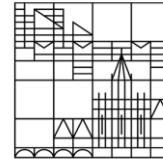


Topics Seminar / Lecture / Thesis



Universität
Konstanz



BERGISCHE
UNIVERSITÄT
WUPPERTAL

Media Bias



G7 summit, 2014



WILLKOMMEN IN DEUTSCHLAND!

Tausende Flüchtlinge erreichen München

“Welcome to Germany! Thousands of refugees arrive in Munich“

VS.

Islam-Invasion nach Deutschland

5. September 2015 348

“Islam-Invasion to Germany“

Media Bias

Media bias is the bias of news producers in the selection of events and stories, and their style/word choice. This overall project aims on uncovering bias on a large scale. We are a bigger team, always up for a beer, and we believe in the societal relevance of the topic.

We're always looking for someone to join our team for a master or bachelor thesis. The project is taking place internationally, between the NII in Tokyo, the Technion in Haifa and the Universities of Passau, Konstanz & Wuppertal.

We offer close guidance for you and your work – so far, every thesis project that was done within the project was published in an international conference.

Get in touch by contacting us by Timo.Spinde@uni-Konstanz.de

MB1: Automating the Creation of Bias Lexica

Background

Do you say: "Refugees will be distributed among all of Germany"? Or do you say: "Islam invades Germany?" Media bias is one of the most trending scientific topics today, it affects us, everywhere, anytime. We developed an approach that improves performance of identifying medium-specific bias language. In the next step, this approach will become automated. To achieve this, one task within the project is to find topic-related words that are likely to be semantically similar to potential bias words.

Goal

This part of the project aims on identifying if potential bias words within a seed word selection process are closer to words that are linguistically likely to induce bias.

Tasks

- Get familiar with media bias & semantic models.
- Scrape an English data set.
- Build a robust automated or semi-automated algorithm to compare words via word embeddings.
- Test and evaluate the approach.



Timo Spinde

Timo.Spinde@uni-konstanz.de



MB2: Efficient storage and retrieval of textual data

Background

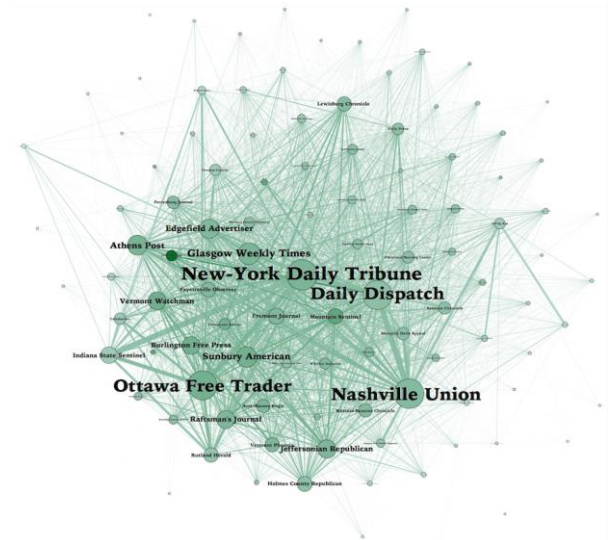
Recent efforts in data exploration use concepts derived from graph theory, to extract meaning from textual data. Examples are the identification of similarity subgraphs from lexica (Veremeyev et al. 2019), or the clustering of a large document archive (13M+ health records) using graph partitioning (Altuncu et al. 2018). While we are working extensively on the identification of unbalanced coverage, a fitting representation will be as important. Within this project, we will gather a large data set, transform it into a fitting graph database structure, and create a suitable user interface.

Goal

Enable users to upload and use large collections of news articles within a graph/network structure.

Tasks

- Review and assess relevant literature with regards to NLP and graph representation
- Assess available Graph Databases and pick a suitable candidate
- Implement a workflow to transfer semi-structured text to graph representation within the DB
- Benchmark OR
- Develop an exemplary application



Timo Spinde
Timo.Spinde@uni-
konstanz.de

MB3: Visualizing Media Bias in an online setting

Background

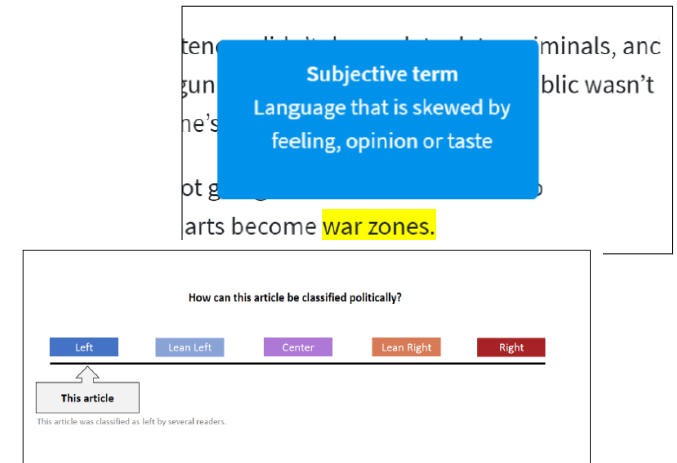
Do you say: "Refugees will be distributed among all of Germany"? Or do you say: "Islam invades Germany?" Media bias is one of the most trending scientific topics today, it affects us, everywhere, anytime. We are developing algorithms to identify such bias, and we have researched how one could present our results to users. In this project, we will bridge the gap between analysis and visualization: We will create an online tool or browser add-on, uncovering media bias to readers, while getting their feedback and reactions to improve the machine learning / deep learning process.

Goal

Enable users to understand media bias in an online environment

Tasks

- Get familiar with web development
- Learn about media bias and perception of media bias
- Evaluate and test the environment you create



MB4: Lexical News Medium Comparison by Word Embeddings

Background

In this project, lexical comparison between left- and right-wing news mediums will be conducted with the focus on bias inducing lexica and with the use of word embeddings. In our previous work, no direct relation is found between bias inducing words and the words whose vectors are distant from each other in the word embedding spaces trained separately on two left and right article data sets. Among the most similar words (in the word embedding spaces) to the words related to controversial opinions between left- and right-wing communities, no salient differences were found, but there is lots to do. You will take a detailed look into transforming vector spaces, creating a powerful data set, and experimenting with the comparison of word embeddings.

Goal

Identify the complex context of media bias wording

Tasks

- Scrape a dataset of many news outlets
- Implement methods to compare word context



Timo Spinde

Timo.Spinde@uni-konstanz.de

