreten Maskierungsprozess hält die Datenbankspezialistin **Andrea Held**[7] folgende Tipps bereit.

- Die Maskierung oder Anonymisierung der Daten muss irreversible sein. Es darf nicht möglich sein, sensible Original-Daten aus dem anonymisierten Datenbestand zurück gewinnen zu können.

- Das Ergebnis sollte repräsentativ für die Quell-Daten sein. Maskierung spiegelt im Gegensatz zu einem einfachen Randomize die Verteilung der Quelldaten wieder. Nur so ist die Konstanz der Ausführungspläne und damit der Datenzugriffe zwischen produktiven und Test- oder Entwicklungssystem gesichert. Das Verhalten der Datenbank kann bei einer abweichenden Verteilung der Daten bei Datenzugriffen ebenfalls abweichen.
- Die referentielle Integrität muss beibehalten werden. Master- und Detail-Daten müssen einander zugeordnet werden können.
- Nicht-sensitive Daten müssen nur dann maskiert werden, wenn aus ihnen sensible Daten abgeleitet werden können.
- Maskierung muss ein wiederholbarer Prozess sein. Entwicklungs- und Test-Daten müssen regelmäßig aus der produktiven Umgebung übernommen und erneut maskiert werden können. Nur so ist es möglich, Änderungen der produktiven Datenverteilung und damit des Verhaltens des Optimizers im Test- und Entwicklungssystem zu reproduzieren.

## Links im Artikel:

[1] **https://www.computerwoche.de/schwerpunkt/i/IT-Migration.html**
[2] **http://www.microfocus.com/**
[3] **http://www.forrester.com/rb/research**
[4] **http://en.wikipedia.org/wiki/Data_masking**
[5] **https://www.computerwoche.de/virtualdatacenter/sicherheit/news/1898018/**
[6] **http://www.gartner.com/it/regionalization/notice/de.jsp**
[7] **http://www.held-informatik.de/**

---