

ProSA

KSWS/Projekt/NEidl SoSe 2021

Betreuer:

- Tanja Auge

Beschreibung:

Data Provenance beschreibt die Herkunft eines konkreten Anfrageergebnisses. Im Rahmen des ProSA-Projekts (**P**rovenance management using **S**chema mappings with **A**nnotations) entwickeln wir Techniken zur Invertierung von Anfragen. Ziel ist die Bestimmung der für das Anfrageergebnisses notwendigen Quelldaten. Wir konzentrieren uns hauptsächlich auf die **Why**- (siehe Cheney et al.) und **How**-Provenance (siehe Green et al.), welche wir durch (minimale) Zeugenbasen oder Provenance-Polynome angeben können.

Ziel dieses Projekts ist die Entwicklung einer GUI, welche die Eingabe bzw. Anbindung einer Datenbank, die Eingabe einer Anfrage sowie die Auswahl des Provenance-Typs (**why, where, how**) ermöglicht. Die Ausführung der Anfrage übernimmt das Java-Tool ChaTEAU. Dieses muss ebenfalls über eine Schnittstelle angebunden werden. Abschließend soll die GUI durch verschiedene Beispielanfragen getestet werden.

Arbeitsschritte:

- Einarbeiten in das Thema Data Provenance
- Aufstellen eines Zeitplans
- Konzeption einer GUI
- Anbindung einer Datenbank
- Implementierung einer ChaTEAU-Schnittstelle
- Entwicklung von Unittests

Voraussetzungen:

- Grundlegende Kenntnisse über Datenbanken
- Grundlegende Programmierkenntnisse (vorzugsweise in Java)
- Grundlegende Kenntnisse über UI-Patterns
- Grundlegende Kenntnisse im Bereich Data Provenance (Master)

Literatur (siehe Stud.IP):

- Provenance in Databases: Why, How, and Where (Cheney, Chiticariu, Tan)
- A survey on provenance (Herschel, Diestelkämper, Lahmer)
- Provenance semirings (Green, Karvounarakis, Tannen)
- Navigator — Ursprung von Daten (Sila)

Prüfung:

- Präsentation des aktuellen Zwischenstandes (1x pro Monat, Vortrag (10min pro Person))
- Ergebnispräsentation im September:
 - Vortrag (15min pro Person (Bachelor), 30min pro Person (Master)) sowie
 - schriftliche Ausarbeitung (5 Seiten pro Person (Bachelor), 10 Seiten pro Person (Master))

Meilensteine:

- 15.04.2020: Auftakttreffen
- 15.05.2020: Vorstellung der Problemstellung sowie des Zeitplans (Meilenstein I)
- 15.06.2020: Abschluss der Literaturrecherche und Vorstellung des GUI-Entwurfs (Meilenstein II)
- 15.07.2020: Anbindung der Datenbank und Schnittstelle zu ChaTEAU (Meilenstein III)
- 15.08.2020: Implementierung der GUI und Vorstellung erster Testanfragen (Meilenstein IV)
- 15.09.2020: (Meilenstein V)
 - Abgabe des Ergebnisberichts (5 Seiten pro Person (Bachelor), 10 Seiten pro Person (Master))
 - Vorstellung der Ergebnisse in Form eines Vortrags (15min pro Person (Bachelor), 30min pro Person (Master))

Bei Interesse bitte eine Mail an tanja.auge@uni-rostock.de!!