

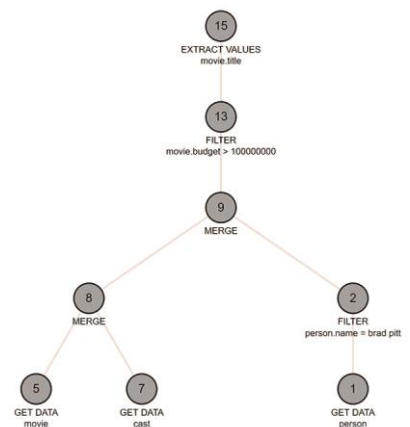
DB4Dummies: Interaktive Konstruktion von Datenbank-Abfragen

Datenbanken enthalten sehr viele strukturierte Daten, die für Personen aus verschiedensten Aufgabengebieten interessant sein können. Der Zugang zu diesen Daten ist für Personen ohne Kenntnisse einer Datenbanksprache allerdings stark begrenzt. Um diesen Zugang auch Laien bieten zu können, wäre eine Google-ähnliche Applikation wünschenswert, in der ein Benutzer eine natürlichsprachliche Frage, welche sich mit Hilfe von Informationen einer Datenbank beantworten lässt, eingibt und anschliessend die gewünschte Antwort erhält. Diese Applikation benötigt jedoch eine grosse Menge an Trainingsdaten, um mittels Machine Learning lernen und sich verbessern zu können. Um diese Daten generieren zu können, soll der bestehende Prototyp aus der gleichnamigen Projektarbeit erweitert werden, so dass anhand einer natürlichsprachlichen Frage nicht nur die korrekte Antwort aus der Datenbank geliefert wird, sondern auch die einzelnen Teilschritte des Benutzers in einer geeigneten Form aufgezeichnet werden. Diese Arbeit beschäftigt sich mit dem Thema 'Interaktive Konstruktion von Datenbankabfragen'. Das bestehende Konzept wurde teilweise übernommen, so dass sich ein Benutzer seine natürlichsprachliche Frage, welche einer komplexeren Datenbankabfrage entsprechen würde, anhand einer geeigneten Kombination von atomaren Operationen beantworten kann. Der Benutzer stellt sich so die komplette Abfrage zusammen, ohne auch nur eine Zeile SQL-Code schreiben zu müssen und erhält am Schluss das gewünschte Resultat. Der Prototyp wurde angepasst und erweitert, so dass die Einzelschritte des Benutzers in einer geeigneten Form aufgezeichnet werden und die Applikation künftig als Annotationstool genutzt werden kann, um Trainingsdaten für zukünftiges Machine Learning zu generieren. Zusätzlich wurden Funktionen implementiert, um die generierten Logfiles analysieren zu können. Das Ergebnis dieser Arbeit stellt eine Webapplikation dar, welche die geforderten Funktionen erfüllt. Weiterentwicklungen in Richtung Benutzeroberfläche könnten die Benutzerfreundlichkeit noch weiter verbessern.



Diplomierende
Nicolas Kaiser
Philippe Schläpfer

Dozierende
Mark Cieliebak
Kurt Stockinger



Grafische Darstellung der
Baumstruktur, welche in den
generierten Logfiles gespeichert wird.