

**Philosophische** Fakultät III

Sprach- , Literatur- und Kulturwissenschaften

Institut für Information und Medien, Sprache und Kultur (I:IMSK)  
Lehrstuhl für Medieninformatik

Projektseminar Mediengestaltung: Computational Social Science: Text Mining und Social Network Analysis als Analysemethoden in den Sozialwissenschaften

Modul: MEI-M05.3

Wintersemester 2017/2018

Explore the Twitter – Use of Politicians

Marian Wolff

Matrikelnummer: 1752031

Angewandte Bewegungswissenschaften und Medieninformatik

7. bzw. 6. Bachelor of Arts

E-Mail: [marian.wolff@stud.uni-regensburg.de](mailto:marian.wolff@stud.uni-regensburg.de)

Simon Dobiosz

Matrikelnummer: 1740941

Informationswissenschaft und Medieninformatik

6. Bachelor of Arts

E-Mail: [simon.dobiosz@stud.uni-regensburg.de](mailto:simon.dobiosz@stud.uni-regensburg.de)

Abgegeben am 31.12.2017

Inhalt

1 Kurzbeschreibung 3

2 Datenbeschaffung 4

3 Netzwerk der Politiker anhand der Retweets 5

4 Tweetanteil der Parteien 5

5 Top-Hashtags 6

6 Hashtagsuche und nächste Partner 7

Abbildungen

Abbildung 1: Netzwerk der Parteien 8

Abbildung 2: Tweetanteil pro Woche 9

Abbildung 3: Tweetanteil gesamt 10

Abbildung 4: Top 5 Hashtags nach Parteien 11

Abbildung 5: Hashtag-Suche 12

# Kurzbeschreibung

Im Rahmen des Projektseminars Mediengestaltung: „Computational Social Science: Text Mining und Social Network Analysis als Analysemethoden in den Sozialwissenschaften“ wurde das Twitterverhalten von über 1000 Politikern anlässlich der Bundestagswahl 2017 untersucht. Es wurden die Verhältnisse untereinander anhand der Retweets und die meist verwendeten Hashtags analysiert und eine Hashtag-Suchfunktion bereitgestellt. Diese Daten wurden über RStudio bearbeitet und anschließend mit JavaScript und der D3-library visualisiert.

# Datenbeschaffung

Der Datensatz der Tweets wurde von der Universität Regensburg bereitgestellt. Insgesamt wurde 389.874 von 1068 deutschen Politiker der sieben großen Parteien (AfD, CDU, CSU, FDP, Grüne, Linke, SPD) untersucht. Die Tweets wurden zur Bundestagswahl 2017 am 24.09.2017 gesammelt. Der Zeitraum beginnt 37 Wochen vor und endet drei Wochen nach der Wahl.

Der Datensatz (twitter\_data.csv) wird mit RStudio bearbeitet und gibt Aufschluss über Name, Partei, Kalenderwoche des Tweets und weitere Variablen bezüglich des Tweets wie beispielsweise isRetweet oder hashtagCount. Die Daten wurden anschließend im JSON-Format exportiert und in die Umgebung des HTML-Projekts gelegt, so dass diese für die Visualisierungen zur Verfügung stehen.

# Netzwerk der Politiker anhand der Retweets

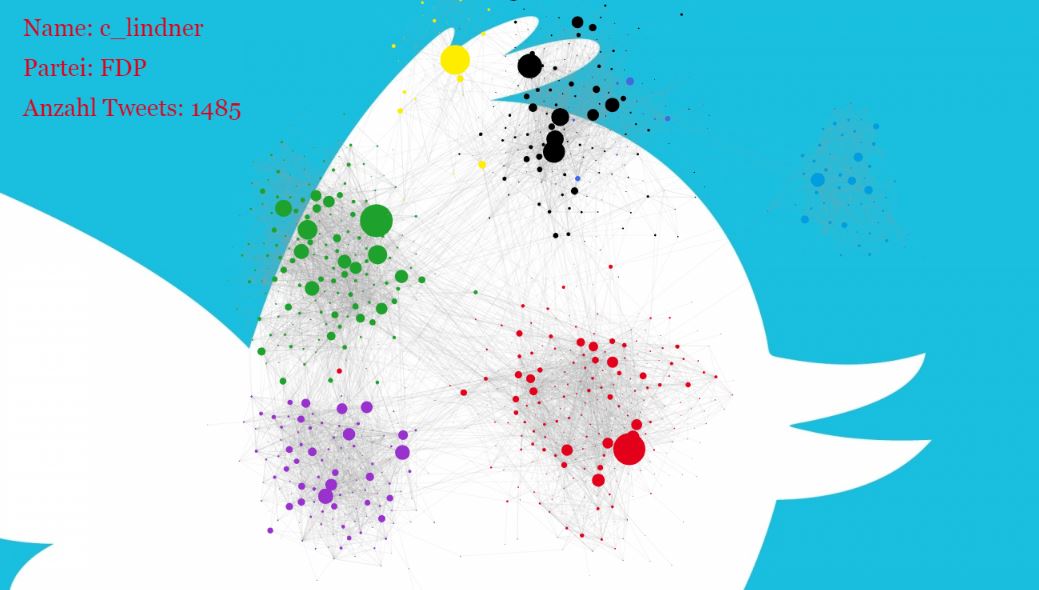
Das Netzwerk zeigt das Twitterverhalten der einzelnen Politiker. Über die Anzahl der Retweets wurde mit Hilfe der igraph-library ein PageRank erstellt, der über die Größe der Nodes deutlich wird. So haben beispielsweise die Spitzenkandidaten wie Christian Lindner (c\_lindner) oder Martin Schulz (MartinSchulz) eine (visuell) größere Bedeutung. Über die Links erkennt man, welche „Retweet-Verbindung“ (Welcher Politiker hat wen geretweetet?) zwischen den Politikern herrscht. Hier ist eine deutliche parteiinterne Nähe erkennbar (vgl. Abbildung 1).

Abbildung 1: Netzwerk der Parteien

Durch Anklicken eines Nodes kann dieser bewegt werden und es werden in der oberen linken Ecke des Canvas Informationen zu entsprechenden Usern angezeigt.

# Tweetanteil der Parteien

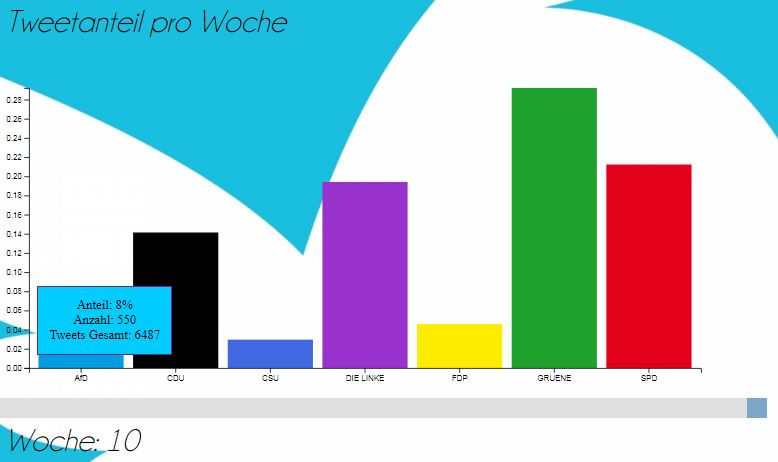
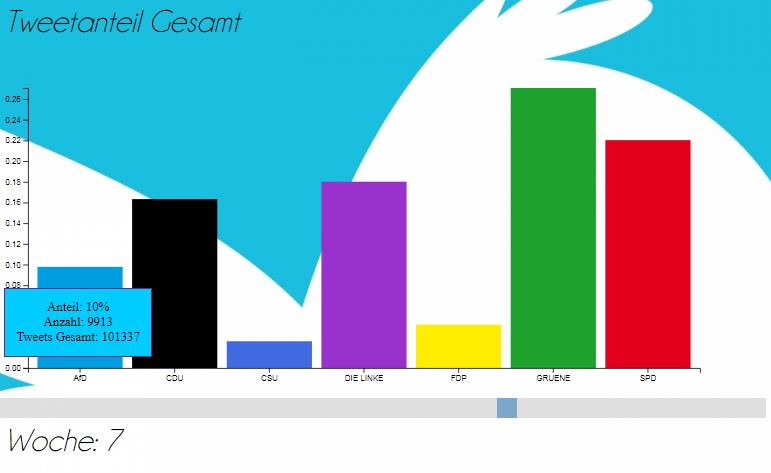
Nach den einzelnen werden nun die Tweets der Parteien miteinander verglichen. Der Zeitraum geht von 6 Wochen vor bis 3 Wochen nach der Wahl. Hier der Anteil der Tweets der Politiker einer Partei mit der Gesamtheit unterschieden. Man kann hierbei zwischen den einzelnen Wochen (siehe Abbildung 2) unterscheiden oder den kompletten Zeitraum (Abbildung 3) vergleichen.

Abbildung 3: Tweetanteil gesamt

Abbildung 2: Tweetanteil pro Woche

# Top-Hashtags

Die Top-Hashtags geben die fünf am häufigsten verwendeten Hashtags im Verlauf der letzten sechs Wochen vor bis drei Wochen nach der Wahl an. Es kann zwischen den einzelnen Wochen sowie den Parteien unterschieden werden. Hierbei wurde der Datensatz entsprechend mit R aufbereitet und über JavaScript mit simplen String und Variablenvergleichen die gewollten Hashtags angezeigt (Abbildung 4).

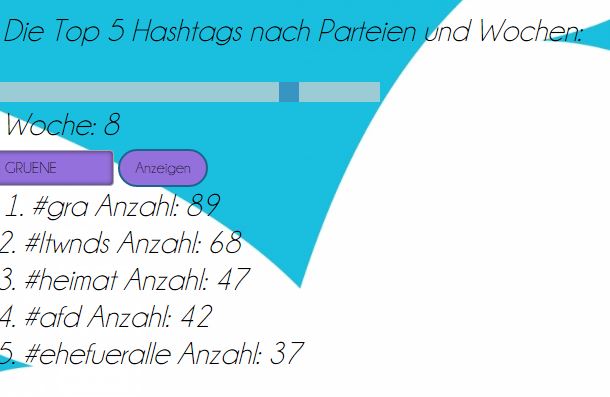


Abbildung 4: Top 5 Hashtags nach Parteien

# Hashtag-Suche und nächste Partner

Bei der Hashtag-Suche wurden zuerst alle Tweets die mehr als einen Hashtag haben, herausgefiltert. Anschließend wurden alle Hashtags, die immer gleichen Tweet vorkommen ausgewählt und gezählt. Somit wird die relative Nähe der einzelnen Hashtags erzeugt und bei der Suche erkennt man, welche Hashtags häufig mit anderen Hashtags im selben Tweet vorkommen (Abbildung 5).

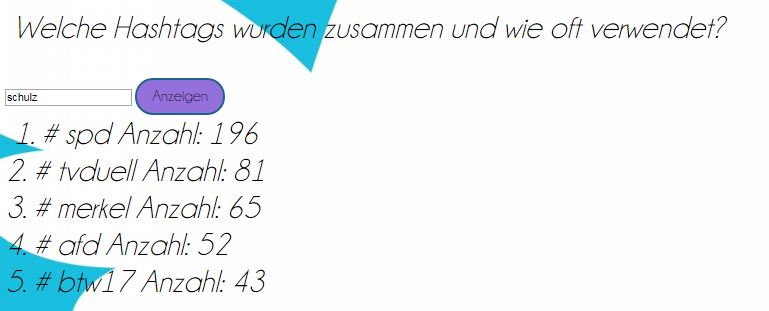


Abbildung 5: Hashtag-Suche