



Datenbanksysteme

Projekt: Übersicht

Freie Universität Berlin
Fachbereich Mathematik und Informatik, Institut für Informatik
Arbeitsgruppe Datenbanken und Informationssysteme
Prof. Dr. Agnès Voisard, Nicolas Lehmann
Sommersemester 2017

Abgabe bis spätestens **Freitag**,
den **14.07.2017, 14:00 Uhr**.

Iteration	1	2	3	Σ
Punkte	100	100	100	300

Hinweise:

Das Projekt soll in Dreiergruppen bearbeitet werden.

Jede Projekt-Iteration muss zu mindestens 70% erfüllt werden.

Die Ergebnisse jeder Projekt-Iteration müssen im Tutorium in Form einer 5 bis 10 minütigen Präsentation vorgestellt werden.

Projektziel

Es soll eine Web-Anwendung entstehen, die die Daten aus dem Datensatz "american-election" geeignet visualisiert und das Abfragen von Informationen ermöglicht. Typische Abfragen an die Web-Anwendung sollen sein:

- Welche Tweets waren am wichtigsten?
- Welche "Hashtags" wurden wann am meisten verwendet?
- Wann traten insgesamt am meisten "Hashtags" auf?
- Welche "Hashtags" treten besonders häufig gemeinsam auf?
- Wie hat sich die Häufigkeit der Verwendung eines speziellen "Hashtags" über die Zeit entwickelt?

Dokumentation

- erfolgt schriftlich in LaTeX
- enthält das Projektziel
- stellt das Team vor
- enthält alle wichtigen Entscheidungen
- beschreibt Schwierigkeiten, Probleme und deren Lösungen
- enthält Links zu Repositories mit entsprechendem Quellcode

1. Iteration: Modellierung

Für weitere Details siehe Dokument "Projekt: 1. Iteration: Modellierung".

1. Explorative Datenanalyse
2. ER-Modell
3. Relationales Modell
4. Datenbank erstellen

2. Iteration: Datenimport

Für weitere Details siehe Dokument "Projekt: 2. Iteration: Datenimport".

1. Datenbankschema erstellen
2. Daten bereinigen
3. Daten importieren
4. Webserver

3. Iteration: Data Mining und Visualisierung

Für weitere Details siehe Dokument "Projekt: 3. Iteration: Data Mining und Visualisierung".

1. Metrik für Ähnlichkeit zwischen zwei "Hashtags" entwickeln
2. Clusteranalyse durchführen
3. "Hashtag"-Netzwerk visualisieren
4. Häufigkeiten und Auftreten von "Hashtags" im Verlauf der Zeit visualisieren