



Übung zur Vorlesung *Grundlagen: Datenbanken* im WS13/14

Henrik Mühe (muehe@in.tum.de)

<http://www-db.in.tum.de/teaching/ws1314/dbsys/exercises/>

Blatt Nr. 1

Gruppenaufgabe 1 (nicht Zuhause vorbereiten)

Sie designen ein Tool zur Univerwaltung, welches Sie aus gegebenen Anlass *TUMoffline* nennen wollen. Früh entschließen Sie sich zum Einsatz eines Datenbanksystems als Backend für Ihre Daten. Ihr Kollege ist skeptisch und würde die Datenverwaltung lieber selbst implementieren. Überzeugen Sie ihn von Ihrem Entschluss. Finden Sie stichhaltige Antworten auf die folgenden von Ihrem Kollegen in den Raum gestellten Äußerungen:

- Die Installation und Wartung eines Datenbanksystems ist aufwendig, die Erstellung eines eigenen Datenformats ist straight-forward und flexibler.
- Mehrbenutzersynchronisation wird in diesem Fall nicht benötigt.
- Es ist unsinnig, das jeder Entwickler zunächst eine eigene Anfragesprache (SQL) lernen muss, nur um Daten aus der Datenbank zu extrahieren.
- Redundanz ist hilfreich, wieso sollte man auf sie verzichten?

Hausaufgabe 1

Schätzen Sie die Größe der Datenbank von Facebook ab. Ermitteln Sie hierzu möglichst aktuelle statistische Werte über die Zahl der Nutzer, Fotos, Videos etc.pp auf Facebook und schätzen Sie fehlende Informationen geeignet ab. Nutzen Sie dann ihre Kenntnis über den Speicherverbrauch von Daten in einem Computersystem, um die "Datenbankgröße" von Facebook zu ermitteln.

- Geben Sie die abgeschätzte Größe sowie Ihren Rechenweg an.
- Welche Faktoren beeinflussen diese Größe. Denken Sie an Metadaten, zusätzlich generierte Informationen, Backups, Redundanz etc.pp.

Hausaufgabe 2

Erläutern Sie den Unterschied zwischen dem Relationalen Modell und dem Graphstrukturieren Modell.

- Nennen Sie ein typisches Einsatzgebiet für das jeweilige Modell.
- Im Datenbankbereich unterscheidet man zwischen Modellen, welche ein festes Schema voraussetzen und anderen, die kein Schema benötigen. Ein Schema ist hierbei eine Vorgabe, wie Daten repräsentiert werden, beispielsweise, das jede Vorlesung genau eine eindeutige Nummer hat, einen Namen von weniger als 200 Zeichen und eine Semesterwochenstundenzahl.
 - Was ist der Vorteil/Nachteil einer solchen Vorgabe für den Anwender?
 - Was ist der Vorteil/Nachteil einer solchen Vorgabe für den Entwickler des Datenbanksystems?

Gruppenaufgabe 2 (nicht Zuhause vorbereiten)

Finden Sie ein Beispiel für ein Problem (bzw. eine Inkonsistenz), die auftreten kann, wenn unkontrolliert parallel auf Daten zugegriffen wird. Ein traditionelles Beispiel hierfür ist eine gegenseitige Bank-Überweisung zwischen zwei Konten A und B. Wenn A einen Betrag x zu B überweist und B einen Betrag x' zu A, sollte immer gelten $Kontostand(A) + Kontostand(B)$ ist konstant, da sonst Geld verschwunden ist. Konstruieren Sie einen Ablauf zweier gegenseitiger Überweisungen, bei dem die Eigenschaft, dass die Kontostandssumme konstant sein soll nach dem Abschluss der zwei Überweisungen verletzt ist.