

Der gläserne Politiker - Informationsgewinnung auf Daten des Schweizer Parlaments

Unsere Wahrnehmung der Politik ist dadurch beeinflusst, über was die Medien berichten. Welches Thema wie viele Zeitungsspalten oder wie viel Sendezeit erhält, hängt stark von den vermuteten Leserzahlen oder Einschaltquoten ab. Ebenfalls sind die Inhalte geprägt von den politischen Ansichten der News-Organisationen und Journalisten. Was kann man tun, wenn man sich objektiv und ungefiltert über politische Vorgänge informieren will? Das Schweizer Parlament stellt der Allgemeinheit eine sehr umfangreiche Sammlung an Daten zu Geschäften, Anträgen, Debatten, Abstimmungen etc. zur Verfügung. Diese Daten sind zwar alle sehr leicht zugänglich, detaillierte Analysen lassen sie jedoch nur mit sehr grossem Aufwand zu. Diese Bachelorarbeit versucht, dem politisch interessierten Benutzer diesen Aufwand etwas zu verringern.

Bei den erwähnten Daten handelt es sich zum einen um hoch strukturierte Daten wie Abstimmungsergebnisse, Informationen zu Personen und Parteien etc. Zum andern jedoch sind viele Informationen auch in unstrukturierter Form verfügbar, wie eingereichte Texte oder mitgeschriebene Reden. Das Ziel ist es, diese Daten auf sinnvolle Art zu kombinieren.

Es wurde ein Data Warehouse entwickelt, welches einen geeigneten Aufbau hat, um die strukturierten Daten möglichst vollständig zwischenspeichern und auf einfache Art zugänglich zu machen. Anschliessend wurden die Texte mit Methoden des Information Retrieval, im speziellen des Named Entity Recognition (NER), untersucht und die Erkenntnisse daraus ebenfalls in die Struktur des Data Warehouses abgelegt. Die durch diese Methode erkannten Entitäten enthalten Namen von Organisationen, Personen und Ortschaften und geben direkt Auskunft darüber, was in den Parlamentstexten tatsächlich geschrieben oder in den Ratsdebatten gesagt wurde.

Das Resultat ist ein Prototyp, welcher komplexe Abfragen auf einfache Art zulässt. Es sind damit unzählige Auswertungen möglich, die verschiedene Themen sowohl aus der Struktur der Daten als auch aus den unstrukturierten Texten kombinieren. Neben unzähligen Informationen zu jedem Politiker, wie beispielsweise Alter, Geschlecht, Partei oder Interessenbindungen, können auch alle Abstimmungsergebnisse abgefragt werden. Ausserdem können mittels der gefundenen NER-Entitäten Erkenntnisse über den Inhalt der Texte und Reden in die Auswertungen integriert werden. Dem politisch interessierten Benutzer sind dabei kaum Grenzen gesetzt.



Diplomierende
Nico Thomas Rüede
Serge Schoop

Dozierende
Martin Braschler
Kurt Stockinger



Bundeshaus, Sitz des
Schweizerischen Parlaments