

Implementierung eines Profilers zur Performanceuntersuchung der Anfragealgebra CHQL

Ziel dieser Abschlussarbeit ist das Profilen/Benchmarken von Operatoren eines neuartigen Informationssystems. Sie untersuchen unterschiedliche Aspekte, wie z.B. die Laufzeit und den Speicherbrauch, einzelner Operatoren und Operatorkombinationen. Der Kern dieser Arbeit ist die Implementierung eines Profilers.

Im Rahmen dieser Abschlussarbeit arbeiten Sie mit der Anfragealgebra CHQL, eine hier am IPD entworfene Algebra zum Anfragen von temporalen Textkorpora. Mit CHQL können Sie z.B. die ähnlichsten Nomen zu dem Wort „Apple“ in der Zeit von 2000--2005 anfragen.

Für die Bearbeitung der Abschlussarbeit existieren bereits umfangreiche Vorarbeiten. Es existiert eine Referenzimplementierung der Anfragesprache, d.h. Implementierung der Operatoren und des Parsers. Für eine einfachere/schnellere Einarbeitung existiert darüber hinaus ein grafisches Frontend zur Erstellung von Anfragen und Visualisierung der Ergebnisse. Sowohl für die Programmiersprache Java als auch für das MapReduce-Framework Apache Flink existiert Software zum umfangreichen Profilen.

Current Expression:

1. word search (case insensitive regex) ^Apple\$
2. subsequence [from, to] 2000 - 2005
3. kNN 5
4. part of speech filter NOUN
5. subsequence [from, to] 2000 - 2005

Show first **Results**

Mit Bearbeitung dieser Abschlussarbeit haben Sie die Möglichkeit einen detaillierten Einblick in die Anfragebearbeitung moderner Informationssysteme zu bekommen. Sie beschäftigen sich mit zentralen Themen aus diesem Bereich: Performanceuntersuchungen von Anfragen. Darüber hinaus arbeiten Sie mit der gängigen Programmiersprache Java und dem aktuellen MapReduce-Framework Apache Flink.

Schreiben Sie mir bei Interesse eine E-Mail. Gerne zeige ich Ihnen weitere Details in einem persönlichen Gespräch.

Ansprechpartner

Jens Willkomm jens.willkomm@kit.edu +49 721 608-44065 Raum: 339

Am Fasanengarten 5 76131 Karlsruhe Gebäude: 50.34