Research Center
Finance & Information Management

Fraunhofer
FIT

Project Group
Business & Information
Systems Engineering

User Roles in Online Political Discussion: A Typology based on Twitter Data from the German Federal Election 2017

by

Henner Gimpel, Florian Haamann, Manfred Schoch, Marcel Wittich

presented at: 26th European Conference on Information Systems (ECIS),
Portsmouth, UK, June 2018

University of Augsburg, D-86135 Augsburg
Visitors: Universitätsstr. 12, 86159 Augsburg
Phone: +49 821 598-4801 (Fax: -4899)

University of Bayreuth, D-95440 Bayreuth
Visitors: Wittelsbacherring 10, 95444 Bayreuth
Phone: +49 921 55-4710 (Fax: -844710)

WI-743

USER ROLES IN ONLINE POLITICAL DISCUSSIONS: A TYPOLOGY BASED ON TWITTER DATA FROM THE GERMAN FEDERAL ELECTION 2017

Research paper

Gimpel, Henner, FIM Research Center, University of Augsburg, Augsburg, Germany, henner.gimpel@fim-rc.de
Haamann, Florian, FIM Research Center, University of Augsburg, Augsburg, Germany, florian.haamann@fim-rc.de
Schoch, Manfred, FIM Research Center, University of Augsburg, Augsburg, Germany, manfred.schoch@fim-rc.de
Wittich, Marcel, FIM Research Center, University of Augsburg, Augsburg, Germany, marcel.wittich@fim-rc.de

Abstract

Twitter is well recognized as a microblogging site, an online social network (OSN), and increasingly as a digital news platform. With the changing media usage behavior over the past decade, political actors have now recognized the need to enrich their election campaign efforts by including social media strategies. However, previous research has shown that users behave heterogeneously in online political discussions. To better understand how users behave and interact in such debates, we conduct an exploratory study to identify emergent user roles from Twitter data. We develop a dynamic selection query to collect a representative data set on the German federal election of 2017. We define features of structure, function, and time for Twitter discussions and conduct a cluster analysis to derive eleven emergent roles from the 30,553 most active users. We then refine those roles by further data-driven analyses to enhance and deepen their understanding. Our results indicate dominance of the online discussion by the populist party Alternative für Deutschland. We also find that media outlets and political parties show somewhat similar behavior, and that the offline popularity of prestigious actors is extended into the online world.

Keywords: social media, user behavior, emergent roles, digital trace data, political discourse.

1 Introduction

Since its release in July 2006, the microblogging platform Twitter has become one of the most popular online social networks (OSN) with more than 330 million monthly active users (Twitter, 2017b). Its increasing popularity manifests itself in more than 500 million tweets per day (Twitter, 2013). Twitter is well recognized as a microblog and OSN, but increasingly also as a digital news platform where user-generated news is provided by millions of users. With Millennials and its successor Generation Z coming of age, the behavior of media consumers has changed drastically and social media, such as Twitter, has become an important source of news. According to a study conducted by the Reuters Institute for Journalism almost a quarter of a representative sample of the global population with access to the internet use social media as their primary news source (Newman et al., 2017). In the United States of America, where the adoption of social media is at a high level, about 50% use it as a news source. In contrast, only 29% of the respondents in Germany use social media for news, significantly lower values than in any other surveyed country. Further usage data shows that the demographic group of 16 to 29-year olds (the Millennials and Generation Z) is the largest group that uses Twitter on a regular basis. This group now owns the right to vote, and represented almost 5% of the total voters in recent German elections (Bundeswahlleiter, 2017b). This changing consumer behavior affects the individuals' daily lives and the way our society functions. As Kane et al. (2014, p. 298) state, the „most significant impact and implications of social media on business and society are still to come". Consequently, this development becomes evident in the way political campaigns have responded to this new media consumption behavior.

As the recent presidential election in the United States of America has shown, political parties and players enrich their campaigns by including the use of social media platforms to their campaign efforts to reach millions of their voters. Thus, social media has become an important influential factor and tool for politicians. The Twitter accounts "@POTUS" (President of the United States of America) and "@realDonaldTrump" (private account of Donald Trump) for example are amongst the users with the most followers of all heads of governments on Twitter, with 20.2 million and 37.8 million followers respectively. In contrast, German chancellor Angela Merkel does not have her own Twitter account, but rather tweets through her government spokesman Steffen Seibert who has barely a million followers. Compared to that, with a single tweet the US president can reach a number of Twitter users that is equivalent to more than 15% of the United States of America's 231 million voters – not to mention the media coverage that his tweets receive. During his election campaign, Donald Trump has revolutionized the way political debates are carried out in OSNs. Trump himself claims his "use of social media is not Presidential – it's MODERN DAY PRESEDENTIAL" (Twitter, 2017a). It becomes evident that despite the importance of OSNs for campaigns, not every politician is using the network to a similar extent.

The change in media and news consumption affects the way voters carry out political debate and has intensified "cultural polarization of democratic societies" (the divergence of opinions to political extremes) (World Economic Forum, 2017). If political actors want to reach young voters successfully, they need to adapt their campaigns to these changed circumstances. In addition to politicians, media outlets also need to realign to reach substantial audiences in this changed environment, as television and print have remarkably declined as news sources (Newman et al., 2017) and voters nowadays have easy access to multiple digital sources of information. Therefore, there is an increasing need to understand how users, particularly political actors, and media outlets, adapt their social media appearance, and how they behave and interact in political debates in OSNs.

Our research aims to understand in which roles users interact with each other in the online political discussion. Hence, we examine what groups of similar behavior can be found, how the behavior of political parties differs from one another, and how their behavior compares to that of media outlets.

Based on real world usage data from the OSN Twitter on the subject of the German federal election 2017, we conduct an exploratory study. We gathered a data set that includes status updates (e.g., tweets) collected through a dynamic set of hashtags for a duration of eight weeks leading up to the election. By applying a hierarchical clustering based on eight quantitative measures of the users' behavioral patterns, we derive a user typology of eleven emergent user roles. Subsequently, those results are refined and enhanced with qualitative insights from content analysis. Although Germany is not considered a country with high Twitter adoption rate (Newman et al., 2017), the data set clearly highlights broad discussion among thousands of linked users. The recent German federal election is of particular interest for practitioners and researchers because of Germany's pluralistic democracy. The rise of populist parties affected the political landscape in Germany like in several other countries in recent years and the election outcome will have an important impact on the future of one of the world's largest economies. Furthermore, all authors of this paper possess expertise regarding the German political landscape and media outlets. This was particularly helpful for the successful data collection, data analysis, and the interpretation of the results.

This paper contributes to the ongoing research on political discussion in OSNs as follows: first, we characterize emergent user roles through an exploratory cluster analysis and distinguish their unique characteristics. Second, we show that political parties, news agencies and media outlets fall into a limited number of user roles. Third, we further describe and explain the user roles' behavior based on qualitative observations of the Twitter discussion. Fourth, we find that populist parties are discussed much more frequently and behave significantly differently from other parties.

The remainder of this paper is structured as follows: Section 2 gives an overview on political discussions in OSNs and the political landscape in Germany and reviews the existing literature regarding user roles.

Section 3 explains our approach and data set. Section 4 contains the results of the study. We proceed to discuss the contributions derived from our results in Section 5. Lastly, Section 6 assesses our study critically regarding its limitations and concludes the paper.

2 Problem Context

2.1 Research on Political Discussions in OSNs

Sunstein (2001) highlights the relevance of active political discourse and its importance for a well-functioning democracy and emphasizes the significance of upcoming changes and threats through OSNs. The online political discourse has been a highly discussed topic for scholars in many regards. For example, Kelly et al. (2005; 2006) analyzed political newsgroups, found that most active users in political discussions can be assigned to ideological positions and further, that discussions mainly happen between polarized groups. Gonzalez-Bailon et al. (2010) observed that online discussions on politics are more long-lasting and include more participants than discussions on other topics. Tumasjan et al. (2010) analyzed social networks with respect to elections, particularly the German federal election of 2009. While their work focuses on the reflection of political sentiment and predictive power of Twitter, they also find a lively political debate dominated by a small number of users. Shi et al. (2014) introduced the term "social broadcasting network" for OSNs, which refers to the blurred boundaries between OSNs and news media. Specifically, the platform Twitter contributes to this development with its functionalities for easy and quick information distribution. Yet, the debate on the impact of such OSNs for politics and democracy is ongoing (e.g. Sunstein, 2001; Adamic and Glance, 2005; Yardi and Boyd, 2010; Conover et al., 2011; Barberá, 2015; Bakshy et al., 2015; Flaxman et al., 2016; Schmidt et al., 2017). Several of those studies analyze user interactions in OSNs regarding their impact on the political discussion. However, they disagree on whether such discussions increase political deliberation (Kelly et al., 2005; 2006; Kushin and Kitchener, 2009; Barberá, 2015) or political polarization (Conover et al., 2011; Grömping, 2014; Schmidt et al., 2017). To enhance the understanding of OSNs in the context of politics, its debates and discourses, it is necessary to start the analysis at the origin of these effects: the behavior of the individual users within OSNs. The aforementioned studies have analyzed user behavior on an aggregated level, however, individual user behavior often cannot be clearly distinguished, as individual differences get lost through aggregation (Kelly et al., 2005). Therefore, Kelly et al. (2005) called for more fine-grained analysis of user behavior in online political discussions.

2.2 Related Work on User Roles

Modern OSNs and their APIs enable researchers to overcome previously existing data limitations while exploring heterogeneous behavior (Himelboim et al., 2009) and have led to exploratory studies on user roles and typologies in various disciplines. Gleave et al. (2009, p. 2) based on Lerner (2005) reason that classifying "types of social relations and behaviors into a smaller set of roles reduces the analytic complexity of social systems" and eventually enable comparative studies of populations across time and setting (Gleave et al., 2009). Roles are valuable to researchers because they combine expectations, responsibilities and skills and facilitate a-priori knowledge about people (Golder and Donath, 2004). Role typologies originate from the social sciences. In his theory of social structure, Nadel (1957) introduced relational and non-relational roles. Later, Golder and Donath (2004), Gleave et al. (2009), and Forestier et al. (2012) all outlined definitions for roles in the context of OSNs. Roles that emerge through interaction are according to Nadel (1957) non-relational roles, with interactions being structural signatures of actors (Welser et al., 2007). Therefore, roles can be identified based on patterns and similarities in these structural signatures (Welser et al., 2007). These types of user roles are also known as emergent user roles in (online) social networks (Gleave et al., 2009; Mirbabaie et al., 2014; Frank et al., 2017). In addition to roles based on user behavior, relational roles are based on existing social relations between actors (Barley, 1990; Mirbabaie et al., 2014). For example, the role of a parent is only associated with an actor if the actor has a child. Both types of roles are relevant and some argue that they cannot be completely separated (Barley, 1990). However, while non-relational user roles may emerge from interactions, and therefore can be observed by replicating interactions through digital trace data (Gleave et al., 2009), relational roles cannot. Hence, researchers identifying emerging user roles based on individual behavior (e.g. Füller et al., 2014; Arazy et al., 2016; Frank et al., 2017) may include relational elements of roles to verify whether there is an association between the identified emergent roles and the existing relational roles.

Early forms of modern OSNs, such as Usenet, were already of particular interest to researchers. Golder and Donath (2004) created a theory-based taxonomy of roles in electronic communities. They validated their taxonomy by applying it to Usenet and identified the roles: *celebrity*, *newbie*, *lurker*, *flamer*, *troll* and *ranter*. Usenet's available digital trace data allowed to quantitatively assess interaction and use patterns or network visualizations to identify roles. Nolker and Zhou (2005), Kelly et al. (2006), Fisher et al. (2006), Welser et al. (2007), and Himelboim et al., (2009) all conducted studies identifying user roles and typologies in Usenet discussion groups based on trace data. Nolker and Zhou (2005) used pattern-based metrics of social network analysis to validate *leaders*, *motivators* and *chatters*. Welser et al. (2007) used network visualizations and regression analysis to show that these methods can be used to identify roles and to encourage further research. Some researchers focused on user roles specifically in the context of online political discussions. For example, Kelly et al. (2006) found heterogeneous user behavior by analyzing visualizations of ego-networks and identified *fighters*, *friendlies* and *fringes*. Himelboim et al. (2009) identified and described the role of *discussion catalysts* using basic quantitative measures and content analysis. Others have explored user roles on Twitter (Tinati et al., 2012; Beller et al., 2014; Mirbabaie et al., 2014; Pervin et al., 2014; Mirbabaie and Zapatka, 2017). Tinati et al. (2012) examined insights of marketing and described the roles *idea starter*, *amplifier*, *curator*, *commentator* and *viewer* by utilizing metrics of social network analysis. They highlighted the need for refinement and suggested the analysis of larger data sets. Beller et al. (2014) explored status updates for self-identification patterns and identified users social roles (e.g., *man*, *woman*, *celebrity,* or *teacher*). They demonstrated the potential of content analysis on Twitter and emphasized the need for further analysis in this area. Mirbabaie et al. (2014) used network visualizations of users' behavior to identify roles on Twitter. By manual assessment, they detected five communication roles: *emergency service agencies*, *media organizations*, *political groups and unions*, *individuals* and *commercial organizations*. These roles are not only a result of analyzing interactions, but also through the actors' given occupation and status in society. Similarly, Stieglitz et al. (2017) reveal such social roles among Twitter power users during crises communication. Mirbabaie and Zapatka (2017) applied the role typology of Pervin et al. (2014) to map Twitter users participating in crises communication to *information starters, amplifiers* and *transmitters* and describe their characteristics. Lately, Füller et al. (2014) and Frank et al. (2017) followed a data-driven exploratory role identification approach utilizing cluster analysis to identify emergent user roles in online communities. Füller et al. (2014) analyzed user interactions in online innovation communities and found *socializers, (passive) idea generators, masters, efficient contributors,* and *passive commentators*. Frank et al. (2017) identified user roles in enterprise social software.

To conclude, many studies exist that explore user behavior in online discussions with a multitude of methods. Some focus on OSNs, others on political discussions. However, an analysis of user interactions on Twitter in the explicit context of online political discussions has yet to be conducted. Previously, most researchers focused their analysis on confirming existing behavioral patterns or manually assessing users to identify roles. The exploratory approach of Füller et al. (2014) and Frank et al. (2017) only considers user behavior and does not require existing hypotheses on user roles nor extensive manual assessment. Applying their approach to the political debate on Twitter, we aim to identify emergent user roles in the online political discussion for the German federal election 2017.

2.3 Description of the Political Landscape in Germany

Germany's legislature is organized as a parliamentary democratic system that is characterized by many different parties. Governments are mostly formed through coalitions of several parties, which have been led by either one of Germany's major parties, the union of *Christlich Demokratische Union Deutschlands (CDU)* and *Christlich-Soziale Union in Bayern e.V. (CSU),* or the *Sozialdemokratische Partei Deutschlands (SPD).* Besides these major parties, smaller parties such as *Freie Demokratische Partei (FDP)* or *Bündnis 90/Die Grünen (Grüne)* have participated in government coalitions as well. The party

Die Linke was often represented in the parliament (i.e. Bundestag), but never part of the government; and the newly found (2013) party *Alternative für Deutschland (AfD)* gained remarkable attention during past regional elections in the federal states. In the most recent German federal election on 24 September 2017, a new high of 42 parties participated (Bundeswahlleiter, 2017a), of which most received only a small fraction of the votes. If a party reaches a threshold of five percent of the total votes, it gains proportional representation in the parliament. The election yielded the following results of the parties to be represented in the parliament: CDU/CSU 33.0%, SPD 20.5%, AfD 12.6%, FDP 10.7%, Die Linke 9.2%, Grüne 8.9%, and all others 5% in total (Bundeswahlleiter, 2017c). The AfD has entered the German parliament for the first time in its history. Prior to the election, all parties who were voted into the parliament have already denied to form a coalition with the AfD because of its populist right-wing orientation and several public extremist statements of AfD politicians (Handelsblatt, 2017; Münchner Merkur, 2017; Tagesschau, 2017c). The political spectrum of the seven parties in the newly formed Bundestag can roughly be aligned from left-wing to right-wing as follows: *Die Linke*, *Grüne*, *SPD*, *FDP*, *CDU/CSU*, and *AfD* (infratest dimap, 2015).

3 Methodology

3.1 Data Set

In this study, we rely on publicly accessible data collected through Twitter's APIs. The global stream of public status updates is made available to developers through the Streaming API to capture posted content in real-time. In this paper, we differentiate between four mutually exclusive types of status updates, which we explain later in greater detail: *tweets*, *retweets*, *replies*, and *quote tweets*. We collected all status updates from August 2, 2017 until September 24, 2017 at 6 PM (closure of polling stations) that included at least one of a set of hashtags related to the German federal election. To capture a comprehensive sample of Twitter data, we set up a data collection system that automatically identified frequently used hashtags related to the election. We initially started collecting data by using the following unbiased base set of hashtags as selection criteria to query Twitter's API: "#btw17", "#btw2017", "#Bundestagswahl", "#Bundestagswahl17", "#Bundestagswahl2017", "#Wahl17", "#Wahl2017", "#Wahlen17", "#Wahlen2017". Every hour, the growing data set was analyzed for the hashtags most frequently used in combination with this base set. If a threshold of 250 status updates containing a hashtag was met during the past 24 hours, the hashtag was considered a candidate to be added to the data collection query. All such candidates were manually assessed by two of the researchers and added to the collection set if they were clearly and without ambiguity related to the election. A total of 129 hashtags were added through this method. These hashtags include all major parties such as "#AfD", "#CDU", or "#SPD", campaign slogans such as "#DarumGrün", "#TrauDichDeutschland", or "#fedidwgugl", but also trending topics such as broadcasted television debates "#Kanzlerduell", "#TVDuell", or "#Elefantenrunde". By following this approach, we made sure not to miss trends in the online discussion and collected a large-scale representative data set of the political discussion on the German federal election. In total, we collected 3,090,932 status updates which can be distinguished into 882,155 tweets, 1,848,625 retweets, 201,368 replies, and 158,784 quote tweets. All statuses were posted by a total of 305,529 distinct users. This data set was then used to identify emergent user roles amongst the users who participated in the Twitter discussion.

3.2 Role Identification and Clustering Dimensions

A comprehensive overview on data-driven role identification in the tradition of social network analysis (SNA) is provided by Wasserman and Faust (1994) and Forestier et al. (2012). For OSNs, Füller et al. (2014) and Frank et al. (2017) demonstrated the applicability of hierarchical clustering for exploratory role identification, which in contrast to the well-known k-means clustering, is independent on the desired number of clusters as an input parameter. Further, users do not change their affiliation to a cluster even if the cluster solution is changed, which consequently helps to determine the appropriate number of clusters (Frank et al., 2017). As Forestier et al. (2012) and Füller et al. (2014) suggest, combining quantitative methodologies, such as cluster analysis and SNA, with qualitative interpretations and methodologies, such as content analysis, is considered a reasonable approach for identifying and understanding of emergent user roles. Gleave et al. (2009) also provide guidance for role analysis in online communities. They reason that an iterative process of both quantitative and qualitative measurements combined with both a bottom-up and a top-down approach results in the best role understanding. Thus, in this study we refine the result of our exploratory analysis, with further qualitative aspects, such as the content of the discussions, and relate it to previous research and typologies on Twitter and other OSNs.

The basic functionalities of Twitter serve as input dimensions for our cluster analysis. Twitter allows users to globally network with others by subscribing to their status updates (follow) and publish their own content through *tweets*. Tweets users send may contain up to 140 characters of plain text, links, or media (i.e. images or videos). Twitters' *retweet*-functionality allows users to share content across their personal network. By retweeting another user's tweet, the tweet appears as a copy on the retweeter's own personal timeline. A retweet itself can be retweeted, where only the original author's status is copied, creating a chain effect by amplifying information content sharing (Shi et al., 2014). When a tweet is annotated with own content and shared, the user's status update is considered a *quote tweet*. In contrast to retweets, quote tweets require more engagement from a user to be created. Additionally, members of the network can engage in discussions by replying to status updates. *Replies* to other users' status updates appear below the original post, creating a discussion board. Since all these types of status updates are disjunctive in Twitter's public stream data, we can assign a single type to each of the data record. Our eight clustering dimensions are structured as follows:

Structure features. In SNA, a set of two linked actors is called a dyad (Wasserman and Faust, 1994). Borgatti et al. (2009) differentiate between four types of dyadic ties through which actors can be linked: *proximities*, *relations*, *interactions*, and *flows*. In our study, we focus on interactions, which are "discrete, transitory relational events" (Kane et al., 2014, p. 282) on Twitter. For our analysis, relational ties between two actors are established when users interact with one another by using the basic functionalities of retweet, reply, or quote tweet. We mine the data set for status updates containing retweets, replies, or quote tweets, and consider the direction of the interaction to identify ties between users (i.e. who retweeted and who was retweeted). Activity or *outdegree* centrality of an actor in a network graph is indicated by a high level of outgoing ties. In contrast, the number of incoming ties, or *indegree* centrality, determines an actor's prestige (Wasserman and Faust, 1994). For our analysis, each retweet, reply, or quote tweet links two users and increases the out- and indegree of the particular users. *Function feature*. Besides above-mentioned ties, which are existing relationships between actors, our data set also contains status updates that have not (yet) resulted in a tie. A high number of *tweets* indicates a high level of activity in the political discussion. Therefore, we consider tweets a functional feature. *Time feature.* To capture users' temporal tweeting behavior, we applied the novel definition of the burstiness measure introduced by Kim and Jo (2016), which is able to deal with a finite number of events. Each status update is considered an event. A series of events is considered a burst if there are periods of high activity followed by periods of no or low activity. The burstiness measure is defined between $B = 1$ (most bursty events) and $B = -1$ (most periodic events).

Through the eight clustering dimensions (1) tweet, (2) outdegree retweet, (3) outdegree reply, (4) outdegree quote tweet, (5) indegree retweet, (6) indegree reply, (7) indegree quote tweet, and (8) burstiness we are able to capture the activity, prestige, and timed involvement of all users in the political discussion regarding the German federal election 2017 on Twitter. Subsequently, we adopt Frank et al. (2017) and apply agglomerative hierarchical clustering with the Ward.D2 minimum variance method, where Euclidean distances calculated on the normalized vectors of the cluster dimensions of each user serve as input parameters. To determine an appropriate number of clusters we follow Füller et al. (2014) and apply both Calinski and Harabasz (1974) pseudo F-statistic as well as Duda and Hart (1973) pseudo t^2-values as stopping rules.

4 Results

4.1 Descriptive Statistics

By mining the data set, we identified 305,529 users that are connected by 2,932,148 ties with interaction mostly happening through retweets (83.7%) followed by replies (9.1%) and quote tweets (7.2%). We find that the average user engages in interactions with 5.4 other users.

Table 1 displays descriptive statistics of the structure, function, and time features. We found that the density of all features, except for the burstiness measure B, is heavily skewed to the right (see minimum skewness of 43), plus the median for almost every dimension is zero. This indicates that many users within our data set have only used a limited number of features with marginal numbers of interactions. This pattern is common for OSN data sets. The data also reveals that on average a user posted three tweets, six retweets, one reply, and one quote tweet. Every interaction in our data set has an outbound and inbound user, therefore the means of in- and outdegree are identical. In contrast, median and standard deviation may differ in response to a high popularity of single posts or users (Frank et al., 2017).

	Function	*Structure*						*Time*
		Outdegree			*Indegree*			
	Tweet	**Retweet**	**Reply**	**Quote**	**Retweet**	**Reply**	**Quote**	**B**
Min / Max	0 / 5,885	0 / 8,282	0 / 1,201	0 / 4,656	0 / 64,115	0 / 3,856	0 / 3,004	-1.00 / 0.93
Mean / Median	2.89 / 0.00	6.05 / 1.00	0.66 / 0.00	0.52 / 0.00	6.05 / 0.00	0.66 / 0.00	0.52 / 0.00	0.03 / 0.00
SD	37.32	55.45	9.48	14.80	190.58	17.35	12.84	0.16
Skewness	66.16	43.82	45.48	209.67	166.63	122.87	103.32	-0.54
99.5% Quantile	465	1054	173	121	1764	151	138	NA

Table 1. Descriptive statistics of the eight clustering dimensions

The statistics indicate a strong focus of degree-centrality (mainly of indegree centrality) on particular users. However, quote tweets are somehow special in this regard, as they are more focused on users discussing status updates from many other users. The burstiness measure (Kim and Jo, 2016) spans almost across the entire codomain, therefore strict periodic behavior as well as highly bursty behavior can be found in the data set. However, since the number of Tweets per user varies heavily, this observation should be taken with a grain of salt. In summation, the descriptive statistics strongly support the perception that heterogeneous user behavior is present within the data set.

For our cluster analysis, each user is represented as a vector with the previously introduced eight features. To keep computation times manageable and to focus our study on relatively active users, we follow Nonnecke and Preece (2000) and focus on the 10% most active users. This results in a final data set of 30,553 users. In all dimensions, we noticed a heavy-tailed distribution and extreme outliers. To deal with those outliers, we truncated all values above the 99.5% quantile (see Table 1).

4.2 User Typology

As Bonner (1964) and Rokach and Maimon (2005) suggest "there is no universal definition for a good clustering size" (Rokach and Maimon, 2005, p. 326; Bonner, 1964). The stopping rules resulted in inconclusive results, however indicated a range of $k = 2$ to $k = 36$ clusters. But by exploring the paths of the main political actors, parties, media and journalists in the dendrogram, we found no remarkable shifts after $k = 11$ clusters. Further manual exploration of higher cluster solutions showed that while early splits are based on actual structural and behavioral differences, later splits are mainly related to varying levels of participation within the same user roles. In other words, neighboring branches have similar value peaks, but different heights in those peaks. Table 2 displays the result of the cluster analysis from which we derive the following typology.

Efficient Headliner. This user type comprises only a small part of the political discussants (0.4%). They participate fairly active in the political discussion (tweet 0.38) primarily by creation of own content, less by interactions with other users, but receive by far the most attention of all users (mean indegree 0.93).

Hence, their participation and engagement is very efficient and they are the most prestigious users (Wassermann and Faust 1994). They mainly create original content (tweet 0.38) and engage with or share other content considerably less (outdegree retweet 0.14, reply 0.09, quote 0.20).

			Function	*Structure*						*Time*
				Outdegree			*Indegree*			
Role		**#**	**Tweet**	**Retweet**	**Reply**	**Quote**	**Retweet**	**Reply**	**Quote**	**Burst.**
Efficient Headliner		115	0.38	0.14	0.09	0.20	0.98	0.86	0.96	0.66
Discussion Catalyst		169	0.10	0.04	0.01	0.03	0.21	0.75	0.66	0.63
Niche Headliner		295	0.20	0.06	0.05	0.09	0.44	0.21	0.29	0.66
Attention-seeker		73	0.68	0.72	0.63	0.70	0.68	0.50	0.53	0.80
Annotator		213	0.25	0.14	0.33	0.85	0.14	0.12	0.15	0.75
Discussant		198	0.13	0.23	0.90	0.09	0.08	0.24	0.06	0.73
Distributor		479	0.06	0.67	0.09	0.09	0.05	0.07	0.06	0.75
Screamer		121	0.99	0.09	0.09	0.02	0.07	0.04	0.03	0.78
Low-activity All-rounder		2,169	0.11	0.05	0.12	0.11	0.06	0.07	0.06	0.68
Occasional Participant	OP1	17,513	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.57
	OP2	9,208	0.03	0.04	0.01	0.01	0.01	0.01	0.01	0.71

Table 2. User typology with corresponding features of function, structure, and time

Discussion Catalyst. Only 0.6% share this user role which is fairly low in its engagement (mean outdegree 0.03). Still, they manage to create content which strongly appeals to other users (mean indegree 0.54) but receive attention mostly by replies and quote tweets rather than retweets. It seems that this group creates controversial and stimulating content inciting the political discussion.

Niche Headliner. This group represents one percent of the active users in the political discussions. Niche headliners do not excel in any dimension but are above average in their indegree. They mainly create original tweets, rarely distribute other content or engage in discussions. Although their contribution is rather low they are fairly prestigious (mean indegree 0.31). Mostly, their posts are retweeted, followed by replies and quote tweets. While this pattern is similar to Efficient Headliners, they receive less attention than Efficient Headliners, thus their audience seems to be more limited.

Attention-seeker. This is the smallest group in the network (0.2%). Although they are just a few users, their contribution to the political discussion is above-average for all functionalities. Their high participation and engagement raises awareness in the network (mean indegree of 0.57) and their contribution is the most bursty among all roles. Attention-seekers are prestigious and valuable members in the network who generate lots of content encouraging other users to participate. However, compared to Efficient Headliners, they show substantially more activity and reach less prestige.

Annotator. Annotators are characterized by distributing other users' tweets and commenting on them through quote tweets. Since commenting could be falsely associated with replying, we named this cluster Annotators. They rarely retweet other tweets, are above average in their creation of original content and their discussion activity but limited in the prestige that they receive (mean indegree 0.14). This group makes up only 0.7% of the active users. It remains to be seen whether Annotators are criticizing, praising or questioning other content before sharing, or if they are trying to raise awareness and credibility for their own opinion by quoting others.

Discussant. Discussants (0.6%) are characterized by using the reply function most frequently. Compared to that, they are using other Twitter functionalities less frequently, but still more often than many other users of the network. Discussants are actively taking part in discussions; however, they do not incite them. This is shown by the difference between outgoing and receiving replies.

Distributor. They account for a small group (0.6%) and engage much more by strictly retweeting than by any other activity. When Distributors participate, they do so by spreading other users' tweets in a somewhat bursty manner. Distributors do not receive much attention from others (mean indegree 0.06).

Screamer. This user type is characterized by its extremely high number of tweets. Screamers seldom use other functionalities of Twitter. Although all 121 members (0.4%) of this cluster actively contribute to the political discussion, their voice remains almost unheard (mean indegree 0.05). It can therefore be assumed, that despite their large number of tweets, the inherent information or opinions have little relevance for others and do not stimulate debates.

Low-activity All-rounder. This user role is the second largest in our study (7.1%) with no noteworthy peaks in any dimension. However, this can be due to a blurring effect in the means through the cluster analysis given the large cluster size. They do differ from Occasional Participants in such way, that they participate more actively and receive more attention from others.

Occasional Participants. Although we already focus only on the ten percent most active users, we still found that 87.5% show very low levels of activity in all dimensions, compared to the other user groups. Still, they are no *Lurkers* (Kollock and Smith, 1996) since their participation is higher than that of an average user. This group is not active enough to show meaningful and interpretable behavioral patterns and too large to be manually assessed. The main difference between OP1 and OP2 is the burstiness.

4.3 Refinement of the User Typology

Following Nadel's (1957) definition and Barley's (1990) interpretation of roles, they are compromised by both relational and non-relational elements. Our cluster analysis is solely based on non-relational elements. Therefore, in the following we first conduct an analysis of relational elements of the previously clustered roles by investigating main participants in political discussions. From a social perspective, the political discussion is mainly driven by voters, political parties, and the media, plus the associated individual politicians and journalists. Although voters presumably make up a large part of the users participating in the discussion, their identification is not unequivocally possible. As a first step towards identifying relational elements of the emergent user roles, we gathered the Twitter accounts of all parties participating in the recent election (Bundeswahlleiter, 2017d) and 544 accounts of all former and newly elected (current) members of the Bundestag (Deutscher Bundestag, 2017b, 2017a). The collection was done through lists of Twitter accounts provided by the parties themselves, containing the accounts of their mandate politicians. We retrieved the members of these lists by Twitter's REST API. Of those accounts, 24 parties and 283 politicians remained in our sample of the 10% most active Twitter users. For the second group, media and journalists, two researchers of our team manually assessed all 1,813 verified user accounts from our sample of active users to identify media outlets and self-proclaimed journalists based on user names and the users' self-descriptions. To measure interrater-reliability we calculated Cohen's Kappa and found a κ-value of 0.78, which according to Landis and Koch (1977) is a substantial agreement. The remaining differences were discussed within the team and resulted in a final list of 331 media outlets and 355 journalists. In doing so, we associated a total of 3.3% of the users to one of the four relational roles: party, politician, media, and journalist. Clearly, by this manual assessment we cannot expect to have identified a complete set of relational roles, but we are confident to have gathered a substantial sample. Despite that limitation, there are some remarkable associations between the relational roles and the identified emergent user roles. Table 3 displays how the relational roles were categorized. To summarize, Annotators, Discussants, Distributors, and Screamers are not occupied by any one of the relational roles. Neither media outlets nor journalists fall into the Attention-seekers cluster, and most, except for one, established parties are Efficient Headliner. Additionally, we retrieved each users number of followers to evaluate the average size of the network for each role.

Users consciously employ hashtags to refer to news, trending topics, or their political opinion. For example, the abbreviations of parties might be used to reference discussions to political matters of a particular party. The hashtag of the populist party "#AfD", dominated the online political discussion (60.1% of all party hashtags) and was of particular interest for users in every role. Interestingly, the distribution of party hashtags used in status updates is not at all representative of the election result. While CDU/CSU and SPD received more than half of all votes, the number of hashtags referencing these parties in our data set constitutes for only about a quarter of all hashtags. We employed a χ^2-independence-test to examine the existence of dependencies between user roles and their respective hashtag usage. To enable the test, we excluded hashtags which have been used less than five times by any role and found dependencies between the user roles and the hashtags they used (p-value < 0.001). By comparing the expected frequencies to observed frequencies, we identified surpassing and substandard utilization of hashtags in each role. In the following, we combine the findings from selected content analyses with the emergent user roles in detail to further characterize and enhance the user typology and get a comprehensive understanding of the behavioral patterns.

Role		Others	Parties	Politicians	Media	Journalists
Efficient Headliner		76	7	11	17	4
Discussion Catalyst		79	5	26	46	13
Niche Headliner		253	2	19	12	9
Attention-seeker		70	1	2	-	-
Annotator		213	-	-	-	-
Discussant		198	-	-	-	-
Distributor		479	-	-	-	-
Screamer		121	-	-	-	-
Low-activity All-rounder		2,046	2	24	50	47
Occasional Participant	OP1	16,958	6	159	183	207
	OP2	9,066	1	43	23	75
	#	29,559	24	284	331	355

Table 3. Relational roles occupied by the emergent user roles

Efficient Headliner. As the label suggests, this user role has a substantial network reach with 150,000 followers on average. Known politicians, established parties, mainstream media, and renowned journalists fall into the Efficient Headliner cluster. Almost all of the elected parties are categorized as Efficient Headliner. Given the subject of online political discussions, it is not surprising that parties are the main headliners of those discussions. The right-wing party AfD is the only elected party not categorized as an Efficient Headliner. However, the female frontrunners of AfD (Alice Weidel, Beatrix von Storch, Frauke Petry (left the party after the election)) are all categorized as Efficient Headliners. Only the lead candidates of the other minor parties achieved the same substantial attention (Cem Oezdemir of Grüne, Sarah Wagenknecht of Die Linke, Christian Lindner of FDP). Interestingly, not a single politician of the major parties CDU/CSU or SPD occupies this role. We further found that the hashtags "#AfD" and "#traudichdeutschland" (campaign slogan of AfD) were used more than expected by this user role. In summation, we can reason that Efficient Headliners do not need to participate in a highly active manner to receive remarkable attention since their recognition from the offline world is extended into the OSN.

Discussion Catalyst. By categorizing just a small fraction of the user base in the sample data set as either party, politician, media, or journalist we were able to identify most Discussion Catalysts (53.3%). The user role inhabits most mainstream media outlets, journalists, and a large number of well-known politicians. Official Twitter accounts of TV shows hosting political debates, such as Anne Will or Maybrit Illner share this role too. These TV formats function as a catalyst for political discussions during the election which is reflected in a high indegree due to attention received from the online community. We also find most mainstream media such as Das Erste, ZDF (both TV stations), Bild, Der Spiegel, or Süddeutsche Zeitung (all newspapers) in this cluster. Taking political actors into account, we identified a majority of renowned politicians of mainly the SPD as Discussion Catalysts; for example, the SPD's candidate for chancellor, Martin Schulz, or other widely known politicians such as Sigmar Gabriel (SPD), Peter Altmaier (CDU), or Katrin Göring-Eckardt (Grüne). Surprisingly, no AfD politician showed this behavior. Hence, Discussion Catalysts refer to "#AfD" or related topics less frequently than expected, but employ hashtags such as "#btw17", "#Schulz", or "#Merkel" more frequently.

Niche Headliner. The relational features of Niche Headliners underline the labelling of this user role. 19 politicians of all parties, except for CDU/CSU, are Niche Headliners and only two niche parties (Allianz Deutscher Demokraten, Demokratie in Bewegung) are among this user group. Some media outlets in this cluster have a regional focus, such as Hannoversche Allgemeine Zeitung or Stuttgarter Zeitung. In their hashtags Niche Headliners address political matters of Die Linke, FDP, or Grüne more than those of the major parties CDU/CSU and SPD. Their content does not significantly differ from other roles and hashtags referring to elected parties are not salient. Lastly, Niche Headliners have a substantial reach of 42,000 followers on average, which is smaller than that of Efficient Headliners, possibly due to the limited geographical reach of the outlets.

Attention-seeker. Two rather less known AfD politicans and most interestingly, the official Twitter account of the party AfD fall into the Attention-seeker cluster. No other party shares that behavioral pattern and no other party was as controversially discussed and covered by the press than the AfD. However, no media outlet or journalist occupies this user role. The hashtag usage shows a tendency towards right-wing political topics, which is no surprise considering the occupation by the AfD. The role averages only 4,800 followers, led by the AfD with 80,000, which is relatively low compared to all other elected parties who have a direct reach of at least 200,000 followers.

Annotator. None of our classified relational roles are included in the Annotators. We randomly selected several Annotators trying to identify their affiliation and social status but found mainly trolls and social bots in this user group. Their reach is limited due to their average number of followers (770) which might be a reason for their overall low indegree. We found that Annotators' content is mainly related to political subjects of the right-wing (e.g. "#DarumAfD", "#Reconquista", "#AfDwaehlen"). However, we have not observed more frequent status updates relating to "#AfD" than to other parties.

Discussant. None of the classified relational roles occupy this role. Discussants debate on candidates of the election more often than expected, but events such as "#TVDuell" are less frequent than expected.

Distributor. The label may indicate a large follower base, but the opposite is the case, with merely 1,000 followers on average. With more than 80% of their status updates containing at least a link or any type of media, they preferably share rich content. Distributors have a strong tendency towards topics of the right-wing and the AfD with the hashtags "#AfD" and "#TrauDichDeutschland" being used way more frequently than neutral references such as "#btw17" or "#DeineWahl".

Screamer. No other user role receives less attention from their activities than Screamers. By manually assessing a random sample drawn from this cluster, we found that a majority of accounts may be social bots, due to their enormous amount of total status updates and their high burstiness. Additionally, we identified some accounts of regional associations of the fringe party Freie Wähler. Screamers indisputably utilized trending hashtags like "#DeineWahl" but also hashtags such as "#notatarget" or "#humanitarian" which are unrelated to the German federal election. Therefore, we reason that Screamers are mostly made up of social bots who operationalize trending topics to get substantial reach.

Low-activity All-rounder. Politicians of all parties, many media outlets, and journalists were categorized into this user group. Various regional, niche, or foreign media outlets share the role of Low-activity All-rounders (i.e. NDR, BBC World, Heise Online, Buzzfeed Germany). Users in this cluster employ "#AfD" less than expected leading to a more homogeneous distribution of hashtag usage relating to parties. We also found that their status updates are often comprised by plain text.

Occasional Participants. Politicians of all elected parties are in this cluster. The content analysis revealed substantial differences between the two clusters categorized as Occasional Participants. OP1 uses the AfD related hashtags "#AfD" and "#TrauDichDeutschland" substantially less than expected and the subjects of their status updates are rather diverse. Additionally, the burstiness of 0.57 indicates a more continuous participation. There are only a few official accounts of parties in this cluster, possibly due to a professional, active management of the party accounts. In contrast, OP2 is a cluster presumably made up of users who were actively discussing the TV debate between chancellor Angela Merkel and her rival Martin Schulz, as indicated by an exceptionally higher than expected employment of the hashtag "#TVDuell". The high burstiness (0.71) may also correspond with the discussion of individual events.

5 Discussion

5.1 Theoretical Contribution

In our exploratory study, we analyzed through which interactions and behavioral patterns users participate in online political discussions on Twitter. Similar to other studies, we can confirm that the typical social structure in OSNs is composed of a small group of highly active users and a large fraction of passive users (e.g. Golder and Donath, 2004; Tinati et al., 2012; Füller et al., 2014; Shi et al., 2014). Some of our resulting emergent user roles are congruent with roles identified in previous studies from different research contexts. We adopted the label Discussion Catalyst from Himelboim et al. (2009) who reason that Discussion Catalysts wield disproportional influence by filtering and amplifying discussions through sharing content of news media in political discussions. Our results indicate that media outlets often act as such Discussion Catalysts, however, they do not need to create much content to receive attention on Twitter. This might be due to their offline prominence which transfers over into the social media realm. Yet, the Discussion Catalysts identified by Himelboim et al. (2009) focused only on a small topic area, while media outlets usually cover a broader range of topics. Annotators and Discussants share similar characteristics with roles described by Tinati et al. (2012): Discussants in our research correspond to curators, whilst Annotators show similarities with commentators. Tinati et al. (2012) reason that both of those roles add their own opinions to existing views and therefore act somehow similar. However, the authors reason that commentators (Annotators) are not contributing to an existing conversation, because they are not replying to an existing discussion thread, in contrast to curators (Discussants), but instead open a new one. It remains an open question whether Annotators and Discussants share the same impact on the flow of conversation on Twitter as Tinati et al. (2012) describes. For online innovation communities, Füller et al. (2014) identified socializers who prefer to engage in discussions with other users, which might also be related to the behavior of Discussants in OSNs. Socializers are interested in "socializing, gossiping, […] and gaining friendships" (Füller et al., 2014). Further content analysis needs to be conducted to verify, whether Discussants share this intention. Idea generators as identified by Füller et al. (2014) have similarities to Screamers due to high amount of content created (ideas or tweets respectively) and comparatively little attention received. Further related to Screamers, Kelly et al. (2006) identified fringes, as a group of users which try hard not to be ignored, but still are – probably because they inherit extreme views and are provoking (Kelly et al., 2006). All three roles have a very high activity in common; nevertheless, their effect differs. Further content analysis of the status updates of Screamers could reveal whether they are adding meaningful contributions or, referring to Golder and Donath's (2004) flamers and trolls, are characterized by their negative behavior. Füller et al. (2014) further introduce efficient contributors, which are described by rather low activity, receive a lot of attention and focus on content creation rather than socializing. Thus, efficient contributors have commonalities with our Efficient Headliners and Niche Headliners. Masters (Füller et al., 2014) and Attention-seekers utilize all given functionalities and receive substantial prestige. Unlike masters, Attention-seekers do not receive the most attention but are by far the most active users in our data. To receive their level of prestige, they have to be rather active. This paper also describes new emergent user roles, only found in this study based on Twitter interactions in the political context. Distributors spread other users' tweets via retweets. Even though they only have a small reach, they are relevant for the information flow as they may incite "chain effects" (Shi et al., 2014).

Bruns and Stieglitz (2012) identified distinct behavioral patterns by analyzing the use of hashtags on Twitter. They found that Twitter functions as a backchannel for live mainstream media events (e.g. TV debates). Users applying such hashtags mostly create own content using tweets, show few interactions with others and their status updates preferably contain plain text rather than links (Bruns and Stieglitz, 2012). The cluster OP2 shares similar characteristics. We found hashtags referencing mainstream media events like "#TVDuell", "#Wahlarena", "#Kanzlerduell" (all TV debates) were used much more often than expected in this cluster. Members of OP2 prefer plain text over rich content and their interactions are remarkably bursty. We conclude that users occupying the cluster OP2 use Twitter as a backchannel of these media events to share opinions and participate in ongoing discussions. Most users are associated to the cluster OP1 and reference neutral topics more frequently than expected and populist or right-wing topics substantially less than expected. Although the overall use of populist or right-wing topics were dominant in our data set, this finding suggests that the majority of the most active users were more deliberating than polarizing. Similar observations were found for Low-activity All-rounders. These results indicate that future research should question how these clusters are assembled regarding user ideology and their effects on political discourse or polarization.

5.2 Practical Implication

Our results help to understand how a multitude of different users contribute to the political discussion in OSNs. The emergent user roles show how different actors utilized social media in the recent election campaigning. Political actors can draw conclusions from our result to better organize their social media presence. Striking content is more appealing in OSNs and engagement in discussions and distributing other content can be as rewarding as creating own content. Some of the well-known public actors, like journalists or media outlets, can transform their offline popularity into online popularity and therefore gain a bonus in the online political discussion. Surprisingly, there is one political party (AfD) that shows significantly different behavior and roles than other parties. It seems to have had a major impact on the online political discussion and had a remarkable outcome in the election. Prior to the election, Alice Weidel (frontrunner of AfD) remarked that her party would employ the use of social bots to enrich their campaign. The AfD revised this statement later. All other parties explicitly denied to employ such tools (Tagesschau, 2017a, 2017b). Given the behavioral pattern of Distributors, Discussants, and Screamers, our results suggest that social bots may have been active, especially in the context of populist right-wing topics. However, this can hardly explain the entire AfD related content, especially since the Occasional Participants also employ AfD related hashtags, though less than expected. This might be because followers of populist parties "present themselves as agents of change" (Barr, 2009, p. 44), want their anti-establishment rhetoric to be heard, and thus might utilize Twitter as an open access social broadcasting network. Due to its large impact on the public debate and its remarkable share of overall content, the AfD took on a special role in the Twitter discussion and became a protagonist in this paper.

6 Conclusion, Limitations, and Outlook

In this study, we conducted an in-depth analysis of digital trace data from Twitter to gain a better understanding of the users' behavior in political discussions. Previous studies examined user roles in other OSNs, but not emergent roles based on Twitter trace data in political discussion. To extend upon this body of knowledge, we collected a data set on the German federal election on Twitter. This paper supports the growing interest of both practitioners and researchers on understanding behavioral patterns and the composition of groups of social media users. Based on basic functionalities of Twitter, we defined features of structure, function, and time to conduct cluster analysis regarding the behavior of the 30,553 most active users in our data set. As a result, we found eleven user roles. With further analysis, we support and enrich our user typology. We observe that amongst political parties and media outlets, Twitter is used similarly. It becomes evident that the offline popularity of some actors assists their presence on social media to a large extent. Further, political discussions in OSNs are not representative of the public opinion and the election results. Our results contribute to research in the field of the social sciences and politics, and studies implications of social media on society. In times of digitalization and OSNs becoming a part of our daily life, forming an own opinion on politics is more important than ever.

Our research has several limitations which allow for further analysis. The dynamic set of hashtags allowed us to capture a representative data set, but status updates were only collected if they employ one of our hashtags. Thus, we have not grasped the full discourse on particular tweets, especially regarding replies, as users usually do not include hashtags in their replies (Bruns and Stieglitz, 2012). Further, our content analysis is solely focused on the employment of hashtags. Additional analyses based on text-mining could be employed to get a better understanding of the content. Since this study is not focused on comprehensive methods to identify social bots, we focused on manual user assessment in this regard. Future research should thus address these limitations and validate our findings in different political discussions, elections, and countries. Additionally, it should be noted that we explicitly do not differentiate between human and bot operated Twitter users. Thus, further analysis regarding the detection of social bots (Varol et al., 2017) and their association to our emergent roles should be addressed. Users whose behavior is characterized by the excessive utilization of one functionality could be an interesting starting point for such an analysis. Although we characterized the most salient behavioral patterns, our study is limited to the most active fraction of Twitter users. Research could use our results to analyze less active Twitter users, as they are the largest proportion of Twitter users and eventually affect election results.

References

Adamic, L. A. and N. Glance (2005). "The political blogosphere and the 2004 U.S. election: Divided they blog." *Proceedings of the 3rd international workshop on Link discovery*, 36–43.

Arazy, O., J. Daxenberger, H. Lifshitz-Assaf, O. Nov and I. Gurevych (2016). "Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Coproduction." *Information Systems Research* 27 (4), 792–812.

Bakshy, E., S. Messing and L. A. Adamic (2015). "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348 (6239), 1130–1132.

Barberá, P. (2015). *How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S.* URL: http://pablobarbera.com/static/barbera_polarization_APSA.pdf (visited on 11/10/2017).

Barley, S. R. (1990). "The Alignment of Technology and Structure through Roles and Networks." *Administrative Science Quarterly* 35 (1), 61–103.

Barr, R. R. (2009). "Populists, Outsiders and Anti-Establishment Politics." *Party Politics* 15 (1), 29–48.

Beller, C., R. Knowles, C. Harman, S. Bergsma, M. Mitchell and B. van Durme (2014). "I'm a Belieber: Social Roles via Self-identification and Conceptual Attributes." *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), 181–186.

Bonner, R. E. (1964). "On Some Clustering Techniques." *IBM Journal of Research and Development* 8 (1), 22–32.

Borgatti, S. P., A. Mehra, D. J. Brass and G. Labianca (2009). "Network analysis in the social sciences." *Science* 323 (5916), 892–895.

Bruns, A. and S. Stieglitz (2012). "Quantitative Approaches to Comparing Communication Patterns on Twitter." *Journal of Technology in Human Services* 30 (3-4), 160–185.

Bundeswahlleiter (2017a). *42 Parteien nehmen an der Bundestagswahl 2017 teil - Pressemitteilung Nr. 10/17 vom 8. August 2017*. URL: https://www.bundeswahlleiter.de/info/presse/mitteilungen/bundestagswahl-2017/10_17_teilnahme.html (visited on 11/20/2017).

Bundeswahlleiter (2017b). *Bundestagswahl 2017: 61,5 Millionen Wahlberechtigte - Pressemitteilung Nr. 01/17 vom 3. Februar 2017*. URL: https://www.bundeswahlleiter.de/info/presse/mitteilungen/bundestagswahl-2017/01_17_wahlberechtigte.html (visited on 11/15/2017).

Bundeswahlleiter (2017c). *Bundestagswahl 2017: Ergebnisse*. URL: https://www.bundeswahlleiter.de/bundestagswahlen/2017/ergebnisse.html (visited on 11/15/2017).

Bundeswahlleiter (2017d). *Bundestagswahl 2017: Wahlbewerberinnen und Wahlbewerber*. URL: https://bundeswahlleiter.de/bundestagswahlen/2017/wahlbewerber.html (visited on 11/15/2017).

Calinski, T. and J. Harabasz (1974). "A dendrite method for cluster analysis." *Communications in Statistics - Theory and Methods* 3 (1), 1–27.

Conover, M., J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini and F. Menczer (2011). "Political Polarization on Twitter." *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*.

Deutscher Bundestag (2017a). *Deutscher Bundestag - Biographien 18. Wahlperiode (2013 - 2017)*. URL: https://www.bundestag.de/abgeordnete/biografien18 (visited on 11/15/2017).

Deutscher Bundestag (2017b). *Deutscher Bundestag - Biographien aktuelle Wahlperiode*. URL: https://www.bundestag.de/abgeordnete/biografien (visited on 11/15/2017).

Duda, R. O. and P. E. Hart (1973). *Pattern classification and scene analysis*. John Wiley & Sons.

Fisher, D., M. Smith and H. T. Welser (2006). "You Are Who You Talk To: Detecting Roles in Usenet Newsgroups." *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*.

Flaxman, S., S. Goel and J. M. Rao (2016). "Filter Bubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80 (S1), 298–320.

Forestier, M., A. Stavrianou, J. Velcin and D. A. Zighed (2012). "Roles in Social Networks: Methodologies and Research Issues." *Web Intelligence and Agent Systems* 10 (1), 117–133.

Frank, L., H. Gimpel, M. Schmidt and M. Schoch (2017). "Emergent User Roles of a Digital Workplace: A Network Analysis Based on Trace Data." *Proceedings of the 38th International Conference on Information Systems*.

Füller, J., K. Hutter, J. Hautz and K. Matzler (2014). "User Roles and Contributions in Innovation-Contest Communities." *Journal of Management Information Systems* 31 (1), 273–308.

Gleave, E., H. T. Welser, T. M. Lento and M. Smith (2009). "A Conceptual and Operational Definition of 'Social Role' in Online Community." *Proceedings of the 42nd Hawaii International Conference on System Science*, 1–11.

Golder, S. A. and J. S. Donath (2004). "Social Roles in Electronic Communities." *Association of Internet Researchers conference Internet Research 5.0*.

Gonzalez-Bailon, S., A. Kaltenbrunner and R. E. Banchs (2010). "The structure of political discussion networks: A model for the analysis of online deliberation." *Journal of Information Technology* 25 (2), 230–243.

Grömping, M. (2014). "'Echo Chambers': Partisan Facebook Groups during the 2014 Thai Election." *Asia Pacific Media Educator* 24 (1), 39–59.

Handelsblatt (2017). *Bundestagswahl: Alle gegen die AfD*. URL: http://www.handelsblatt.com/politik/deutschland/bundestagswahl/alle-schlagzeilen/bundestagswahl-alle-gegen-die-afd/20343226-all.html (visited on 11/21/2017).

Himelboim, I., E. Gleave and M. Smith (2009). "Discussion catalysts in online political discussions: Content importers and conversation starters." *Journal of Computer-Mediated Communication* 14 (4), 771–789.

infratest dimap (2015). *Die Positionierung der politischen Parteien im Links-Rechts-Kontinuum: Afd rückt nach rechts, CDU nach links*. URL: https://www.infratest-dimap.de/uploads/media/LinksRechts_Nov2015_01.pdf (visited on 11/15/2017).

Kane, G. C., M. Alavi, G. Labianca and S. P. Borgatti (2014). "What's Different about Social Media Networks?: A Framework and Research Agenda." *MIS Quarterly* 38 (1), 274–304.

Kelly, J., D. Fisher and M. Smith (2005). "Debate, Division, and Diversity: Political Discourse Networks in USENET Newsgroups." *Online Deliberation Conference*.

Kelly, J., D. Fisher and M. Smith (2006). "Friends, foes, and fringe." *Proceedings of the 2006 national conference on Digital government research*, 412–417.

Kim, E.-K. and H.-H. Jo (2016). "Measuring burstiness for finite event sequences." *Physical Review E* 94 (3), 325–331.

Kollock, P. and M. Smith (1996). "Managing the virtual commons." *Computer-Mediated Communication*, 109–128.

Kushin, M. J. and K. Kitchener (2009). "Getting political on social network sites: Exploring online political discourse on Facebook." *First Monday* 14 (11).

Landis, J. R. and G. G. Koch (1977). "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33 (1), 159.

Lerner, J. (2005). "Role Assignments." In: *Network analysis: Methodological foundations*. Ed. by U. Brandes and T. Erlebach. Springer, p. 216–252.

Mirbabaie, M., C. Ehnis, S. Stieglitz and D. Bunker (2014). "Communication Roles in Public Events." *Information Systems and Global Assemblages. (Re)Configuring Actors, Artefacts, Organizations* 446, 207–218.

Mirbabaie, M. and E. Zapatka (2017). "Sensemaking in social media crisis communication - a case study on the Brussels bombings in 2016." *Proceedings of the 25th European Conference on Information Systems (ECIS)*, 2169–2186.

Münchner Merkur (2017). *Bundestagswahl 2017: Warum gibt es keine Koalition mit der AfD?* URL: https://www.merkur.de/politik/warum-keine-koalition-mit-afd-nach-bundestagswahl-2017-zr-8715056.html (visited on 11/19/2017).

Nadel, S. F. (1957). *The Theory of Social Structure*. Cohen & West.

Newman, N., R. Fletcher, A. Kalogeropoulos, D. A. L. Levy and R. K. Nielsen (2017). *Reuters Institute Digital News Report 2017*. URL: http://po.st/lfJFXh (visited on 11/15/2017).

Nolker, R. D. and L. Zhou (2005). “Social Computing and Weighting to Identify Member Roles in Online Communities.” *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, 87–93.

Nonnecke, B. and J. Preece (2000). “Lurker demographics.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, 73–80.

Pervin, N., H. Takeda and F. Toriumi (2014). “Factors Affecting Retweetability: An Event-Centric Analysis on Twitter.” *ICIS 2014 Proceedings*, 1–10.

Rokach, L. and O. Maimon (2005). *Clustering Methods*. Springer.

Schmidt, A. L., F. Zollo, M. Del Vicario, A. Bessi, A. Scala, G. Caldarelli, H. E. Stanley and W. Quattrociocchi (2017). “Anatomy of news consumption on Facebook.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (12), 3035–3039.

Shi, Z., H. Rui and A. B. Whinston (2014). “Content sharing in a social broadcasting environment: Evidence from Twitter.” *MIS Quarterly* 38 (1), 123–142.

Stieglitz, S., M. Mirbabaie, L. Schwenner, J. Marx, J. Lehr and F. Brünker (2017). “Sensemaking and Communication Roles in Social Media Crisis Communication.” *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*, 1333–1347.

Sunstein, C. R. (2001). *Republic.com*. Princeton University Press.

Tagesschau (2017a). *Manipulation durch Social Bots: Kampf den Meinungsmaschinen*. URL: http://www.tagesschau.de/inland/social-bots-wahlkampf-101.html (visited on 11/18/2017).

Tagesschau (2017b). *Social Bots im Wahlkampf: AfD verzichtet auf Meinungsroboter - oder nicht?* URL: http://faktenfinder.tagesschau.de/social-bots-bundestag-wahl-101.html (visited on 11/18/2017).

Tagesschau (2017c). *Taktisches Wählen: Spekulieren mit dem Stimmzettel*. URL: https://www.tagesschau.de/inland/btw17/taktisch-waehlen-101.html (visited on 11/19/2017).

Tinati, R., L. Carr, W. Hall and J. Bentwood (2012). “Identifying communicator roles in twitter.” *Proceedings of the 21st international conference companion on World Wide Web*, 1161–1168.

Tumasjan, A., T. O. Sprenger, P. G. Sandner and I. M. Welpe (2010). “Predicting elections with Twitter: What 140 characters reveal about political sentiment.” *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*.

Twitter (2013). *New Tweets per second record, and how!* URL: https://blog.twitter.com/engineering/en_us/a/2013/new-tweets-per-second-record-and-how.html (visited on 11/15/2017).

Twitter (2017a). *"My use of social media is not Presidential - it’s MODERN DAY PRESIDENTIAL. Make America Great Again!" of @realDonaldTrump*, Posted on 1st Juli 2017. URL: https://twitter.com/realdonaldtrump/status/881281755017355264 (visited on 11/15/2017).

Twitter (2017b). *Q3 2017 Earnings: Selected Company Metrics and Financials*. URL: https://investor.twitterinc.com/financials.cfm (visited on 11/15/2017).

Varol, O., E. Ferrara, C. A. Davis, F. Menczer and A. Flammini (2017). “Online Human-Bot Interactions: Detection, Estimation, and Characterization.” *Proceedings of the 11th International AAAI Conference on Weblogs and Social Media*.

Wasserman, S. and K. Faust (1994). *Social network analysis:* Methods and applications. Cambridge University Press.

Welser, H. T., E. Gleave, D. Fisher and M. Smith (2007). “Visualizing the Signatures of Social Roles in Online Discussion Groups.” *The Journal of Social Structure* 8 (2).

World Economic Forum (2017). *The Global Risks Report 2017 - 12th Edition*. URL: http://www3.weforum.org/docs/GRR17_Report_web.pdf (visited on 11/15/2017).

Yardi, S. and D. Boyd (2010). “Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter.” *Bulletin of Science, Technology & Society* 30 (5), 316–327.