

## Übung 2

Prof. Dr. Alexandros Nanopoulos, Christoph Freudenthaler  
Wirtschaftsinformatik und Maschinelles Lernen (ISMLL)  
Universität Hildesheim

**Abgabe bis: 10.05.2009**

### **Aufgabe 1: (10 Punkte)**

Was ist der Unterschied zwischen den herkömmlichen, transaktionsorientierten Datenbanken und einem Data Warehouse? Wie sehen die unterschiedlichen Zielsetzungen aus? Wie die daraus ableitbaren Anforderungen und Architekturen? Warum werden Data Warehouses zusätzlich zu (und nicht gemeinsam mit) einer bereits bestehenden OLTP-Lösung verbunden?

### **Aufgabe 2: (10 Punkte)**

Was ist die 2-Schichten- und was die 3-Schichten-Architektur bei der Erstellung eines Data Warehouses? Erkläre unter welchen Umständen die 2-Schichten-Architektur der 3-Schichten-Architektur vorzuziehen ist.

### **Aufgabe 3: (5 Punkte)**

Beantworte die folgenden Aussagen mit *wahr* oder *falsch*:

1. Sobald die Daten einmal in das DWH eingegeben wurden, kann der Benutzer weder die Daten ändern noch löschen.
2. Ein unabhängiger Data Mart stellt sicher, dass alle Endbenutzer die gleiche Version der Datenwahrheit zu sehen bekommen.
3. Ein Operational Data Store (ODS) wird für mittel- und langfristige Entscheidungen im Zusammenhang mit dem zentralen Data Warehouse verwendet.
4. Die Daten eines ODS stammen aus den oper marts.
5. DWH beinhalten unglaubliche Mengen an internen Daten und beschränken sich daher auf diese. Externe Daten sind in der Regel nicht im DWH enthalten.
6. Die beliebtesten DWH-Architekturen sind die 2- und 3-Schicht-Architektur.
7. Bei der Entwicklung von web-basierten DWH-Architekturen wie der 4-Schicht- bzw. Multi-Schicht-Architektur kommt es unter anderem auf die Geschwindigkeit des Datenaustausches zwischen Web-Server und Browsern/Webseiten an.
8. Ein zentrales DWH (ohne Data Marts) unterscheidet sich vom Hub-and-Spoke-Ansatz nur dadurch, dass es keine unabhängigen Data Marts hat.
9. Aufgrund von Performanz- und Datenqualitätsgründen sind sich die meisten Experten einig, dass verteilte Strukturen (dh. ohne physikalisch eigenständigen Datenspeichern wie dem DWH oder Data Marts) geeignet sind, das zentrale DWH zu ersetzen.

## Aufgabe 4: (3 Punkte)

Beantworten Sie die folgenden Multiple-Choice Aussagen:

- Ein DWH ist eine \_\_\_\_\_, integrierte, zeit-variante, nicht-volatile Datenansammlung zur Unterstützung von Managemententscheidungen.
  - Analyse orientiert
  - Objekt orientiert
  - Subjekt orientiert
  - Modell orientiert
- Sobald die Daten im DWH sind, können sie nicht mehr geändert werden. Irrelevante Daten werden entfernt und Änderungen werden wie neue, eigenständige Einträge gesehen. Diese Eigenschaft eines DWH nennt man:
  - veränderbar
  - nicht volatil
  - unverdorben
  - statisch
- Die hohen Kosten für ein DWH begrenzt deren Einsatzgebiet auf große Unternehmen. Als eine Alternative verwenden viele Firmen daher billigere, kleinere Versionen des DWH, die so bezeichnet werden:
  - Data Mart
  - Operational Data Store
  - Dependent Data Mart
  - Independent Data Mart
- Welche der folgenden DWH-Komponenten greift für den User auf die Datenbestände des DWH zu.
  - Middleware
  - Benutzeroberfläche
  - Abfragesprachen wie SQL
  - OLAP
- \_\_\_\_\_ haben inkonsistente Datendefinitionen und unterschiedliche Dimensionen und Maße, die es schwierig machen, Daten zwischen den einzelnen Data Marts zu analysieren.
  - Data Marts
  - Operational Data Marts
  - abhängige Data Marts
  - unabhängige Data Marts

## Aufgabe 5: (2 Punkte)

Füllen Sie richtig aus:

- Sobald DWH in ihrer Größe wachsen, entsteht das Problem der \_\_\_\_\_.
- \_\_\_\_\_ ist ein Ausschnitt aus einem DWH, der den Vorteil genießt, für einzelne Abteilungen maßgeschneidert zu sein, gleichzeitig aber keinen eigenen Datenintegrations- oder -bereinigungsschritt benötigt.
- Ein \_\_\_\_\_ ist ein für unternehmensübergreifende Entscheidungen konzipiertes DWH.
- Ein \_\_\_\_\_ beinhaltet die aktuellsten Kundeninformationen zur Analyse. Meistens wird er als Intermediär zwischen einem klassischen und einem Echtzeit-DWH (real-time DWH) verwendet.