

DebateExplorer: Interactive journalistic investigation of large text collections

Initiative: Wissenschaft und Datenjournalismus

Bewilligung: 20.10.2015

Laufzeit: 9 Monate

Projekt-Website: <https://www.debateexplorer.de>

The main goal of this project is to develop an interactive research environment for text collections that (a) puts state-of-the-art text analysis models from Computational Linguistics in the hands of data journalists, allowing them to quickly tailor search facilities and filters to their research goal, i.e., finding and categorizing textual passages in the collection that instantiate a relevant position towards an issue under exploration. The environment furthermore (b) relates the categorized positions, or claims, to the uttering actors, capturing dates of utterance, the relation to relevant mentioned entities, and (c) provides exploratory facilities and visualization tools for performing time-series analysis and network analysis on aggregated text-analytical results, including differential analysis against trends observed in previous legislation processes. By keeping all backward links from aggregated results to the individual underlying text sources, the environment readily supports (d) a critical assessment of the analysis and (e) a transparent presentation of the data basis of a news story. The project involves a researcher from the University of Stuttgart and a journalist from the "Stuttgarter Zeitung".

Projektbeteiligte

Prof. Dr. Jonas Kuhn

Universität Stuttgart

Fak. 5: Informatik, Elektro-, Informationstechnik

Institut für Maschinelle Sprachverarbeitung

Stuttgart

Eva Wolfangel

Stuttgart

Open Access-Publikationen

Towards a text analysis system for political debates

