

*Modul Datenmanagement (DMG)***Übungsprojekt: NoSQL**

Sie erarbeiten sich selbstständig Know-how in einer NoSQL-Technologie, wenden dieses für einen Anwendungsfall an, und beschreiben ihre Erkenntnisse in einer schriftlichen Arbeit. Dazu arbeiten Sie in den gewohnten Übungsgruppen.

1. NoSQL-Technologie auswählen

- ☞ Wählen Sie eine NoSQL-Datenbanktechnologie, welche ihre Gruppe am meisten interessiert. Die Technologie soll im Ranking von DB-Engines.com (<http://db-engines.com/de/ranking>) in den top 50 vorkommen, und es soll sich um eine NoSQL-Technologie (siehe Skript) handeln.

2. Standardprojekt: Migration UniDB von SQL zu NoSQL

Sie kennen bereits die relationale Uni-Datenbank, mit der wir im Unterricht gearbeitet haben. Als Ausgangslage ist folgende Query gegeben, welche pro Professor die Anzahl Studenten und die Summe der SWS ausgibt.

```
select ProfessorName, AnzahlStudenten, SummeSWS
from
(
    select
        p.Name as ProfessorName,
        count(s.MatrNr) as AnzahlStudenten
    from Professoren p
    join Vorlesungen v on v.gelesenVon = p.PersNr
    join hören h on h.VorlNr = v.VorlNr
    join Studenten s on s.MatrNr = h.MatrNr
    group by p.Name
) A
join
(
    select
        p.Name as ProfessorName,
        sum(SWS) as summeSWS
    from Professoren p
    join Vorlesungen v on v.gelesenVon = p.PersNr
    group by p.Name
) B
using(ProfessorName);
```

Nun sollen Sie mit einer NoSQL-Technologie ihrer Wahl arbeiten. Sie recherchieren und wählen eine NoSQL-Datenbanktechnologie, mit der Sie arbeiten möchten. Sie setzen das ER-Modell der

Uni-DB in Ihrer NoSQL-Technologie als Schema bzw. Datenstruktur um. Dann migrieren Sie die Uni-Datenbank von SQL nach ihrer NoSQL-DB. Sie importieren also alle Daten der Uni-DB (Professoren, Studenten, etc.) in eine NoSQL-Technologie ihrer Wahl. Anschliessend definieren Sie eine Datenbankabfrage, welche das gleiche Resultat wie die oben definierte Query ergibt.

2.1. Uni-Datenmodell (ER-Diagramm) in der DB umsetzen

- ☞ Sie überlegen Sie sich, wie das konzeptionelle Uni-ER-Modell auf ihre Datenbanktechnologie abgebildet werden kann, und setzen dies in der Datenbank um. Sie überlegen sich, wie sie ihr konzeptionelles Uni-Datenmodell in der NoSQL-Datenbanktechnologie ihrer Wahl als Schema, Metadaten oder einer andern Form von Struktur abbilden können. Sie können dazu auch im Internet recherchieren. Beispiele:
 - Bei Key-Value-Stores können Sie den Schlüsselraum entsprechend definieren,
 - bei Document stores können sie die JSON-Struktur am ER-Modell ausrichten,
 - bei XML-Datenbanken das XML-Schema,
 - Bei Graphdatenbanken werden die Knoten- und Kantentypen und Attribute anhand des ER-Diagramms definiert

2.2. Uni-Daten migrieren

- ☞ Erstellen Sie die zugehörige Datenbank, inkl. Datendefinition (Schema bzw. Metadaten)
- ☞ Importieren Sie die Uni-Daten in ihre Datenbank

2.3. Abfrage umsetzen

- ☞ Schreiben Sie eine Query (oder ein Programm über die API), welche die in Abschnitt 2 beschriebene SQL-Abfrage in der NoSQL-Technologie Ihrer Wahl implementiert.

3. Alternative: autonomes Projekt

Wenn Sie möchten, können Sie anstelle des Standard-Projektes Ihre eigene Idee als App mit NoSQL autonom entwickeln. Als Idee, könnten Sie zum Beispiel ein Info-App mit Open Data programmieren:

<http://make.opendata.ch/wiki/data:ch>

- ☞ Vorgaben: Sie sollen (1) eine NoSQL-Datenbank verwenden, (2) ein Datenmodell entwerfen, (3) dieses Modell in der Datenbank umsetzen, (4) Daten aus einer externen Quelle dahin migrieren, und (5) diese Daten mit einer Anfrage abfragen. Die Anfrage soll dabei folgende Elemente aufzeigen:
 - Verbinden zweier Entitätsmengen (Join)
 - Verarbeiten von Wertemengen zu einem Wert (Aggregat)
 - Datensätze filtern (Selektion)
 - Attribute auswählen (Projektion)
 - Verschachtelung von Queries und Subqueries (Unterabfragen)

4. Alternative: Hackathon und / oder autonomes Projekt

Wenn Sie möchten, können Sie anstelle des Standard-Projektes Ihre eigene Idee als App mit NoSQL autonom entwickeln. Sie können das im Rahmen des Post Hackathons *Hacking Utopia* durchführen:

<https://www.hacking-utopia.ch>

☞ Im Idealfall nehmen Sie am Hackathon teil und entwickeln im Team eine NoSQL-basierte Lösung.

Als Idee, könnten Sie zum Beispiel ein Info-App mit Verwendung von Open Data programmieren:

<http://make.opendata.ch/wiki/data:ch>

☞ Vorgaben: Sie sollen eine NoSQL-Datenbank verwenden, Daten aus einer externen Quelle dahin migrieren, und diese Daten mit einer Anfrage abfragen. Die Anfrage soll dabei folgende Elemente aufzeigen:

- Verbinden zweier Entitätsmengen (Join)
- Verarbeiten von Wertemengen zu einem Wert (Aggregat)
- Datensätze filtern (Selektion)
- Attribute auswählen (Projektion)
- Verschachtelung von Queries und Subqueries (Unterabfragen)

☞ Sie schreiben anschliessend eine schriftliche Arbeit darüber, mit den gleichen Kriterien wie fürs Standard-Projekt.

☞ Diese Alternative gilt für alle, unabhängig davon, ob Sie am Hackathon teilnehmen.

5. Schriftliche Arbeit

Schreiben Sie einen Aufsatz zu Ihrem Projekt. Sie können anhand des fiktiven Fallbeispiels aufzeigen, wie die Software installiert wird und wie dann Datenschemas definiert, Daten eingefügt und abgefragt werden können. Zudem vergleichen Sie Ihre Technologie zu den relationalen Datenbanken. Schliesslich geben Sie auch eine Beurteilung aus Ihrer Sicht ab. Dieser Aufsatz soll, bezogen auf Ihr konkretes Projekt, die nachfolgenden Fragen beantworten.

☞ Vorgabe: Ihr Aufsatz darf nicht mehr als 2500 Wörter auf max. 11 Seiten (+ 1 Titelblatt) aufweisen.

☞ Dieser Aufsatz wird bewertet und zählt 1/3 der Schlussnote. Jede der 15 Fragen wird mit einer Note von 1 bis 6 bewertet. Kriterien in absteigender Reihenfolge der Relevanz: Inhalt, Stil, Rechtschreibung, Layout.

☞ Verwenden Sie folgende Struktur:

Titel, Namen, Modul, Semester

1. Einführung
 - 1.1. Was ist der Kontext, warum ist das Projekt relevant, und worum geht es?
2. Datenmanagement
 - 2.1. Um welche Datenbanktechnologie handelt es sich?
 - 2.2. Welche Anwendung (Use Case) unterstützt ihre Datenbank?
 - 2.3. Welche Daten werden migriert / eingefügt, und wie genau?
 - 2.4. Wie interagiert der Benutzer mit der Datenbank?
3. Datenmodellierung
 - 3.1. Welches Datenmodell (ER) liegt ihrem Projekt zugrunde?
 - 3.2. Wie wird ihr Datenmodell in Ihrer Datenbank in ein Schema übersetzt?
4. Datenbanksprachen
 - 4.1. Wie werden Daten anhand einer Query abgefragt?
5. Konsistenzsicherung
 - 5.1. Wie wird die Datensicherheit gewährleistet?
 - 5.2. Wie können Transaktionen parallel / konkurrierend verarbeitet werden?
6. Systemarchitektur
 - 6.1. Wie ist der Server aufgebaut und wie wurde er installiert?
 - 6.2. Wie kann die Effizienz von Datenanfragen optimiert werden?
7. Vergleich mit relationalen Datenbanken
 - 7.1. Vergleichen Sie ihre NoSQL-Technologie mit SQL-Datenbanken.
8. Schlussfolgerungen
 - 8.1. Was haben Sie erreicht, und welche Erkenntnisse haben sie dabei gewonnen?
 - 8.2. Wie beurteilt ihre Gruppe die gewählte Datenbanktechnologie, und was sind Vor- und Nachteile?

6. Präsentation

- ☞ Ausgewählte, besonders gute Projektergebnisse können als live-Demo beim Semesterabschluss präsentiert werden.

Bei Interesse bewerben Sie sich für eine Präsentation per Mail oder mündlich.

7. Abgabe

- ☞ Erstellen Sie ein PDF mit Ihrem Aufsatz
 - Projekt_Gruppe_<Kürzel>__<Thema>.pdf
- ☞ Laden Sie diese Dateien im ILIAS in den Ordner 01 Dateiaustausch / Projekt.
- ☞ Abgabetermin: Siehe Semesterplan Detail.