A Big Social Media Data Study of the 2017 German Federal Election based on Social Set Analysis of Political Party Facebook Pages with SoSeVi

Benjamin Flesch[1], Ravi Vatrapu[1,2], Raghava Rao Mukkamala[1]
[1]Copenhagen Business School, Denmark and [2]Westerdals Oslo School of Arts, Comm & Tech, Norway
{bf.digi, rv.digi, rrm.digi}@cbs.dk

***Abstract*—We present a big social media data study that comprises of 1 million individuals who interact with Facebook pages of the seven major political parties *CDU*, *CSU*, *SPD*, *FDP*, *Greens*, *Die Linke* and *AfD* during the 2017 German federal election. Our study uses the Social Set Analysis (SSA) approach, which is based on the sociology of associations, mathematics of set theory, and advanced visual analytics of event studies. We illustrate the capabilities of SSA through the most recent version of our Social Set Analysis (SoSeVi) tool, which enables us to deep dive into Facebook activity concerning the election. We explore a significant gender-based difference between female and male interactions with political party Facebook pages. Furthermore, we perform a multi-faceted analysis of social media interactions using gender detection, user segmentation and retention analysis, and visualize our findings. In conclusion, we discuss the analytical approach of social set analysis and conclude with a discussion of the benefits of set theoretical approaches based on the social philosophical approach of associational sociology.**

Keywords*—Big social media data, Social set analysis, Big data visual analytics, Facebook, 2017 German federal election, Bundestagswahl, CDU, CSU, SPD, FDP, Grüne, AfD, Linke

I. Introduction

This paper applies *Social Set Analysis* research approach to the 2017 federal election in Germany, more precisely to the activity on the major political parties' Facebook walls. *Social Set Analysis* is a research approach situated in the domains of Data Science [1]–[3] and Computational Social Science [4] with practical applications to Big Social Data Analytics in organizations [5]–[7]. It addresses one of the important theoretical and methodological limitations in the emerging paradigm of Big Data Analytics of social media data [8]. In particular, it address the major limitation in existing research on Big Social Data analytics that computational methods, formal models and software tools are largely limited to graph theoretical approaches [9] (such as SNA [10]), and are informed by the social philosophical approach of relational sociology [11]. There are no other unified modeling approaches to social data that integrate the conceptual, formal, software, analytical and empirical realms [12]. This results in a research problem when analyzing Big Social Data from platforms like Facebook and Twitter as such data consists of not only dyadic relations but also individual associations [13]. For Big Social Data analytics of Facebook or Twitter data, the fundamental assumption of SNA that social reality is constituted by dyadic relations and interactions that are determined by structural positions of individuals in social networks [14] is neither necessary nor sufficient [15]. Previous versions of the *Social Set Visualizer* tool have been introduced to showcase the *Social Set Analysis* approach [16].

For example, consider a Facebook post made on the official Facebook wall of Lionel Messi, the soccer prodigy who plays for FC Barcelona and Argentina's national football team. Each official post by Messi to his Facebook page typically receives more than 100,000 likes, 25,000 comments and 18,000 shares. Such association-based and content-driven social media interactions involving large number of social actors are unlike the other social interactions such as face-to-face, email, phone and instant messaging in the sense that what binds the interacting social actors together in the first instance is not so much the relational ties (strong vs. weak ties) but associations ranging from the player himself, the teams that he plays for, to the cultural, ethnic, national and linguistic attributes. Modeling such Facebook interactions using affiliation networks creates the problem of an extremely low number of nodes with an extremely high number of nodes as spokes. Further, such SNA assumes the central social psychological concept of "homophily" that social actors with similar interests (that is, associations) prefer to interact with each other. To overcome this limitation and address the research problem, this paper proposes an alternative holistic approach to Big Social Data analytics that is based on the sociology of associations and the mathematics of set theory and offers to develop fundamentally new methods and tools for Big Social Data analytics, Social Set Analysis (SSA). Our overarching research question is stated as, *How, and in what way, can methods and tools for Social Set Analysis derived from the alternative holistic approach to Big Social Data analytics based on the sociology of associations and the mathematics of set theory result in meaningful facts, actionable insights and valuable outcomes?*

The rest of the paper is organized as follows. First, we present a philosophical template for holistic approaches to computational social sciences, compare and contrast the dominant approach of social network analysis with the proposed novel approach of social set analysis and discuss the benefits of set theoretical approaches based on the social philosophical approach of associational sociology in Sec II. Second, we present the most recent version of our Social Set Visualizer (SoSeVi) tool in III.

Third, we take a deep dive into Facebook activity concerning the 2017 German federal election held on 24th of September 2017 on a political party level. Section IV illustrates

978-1-5386-2715-0/17/$31.00 ©2017 IEEE

the capabilities of SoSeVi by showcasing growth and retention of audience by political parties, user segmentation into loyalists and persons with positive and negative feelings towards a political party, and further analyses based on first names and gender classification.

Fourth and last, we discuss the findings from our illustrative case study, offer methodological and analytical reflections on social set analysis, identify its limitations, and outline future work directions. We have not provided any dedicated section for related work, but we have referred the relevant literature at appropriate places throughout the paper.

II. Theoretical Framework

Social Set Analysis (SSA) as employed in this paper is concerned with the mobility of social actors across time and space. For mobility across time, we conduct SSA of big social data from the Facebook walls of the seven major political parties in Germany, with an analytical focus on the set of actors that interacted with the parties during the 2017 federal election campaign. Similarly, for mobility across space, we conduct set inclusions and exclusion of actors who interacted with different Facebook walls. This will allow us to uncover not only the interactional dynamics over time and space but also identify actor sets that correspond to marketing segmentations such as loyalists, advocates, critics and activists. The theoretical framework and the formal model behind our proposed approach of Social Set Analysis have been elaborated in previous papers such as [16] [15].

III. Social Set Visualizer (SoSeVi) tool

A. User interface

The Social Set Visualizer (SoSeVi) tool for Social Set Analysis has been under active development since 2014, with older version presented to the research community in several papers. The latest version focuses on Upset-inspired [17] visualization set intersections, and is paired with a built-in Facebook crawler. The set intersection visualization allows researchers to define social media interactions in a set query language, and then perform further analysis based on the set of individuals at hand which resulted from the query.

Figure 1(a) showcases the latest version of the Social Set Analysis user interface provided by the Social Set Visualizer tool. After selection of Facebook pages of interest, the user can compare these Facebook pages in an Upset-inspired [17] set visualization tailored to the Social Set Analysis approach. Social Set Visualizer (SoSeVi) provides means to segment individuals on social media and visualize their interactions. Word cloud visualization and aggregated Facebook page information is shown in figure 1(b).

To summarize, the SoSeVi big data visual analysis dashboard empowers users to use it in many different ways. The dashboard adheres to the user's preferred interaction method without making any assumptions. This means tablet users may also type in their selection of the Facebook walls, or desktop users may use the date picker to manually select a date. The dashboard may be accessed at http://rf2017.roonk.de/.

(a) Selection of Facebook pages and time period of interest as preparation for a visualization of set overlaps and intersection cardinalities.

(b) Facebook page overview with alphabetical word cloud and Facebook reaction visualization.

Figure 1: User interface provided by Social Set Visualizer.

B. Technology

The technology choice for realizing the dashboard visualizations is the D3.js Javascript-based visualization framework which uses dynamic SVG images for data visualization. D3.js constitutes a lightweight and very extendable Javascript visualization framework which can display visualizations for a multitude of browser-based clients. The flexibility provided by D3.js enables the creation of new kinds of interactive visualizations which are able to run on any device with decent processing resources including Windows, MacOS and Linux based systems with screen sizes up to 4K devices. Data is stored in a relational database and heavily indexed using *PostgreSQL*. Queries are cached both in database tables and in-memory using *Redis*.

IV. 2017 German federal election case study

A. Background

The 2017 German federal election held on 24th of September 2017 was the largest political event in recent years. Major topics such as the European migrant crisis [18], central bank policies [19] and workplace equality [20] have put pressure on incumbent Angela Merkel, her cabinet and the political parties *CDU*, *CSU* and *SPD* closely affiliated with her. Both pro-business liberal party *FDP* and the green party *Bündnis '90 / Die Grünen* aim to get more foothold with mainstream voters than in previous years.

More extreme political parties such as recently formed *Alternative für Deutschland* (AfD) and leftist party *Die Linke* contest voters' mind share and aim to get more influence in the future government. Based on these circumstances, we take a deep dive into social media reactions on the major political parties' Facebook pages to better understand the state of mind of the political parties' audiences and ultimately, the German voters.

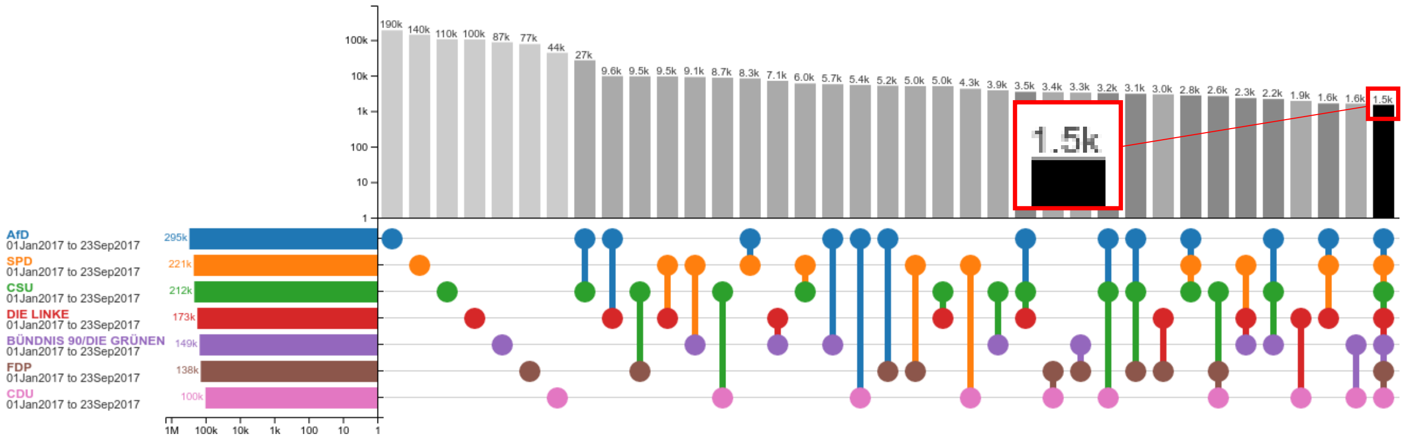


Figure 2: SoSeVi-based set visualization of Facebook audience overlap between major political parties in Germany during the 2017 federal election campaign. The overlap between all seven political parties is represented by the black bar on the right side of the visualization. The grid-based set overlap visualization using interconnected circles is inspired by Upset [17].

B. Methodology

Our research methodology consisted of several steps. First, we fetched the Facebook walls of the major political parties in Germany: Angela Merkel's *CDU*, Bavarian *CSU*, social democratic *SPD*, liberal *FDP*, green party *Bündnis '90 Die Grünen*, leftist party *Die Linke* and ultra-conservative alternative party *AfD*. For this, we use a self-made Facebook crawler. Furthermore we restrict our observation timeframe to the beginning of 2017 up until the day before the federal election, 23rd of September 2017. Second, we analyze collected Facebook activity with our *Social Set Visualizer* (SoSeVi) tool. We visualize overlaps between individual parties' Facebook audiences and illustrate inner-party retention rates throughout the hot phase of the election campaign. Third, we perform deep dives into audience segments of interest and illustrate the capabilities of SoSeVi by addressing party loyalist, audience reactions and demographics such as the most common first names of individuals interacting with the Facebook walls. Fourth, we discuss our findings and deploy the dashboard internally to support ongoing research.

C. Data Collection & Processing

The event timeline of the 2017 federal election was collected through desk research including systematic searches in web and media databases. Facebook data was previously collected through the Social Data Analytics Tool (SODATO) [21]–[23]. For this paper a SoSeVi-internal crawler was used to provide Facebook data shown in table I. The general concept follows the stages of the "Big Data Value Chain" introduced by Miller and Mork [24], with steps of preparation, organization and integration of the data prior to visualization and analysis. The aggregated data is then imported into a database management system (DBMS), from which it can be accessed for visual analytics purposes.

D. Size of political party Facebook audience

In figure 2 SoSeVi is utilized to visualize a total of 958,834 individuals who interacted with German political party Facebook pages during the 2017 federal election. This number

Party	Posts	P.Reactions	Comments	C.Reactions
AfD	970	2,107,255	445,978	1,031,180
CDU	550	374,830	152,904	364,261
CSU	598	985,812	142,078	455,527
FDP	652	592,527	80,403	106,132
GREEN	442	361,351	97,309	214,113
LINKE	609	607,137	104,082	246,823
SPD	531	719,632	121,215	229,401

Table I: Overview of Facebook dataset of major German political parties

is also displayed in figure IV as all-party total. We examine the aggregate number of individuals that interacted with each parties' Facebook page during the examination period up to 23rd of September 2017, as visualized through the left-side horizontal bar chart in figure 2.

It strikes that newcomer *AfD* leads with a total of 295,000 individuals, followed in second place by social democrats *SPD* who interacted with 221,000 individuals. Third largest is Bavarian-only *CSU* party with 212,000 individuals active on their page, the sister party of Angela Merkel's *CDU*. *CDU* themselves are in last place, because only 100,000 individuals interacted with their Facebook page during the 2017 federal election campaign. All minor parties such as the *FDP* with 138,000 individuals, the *Green party* with 149,000 and the leftist party *Die Linke* with 173,000 had Facebook interactions with more unique individuals than Angela Merkel's ruling party *CDU*.

E. Audience overlap between political party Facebook pages

In figure 2 we also visualize overlaps of Facebook audiences between the major political parties in Germany in the 2017 federal election period from 1st of January to 23rd of September 2017. We use Social Set Analysis approach to

	April	May		June		July		August		September		All months	
Party	#	% chg	#	% chg	#	% chg	#	% chg	#	% chg	#	Sparkline	CMGR
AfD	**63.1k**	-8%	**58.4k**	8%	**63.4k**	-6%	**60.0k**	35%	**92.8k**	40%	**155.0k**		**19.7%**
CDU	**11.8k**	16%	**14.1k**	18%	**17.2k**	32%	**25.4k**	-11%	**22.8k**	51%	**46.7k**		**31.7%**
CSU	**39.9k**	-18%	**33.7k**	11%	**37.9k**	33%	**56.8k**	-22%	**46.5k**	30%	**66.0k**		**10.6%**
FDP	**25.1k**	6%	**26.8k**	-25%	**21.5k**	34%	**32.5k**	29%	**45.8k**	25%	**61.1k**		**19.5%**
GREEN	**18.6k**	11%	**20.9k**	17%	**25.1k**	-47%	**17.1k**	58%	**40.4k**	10%	**44.9k**		**19.3%**
LINKE	**18.7k**	52%	**39.3k**	0%	**39.3k**	-2%	**38.5k**	8%	**41.8k**	52%	**86.7k**		**35.9%**
SPD	**31.4k**	-18%	**26.7k**	27%	**36.8k**	-10%	**33.5k**	54%	**72.5k**	13%	**83.8k**		**21.7%**

Table II: Monthly growth rate of unique individuals who interacted with German political party Facebook pages during the 2017 federal election campaign between 1st of January and 23rd of September 2017. Sparklines visualize month with lowest and highest number of individuals on Facebook page. Compound monthly growth rate is calculated and compared.

calculate sets of individuals and visualize overlaps between the sets at hand. Major two-set overlaps between political parties are:

1) We observe that more than 27,000 individuals were active both on the *CSU* and the *AfD* Facebook pages, displaying the biggest audience overlap between two political parties.
2) The second major audience overlap is between *AfD* and leftist party *Die LINKE* with 9,600 individuals.
3) The third largest overlap is between Bavarian *CSU* party and liberal *FDP* party with more than 9,500 individuals active on both parties' Facebook pages.
4) Fourth largest overlap is between social democrats *SPD* and leftist *Die Linke* with *9,500* individuals, followed by fifth largest overlap between *SPD* and the *Green* party with 9,100 individuals active on both Facebook pages.
5) Angela Merkel's *CDU* and her Bavarian sister party *CSU* depict the sixth largest overlap with 8,700 individuals.

Further overlaps between political party Facebook audiences are visualized in the figure, but due to space restrictions we cannot list all of them. The major overlaps identified seem to follow the parties' closeness on the political spectrum, even though at the moment we cannot explain the detailed reason for the relative differences in cardinality between overlaps such as *CSU/AFD* and *SPD/Die Linke*.

F. Audience growth during election campaign

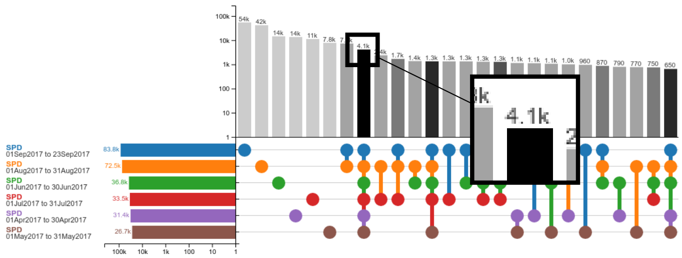
The audience growth rate in terms of the total number of individuals who were active on a certain political party's Facebook page during the campaign is showcased in table II. Using social set analysis we create sets of individuals who interacted with a certain party for each month of the election campaign. Cardinalities of monthly sets for each political party have been taken from the set visualizations of figure 3. Based on this data, a compound monthly growth rate (CMGR) has been calculated to compare each party's audience growth during the time period of the election campaign. We observe the following:

1) For all parties, the final month of campaigning, September, was the best month in terms of total number of individuals they interacted with.
2) No party showcases a steady, consistent growth story. All of them have at least one month where they actually decreased their audience compared to the previous month.
3) Comparing the compound monthly growth rate (CMGR), both leftist *LINKE* (+35.9%) and Angela Merkel's party *CDU* (+31.7%) depict the biggest growth over the whole period of investigation. Both are also the only parties where both April is the overall weakest month and September the overall peak.
4) With only 10.6% over the whole campaign, *CSU* showcased the lowest overall growth rate.
5) All other parties *SPD*, *FDP*, *GREEN*, and *AfD* expressed a compound growth rate of around 20% per month.
6) In August, penultimate month of the 2107 election campaign, current chancellor Angela Merkel's parties *CDU* and *CSU* both decreased in the number of individuals that interacted with their Facebook pages by a total of 69.3k people (-11% and -22% respectively). This is interesting because one would expect that during August, at the peak of campaigning, both sister parties would continue to push very hard. This decrease could be explained with summer holidays for the shared campaigning team.
7) Also in August, *SPD*, the biggest rival of *CDU/CSU*, grew their audience at 54%. With a total of 72.5k individuals, *SPD* reached a larger audience on Facebook than both *CDU* (22.8k) and CSU (46.5k) combined.

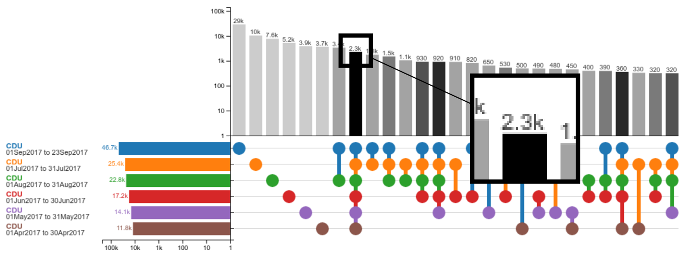
G. Audience retention of political party Facebook pages

We visualize month-over-month retention of Facebook audience for each political party in Germany from 1st of April up to election day 24th of September 2017. For this purpose we create six monthly slices (April, May, June, July, August, September until 23rd) for each party and utilize SoSeVi to perform social set analysis on them.

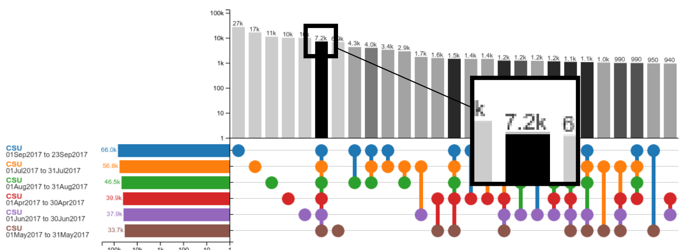
Using the example of social democrat party *SPD*, we visualize in Figure 3(a) the month-to-month development of individuals interacting with the party's Facebook page. The visualization shows that 4,100 individuals interact with the party on a monthly level, and the vast majority of users interact with the party's Facebook page on a very loose basis. Even though we see a steady month-to-month growth between April and September, the retention of individuals seems to be lacking. In September up until election day, a total of 83,000 individuals interacted with *SPD*, but 54,000 of those only did so in September and not in any prior months. The visualization for social democrat *SPD* party can be accessed



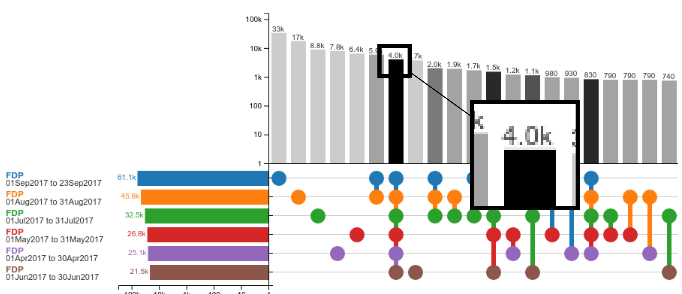
(a) SPD



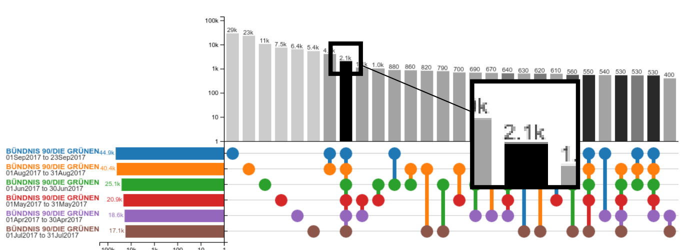
(b) CDU



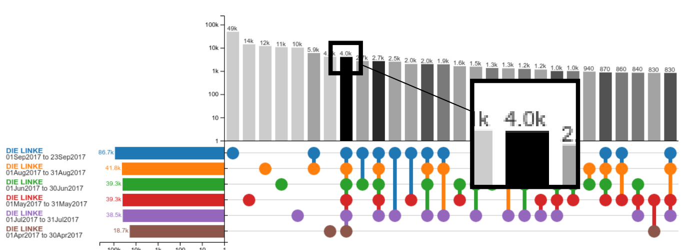
(c) CSU



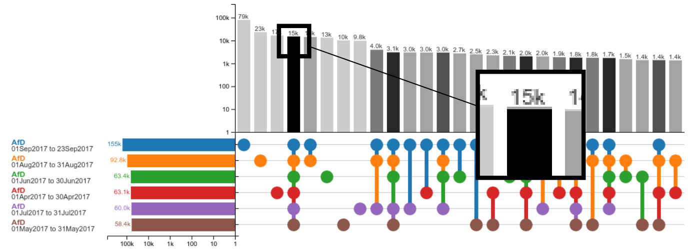
(d) FDP



(e) Green party



(f) Leftist party



(g) AfD

Figure 3: SoSeVi-based visualization of month-over-month development and retention of Facebook audience for German political parties, sliced monthly until election day. Loyalist audience for each party is depicted by the black vertical bar spanning all six month-based sets.

online at rf2017.roonk.de/upset. Likewise we visualize audience retention for other political parties in figure 3.

H. Identification of political party loyalists on Facebook

We define political party loyalists as the set of individuals who interact with a certain party's Facebook page at least once per month. For this purpose we examine monthly slices for the six months preceding election day, same as in previous section IV-G. We determine loyalist audience from the number of individuals who are active on a specific party's Facebook page in every single month within the observation period up to election day. Using figures 3(a), 3(b), 3(c), 3(d), 3(e), 3(f), 3(g) we can determine the total number of loyalists for each political party.

In order to put the absolute size of party loyalist audience in perspective, we compare official party membership numbers to the size of the loyalist audience on Facebook and calculate a ratio. Number memberships has been collected for each party from official publications [25]. Table III showcases that loyalist Facebook audience varies highly between political parties. The massive membership bases of the two major parties *SPD* and *CDU* are not significantly more active on the parties' Facebook pages than the loyalist audiences of smaller parties.

Relative to the total number of party memberships, small parties such as *AfD*, *FDP*, and the Leftist party *Die Linke*

Party	Total party members	# loyalists on Facebook	% of members
SPD	432,706	4,100	0.95%
CDU	431,920	2,300	0.53%
CSU	142,412	7,200	5.06%
GREEN	61,596	2,100	3.41%
LINKE	58,910	4,000	6.79%
FDP	53,896	4,000	7.42%
AfD	26,409	15,000	56.80%

Table III: Comparison of political party loyalist audience on Facebook and official party membership numbers. Membership numbers are from 31. December 2016 and based on official publications [25].

Party	M/F ratio party memberships*	All individuals						Reaction to Posts			Reaction to Comments			Writing Comments		
		Total	Female	Male	N/A	%	M/F Ratio	Female	Male	M/F Ratio	Female	Male	M/F Ratio	Female	Male	M/F Ratio
AfD	5.25	294,951	77,514	188,409	29,028	9.84%	2.43	56,768	146,909	2.59	34,613	70,435	2.03	22,011	60,539	2.75
CDU	2.85	100,494	25,371	62,839	12,284	12.22%	2.48	17,016	40,635	2.39	9,575	22,879	2.39	6,021	19,757	3.28
CSU	4.00	212,019	58,728	136,509	16,782	7.92%	2.32	50,216	115,146	2.29	17,494	36,853	2.11	10,149	29,149	2.87
FDP	3.35	137,717	31,414	96,032	10,271	7.46%	3.06	26,732	80,585	3.01	6,778	21,322	3.15	5,095	20,875	4.10
GREEN	1.56	148,626	59,325	72,102	17,199	11.57%	1.22	51,037	48,682	0.95	11,934	24,215	2.03	6,625	18,450	2.78
LINKE	1.70	172,902	50,006	100,937	21,959	12.70%	2.02	42,904	80,663	1.88	14,925	30,792	2.06	7,639	22,540	2.95
SPD	2.13	220,904	71,674	119,264	29,966	13.57%	1.66	60,460	90,220	1.49	15,237	31,983	2.10	9,370	27,908	2.98
All parties	2.53	958,834	296,772	548,827	113,235	11.81%	1.85	254,271	461,346	1.81	88,916	176,913	1.99	55,121	149,985	2.72

Table IV: First name based gender classification of social media actors on political party Facebook pages during the 2017 German federal election. Official party member gender ratio is based on 2016 data published by German federal ministry for political education (BPB) [26]. *N/A* displays failed gender classification.

interact with a high number of individuals compared to their total memberships. *AfD* in particular is rapidly growing with a low number of official party memberships, thus the high percentage of 56.80%. Compared with peers, the *Green* party receives only a small amount of loyalist interaction on the Facebook page, both in absolute numbers but also as a relative percentage to their peers in terms of party memberships *Die Linke* and *FDP*.

I. Audience reactions to political party Facebook posts

Table V showcases Facebook reactions by individuals to posts by political parties. For this analysis we count the number of individuals who interact with the party post with a Facebook reaction, focusing on the most widely used Facebook reactions LOVE, LIKE, SAD, ANGRY and HAHA. We observe:

1) Far-right *AfD* received reactions from more than 225k audience members. This is 40k more people than the next biggest parties, *CSU* (180k) and *SPD* (175k).
2) Every party except Angela Merkel's *CDU* received reactions from more than 110k individuals to their Facebook postings. In total, only 66k users reacted to *CDU* posts.
3) Used by more than 90% of all individuals, *LIKE* depicts major audience reaction to political party posts.
4) Liberal *FDP* receives a *LIKE* from 95% of their interacting Facebook audience.
5) Receiving *LIKEs* from 202k individuals, far-right *AfD* significantly eclipses Angela Merkel's *CDU* which only receives *LIKE*s from 56k individuals during the campaign.
6) Far-right *AfD* received *ANGRY* reactions from 51k individuals or 23% of their audience.
7) Reactions other than *LIKE* are not very frequently used, major exception being the numerous *ANGRY* reactions towards *AfD* posts.

Party	# individuals with reaction to posts by a political party										# unique individuals
	ANGRY		SAD		HAHA		LOVE		LIKE		
CSU	16,751	9%	8,389	5%	20,858	12%	5,952	3%	165,783	92%	179,348
AfD	51,179	23%	11,854	5%	37,084	16%	21,144	9%	201,968	90%	225,160
SPD	5,332	3%	5,399	3%	12,676	7%	11,689	7%	159,372	91%	175,649
LINKE	8,933	6%	4,044	3%	6,039	4%	12,528	9%	133,882	94%	142,540
GREEN	5,867	5%	4,656	4%	8,961	8%	9,152	8%	100,961	89%	113,722
FDP	3,893	3%	2,556	2%	6,910	6%	5,940	5%	110,054	95%	115,658
CDU	4,093	6%	2,386	4%	7,811	12%	5,077	8%	56,363	84%	66,718

Table V: Audience reactions to political party Facebook posts during 2017 election campaign.

J. Comparing Facebook gender distribution with official party membership data

Table IV displays the results of gender-based Facebook audience segmentation. In the first part of the table, we show aggregate numbers for each political party. The center of the table shows audience reactions to posts and to comments by the political party, aggregated by gender. Furthermore, we show male/female comment authorship. In the final row of the table we plot the official male to female ratio based on party membership publications for comparison. We perform audience segmentation by gender to showcase the full potential of Social Set Analysis. Audience interactions with political party Facebook pages are analyzed along this dimension. Gender inference is performed based on the first name of the Facebook user at hand. We use the $nam\_dict.txt$ database[1] to link first names with genders. This technique for gender inference has been successfully applied by other researches such as [27]. Based on male/female audience segmentation of German political party Facebook walls as shown in table IV we can point out several qualitative findings:

1) Both on an aggregate and on an individual level, discussion and reactions on the political parties' Facebook pages appear male-dominated, with a male-to-female ratio as high as 4.10 for comment authorship on *FDP* page.
2) The only exception to this observation are reactions to posts on the *Green party*'s page. With post reactions from 51,037 females and only 48,682 males, this is the only dimension in table IV where we can count more females than males interacting with the party's posts.
3) Incumbent ruling party *CDU* has the fewest individuals on their Facebook page, less than half as many as their biggest rival, the social democrat party *SPD*.
4) Leftist party *Die Linke* is the only political party where the Male/Female ratio of all dimensions of interaction (post and comment reactions, comment authorship) with their Facebook page is higher than the official Male/Female ratio based on their party memberships.
5) Apart from the Leftist party *Die Linke*, all other parties have a more balanced male-to-female ratio on their Facebook page than the male-to-female ratio based on official party memberships numbers suggests.

[1] nam_dict.txt first-name based gender classification database (c) 2008 Jörg Michael, available at https://www.heise.de/ct/ftp/07/17/182/

(a) Female first name distribution across political parties

Party \ First name	PETRA	SABINE	ANDREA	ANNA	CLAUDIA	JULIA	SANDRA	MONIKA	NICOLE	KARIN	SUSANNE	SARAH	HEIKE	LISA	MARIA	ANJA	MARTINA	KERSTIN	BIRGIT	MARION	CHRISTINE	TANJA	KATHARINA	LAURA	BRIGITTE
AfD	.34	.31	.30	.21	.26	.18	.30	.25	.27	.21	.18	.18	.23	.17	.18	.21	.19	.21	.17	.20	.17	.19	.11	.10	.14
CDU	.29	.31	.28	.24	.26	.21	.19	.23	.22	.22	.22	.15	.19	.14	.22	.17	.19	.17	.17	.17	.18	.14	.17	.14	.15
CSU	.47	.35	.37	.20	.33	.17	.24	.41	.21	.34	.23	.08	.24	.13	.24	.17	.23	.21	.26	.24	.25	.15	.13	.07	.29
FDP	.26	.27	.26	.25	.25	.26	.20	.18	.20	.17	.19	.16	.14	.16	.15	.18	.15	.15	.16	.13	.13	.12	.18	.14	.12
GREEN	.41	.47	.40	.56	.36	.48	.32	.30	.28	.28	.35	.38	.26	.38	.29	.27	.24	.24	.25	.20	.27	.19	.30	.34	.20
LINKE	.33	.30	.28	.29	.25	.25	.23	.23	.20	.22	.20	.22	.20	.23	.18	.19	.19	.20	.18	.18	.16	.14	.16	.16	.15
SPD	.34	.33	.29	.34	.30	.33	.27	.26	.26	.22	.22	.28	.25	.28	.22	.22	.20	.18	.19	.18	.17	.19	.20	.23	.18
All parties %	.35	.32	.31	.31	.29	.28	.27	.26	.26	.23	.23	.23	.22	.22	.22	.21	.20	.20	.20	.19	.19	.18	.18	.18	.18
# Actors [1000s]	3.3	3.1	3.0	2.9	2.8	2.7	2.6	2.5	2.5	2.2	2.2	2.2	2.1	2.1	2.1	2.0	1.9	1.9	1.9	1.8	1.8	1.7	1.7	1.7	1.7
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Party \ First name	MELANIE	DANIELA	ELKE	BARBARA	ANNE	EVA	ANGELIKA	NADINE	MANUELA	STEFANIE	GABRIELE	CHRISTINA	SIMONE	ALEXANDRA	UTE	RENATE	JANA	SABRINA	SONJA	MICHAELA	LENA	NINA	KATJA	INGRID	KATRIN
AfD	.18	.17	.16	.12	.11	.12	.16	.16	.19	.14	.14	.12	.14	.12	.13	.11	.12	.15	.13	.13	.08	.08	.12	.11	.12
CDU	.13	.13	.15	.16	.18	.13	.18	.12	.13	.12	.15	.12	.13	.11	.13	.11	.11	.07	.10	.12	.10	.07	.09	.11	.11
CSU	.13	.17	.22	.23	.12	.17	.26	.10	.20	.14	.21	.14	.15	.15	.16	.23	.09	.10	.15	.18	.05	.07	.09	.21	.10
FDP	.11	.14	.12	.13	.15	.15	.11	.13	.09	.12	.13	.12	.10	.13	.11	.09	.10	.07	.08	.08	.12	.13	.10	.11	.11
GREEN	.24	.20	.19	.23	.29	.27	.16	.19	.13	.20	.18	.19	.18	.16	.17	.13	.19	.15	.21	.15	.27	.23	.19	.13	.18
LINKE	.13	.14	.16	.15	.18	.15	.15	.14	.12	.12	.13	.11	.12	.13	.13	.13	.14	.10	.11	.10	.12	.15	.13	.11	.12
SPD	.18	.16	.19	.16	.18	.16	.17	.19	.13	.15	.14	.15	.13	.14	.15	.15	.16	.15	.13	.14	.17	.15	.14	.14	.12
All parties %	.17	.17	.17	.17	.17	.17	.17	.16	.16	.15	.15	.15	.15	.14	.14	.14	.14	.14	.14	.14	.13	.13	.13	.13	.13
# Actors [1000s]	1.7	1.7	1.6	1.6	1.6	1.6	1.6	1.6	1.5	1.4	1.4	1.4	1.4	1.4	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.2
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50

(b) Male first name distribution across political parties

Party \ First name	MICHAEL	THOMAS	CHRISTIAN	ANDREAS	PETER	DANIEL	STEFAN	FRANK	MARTIN	MARKUS	JAN	HANS	SEBASTIAN	ALEXANDER	MATTHIAS	KLAUS	SVEN	PATRICK	TOBIAS	MARCEL	FLORIAN	JÜRGEN	JENS	MARCO	WOLFGANG
AfD	1.6	1.4	1.1	1.2	1.1	.93	.89	.96	.75	.73	.52	.62	.55	.57	.55	.59	.65	.60	.45	.61	.40	.57	.58	.56	.49
CDU	1.6	1.4	1.1	1.2	.88	.81	.84	.79	.70	.71	.64	.62	.61	.62	.59	.52	.50	.46	.46	.46	.48	.45	.48	.43	.45
CSU	1.8	1.6	1.2	1.4	1.2	.81	1.1	.91	.81	.91	.49	.89	.61	.67	.65	.78	.48	.48	.51	.37	.57	.73	.49	.46	.72
FDP	1.6	1.4	1.5	1.2	.90	1.1	1.0	.82	.87	.83	1.0	.59	.89	.92	.74	.52	.61	.64	.74	.53	.69	.48	.57	.49	.50
GREEN	1.1	1.0	.78	.81	.66	.64	.64	.53	.63	.50	.58	.39	.49	.47	.45	.40	.35	.36	.38	.26	.40	.35	.34	.28	.34
LINKE	1.3	1.0	.88	.94	.83	.82	.74	.71	.67	.53	.62	.53	.58	.54	.46	.47	.43	.47	.41	.43	.41	.43	.42	.42	.41
SPD	1.1	.95	.80	.76	.63	.69	.63	.57	.58	.52	.61	.43	.52	.44	.44	.43	.41	.40	.42	.41	.41	.36	.33	.33	.37
All parties %	1.2	1.0	.90	.90	.77	.75	.73	.66	.63	.60	.58	.54	.53	.52	.48	.47	.46	.46	.43	.43	.43	.42	.42	.40	.40
# Actors [1000s]	12	10	8.7	8.6	7.4	7.2	7.0	6.3	6.0	5.8	5.5	5.2	5.1	5.0	4.6	4.5	4.4	4.4	4.1	4.1	4.1	4.1	4.0	3.8	3.8
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Party \ First name	UWE	RALF	DAVID	DIRK	OLIVER	BERND	PHILIPP	JÖRG	DENNIS	ROBERT	CHRISTOPH	MAX	ALEX	SASCHA	MARC	CHRIS	TIM	MARIO	KEVIN	FELIX	STEPHAN	LUKAS	JOHANNES	PAUL	STEFFEN
AfD	.59	.53	.46	.51	.43	.47	.29	.48	.46	.44	.30	.31	.38	.44	.36	.42	.29	.53	.43	.22	.33	.20	.21	.31	.39
CDU	.41	.43	.37	.43	.46	.44	.42	.41	.39	.32	.48	.37	.30	.37	.35	.27	.37	.28	.28	.36	.38	.36	.38	.31	.29
CSU	.61	.56	.30	.45	.45	.59	.34	.48	.30	.44	.44	.38	.34	.34	.34	.34	.28	.35	.23	.26	.39	.26	.39	.24	.32
FDP	.40	.45	.50	.48	.57	.47	.80	.43	.49	.45	.60	.62	.44	.40	.51	.47	.56	.26	.36	.60	.41	.54	.51	.34	.36
GREEN	.33	.31	.39	.31	.33	.33	.34	.28	.24	.30	.34	.30	.30	.23	.27	.27	.28	.21	.17	.32	.25	.29	.31	.24	.22
LINKE	.42	.40	.43	.36	.35	.36	.35	.35	.36	.35	.31	.38	.40	.34	.29	.36	.33	.30	.32	.36	.28	.33	.25	.34	.27
SPD	.33	.32	.38	.33	.33	.32	.37	.29	.32	.27	.31	.32	.27	.30	.29	.26	.34	.21	.33	.33	.24	.31	.26	.24	.21
All parties %	.40	.39	.39	.38	.37	.37	.36	.36	.36	.35	.34	.34	.33	.33	.33	.33	.32	.31	.31	.30	.29	.29	.28	.27	.27
# Actors [1000s]	3.8	3.7	3.7	3.6	3.5	3.5	3.4	3.4	3.4	3.4	3.3	3.3	3.2	3.1	3.1	3.1	3.1	3.0	3.0	2.9	2.8	2.8	2.7	2.6	2.6
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50

Table VI: Comparative visualization TOP50 most frequent male and female first names across German political party Facebook pages. Gender detection is performed based on first name. Colored areas display each party's Facebook audience having a certain first name as a percentage of the whole dataset. The number of actors at the bottom of each table concerns the absolute number of individuals in our dataset who hold a certain first name.

K. Gender-based differences in Facebook interactions with German political parties

Furthermore we examine whether there is a statistically significant difference between male and female individuals in their interaction with German political party's Facebook pages during the period of the 2017 federal election. For this purpose we perform a chi-square test of gender-based differences in engagement with 6 degrees of freedom. The test shows a significant difference between males and females, with $p < 0.05$ and $\chi^2 = 17825.46$.

Potential limitations of this finding are the extent and veracity of our first name based gender classification approach. We have manually verified gender classification results for the top 100 most frequently used first names, yet the long-tail correctness of classification results has not been thoroughly examined. The name lists underlying our gender classification approach is targeted at German-speaking population and does not capture all names from other cultural backgrounds. Table IV depicts gender classification results. A total of 113,235 actors (12%) have not been successfully classified. To further test for gender-based differences in Facebook interaction, we should assume that all non-gender-classified first names are female, and repeat the chi-square test. Again it shows that the finding is significant with $p < 0.05$ and $\chi^2 = 20944.96$.

L. Frequency analysis of first names across political party Facebook audience

In table VI we visualize top 50 most frequently occurring first names across all individuals interacting with political party pages during the election campaign. More specifically, table V(a) depicts the frequency distribution of overall top 50 female first names and how often these first names are observed in each political party within the time period of the election campaign. Table V(b) provides the same information for all individuals that were classified as males based on their first names.

The visualization of top 50 female first names in table V(a) provides insight into party-specific distribution of first names. Facebook audience of the *GREEN* party exhibits above-average frequency of female first names, e.g. *ANNA* (0.56% vs. global average 0.31%) and *JULIA* (0.48% vs. global average 0.28%). *CSU* displays higher variance than the *GREEN* party: With 0.08% of global audience, names such as *SARAH* are significantly less frequent on the *CSU* page than it would be expected given the 0.23% overall average. Conversely, table V(b) depicts overall top 50 male first names from our dataset and their frequency across political parties. As shown in section IV-J, political party Facebook audience within the 2017 German federal election campaign is overwhelmingly male. Viewing the male first name visualization this becomes apparent through the fact that most frequencies are about two to three times higher than in table V(a). No significant trends are visible to the eye.

The top 50 first names returned by the gender-focused approach in this section largely mimic historic demographics of Germany, and thus don't provide significant findings apart from several outliers and slight trends between political parties.

Table VII: Top 50 most uniquely attributable first names for each political party in the 2017 German federal election. Numbers depict the percentage share of all individuals with a certain first name interacting with the respective party's Facebook page.

(a) AfD

Party \ First name	RONNY	SYLVIO	IAN	RINGO	GARY	ENRICO	SILVIO	RICO	RENE	DENNY	STEVE	PIOTR	MAIK	ROCCO	COLIN	ROY	ROBBY	ONKEL	MIKE	TINO	RICCARDO	MARIO	BRIAN	ANDREW	HANSI
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	64	63	62	60	59	58	57	57	57	57	56	55	54	54	54	54	54	53	53	53	52	52	51	51	51
CDU	9	12	6	5	3	10	8	9	8	7	8	9	10	12	6	4	7	5	10	11	11	9	6	7	9
CSU	20	20	16	22	13	19	21	19	22	16	16	14	19	21	11	16	17	25	24	24	16	24	16	16	35
FDP	9	11	10	9	12	8	9	9	10	6	9	5	11	5	14	9	10	6	14	13	11	12	10	12	9
GREEN	8	10	10	14	8	9	8	10	7	8	9	7	8	12	15	11	13	12	9	10	8	10	9	8	8
LINKE	18	12	9	22	12	18	20	18	16	16	16	11	17	20	15	18	21	13	16	17	18	17	14	14	14
SPD	12	10	11	10	12	15	11	14	15	10	12	16	16	15	17	11	12	10	16	13	8	15	21	12	14
# Total Actors	1824	124	168	123	134	1305	668	792	2371	333	950	136	2456	208	154	337	162	125	3032	882	163	4229	276	221	238
# Actors in AfD	838	56	84	52	67	550	285	331	996	158	422	64	984	81	63	148	65	54	1129	332	68	1555	112	95	87

Party \ First name	ANDY	INGOLF	JEFF	ANTHONY	ADOLF	JAMES	DANNY	THOR	FALKO	MANDY	MIRKO	TOMMY	HEIKO	MARKO	ROGER	TONY	KEN	ANDRZEJ	JIM	BOB	JOHN	PAVEL	EDDY	RONALD	RICKY
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	51	50	50	50	50	50	50	50	49	49	49	49	48	48	48	48	48	48	48	48	48	48	48	48	48
CDU	10	6	12	6	7	10	8	11	11	7	10	7	11	10	9	10	6	9	8	7	10	7	9	10	5
CSU	24	28	11	10	33	15	17	23	19	13	19	21	27	22	24	12	12	15	18	18	15	16	21	29	20
FDP	13	14	13	8	7	8	13	16	11	6	14	12	14	12	16	8	9	9	10	11	12	10	7	14	8
GREEN	9	9	7	11	12	5	10	13	8	12	11	10	12	12	12	9	10	11	11	18	11	9	6	10	17
LINKE	18	19	12	8	9	14	18	21	13	17	18	17	17	14	15	21	22	13	20	14	16	14	22	15	21
SPD	14	13	25	21	18	17	15	21	13	15	19	16	17	16	15	13	18	11	15	18	17	15	17	16	14
# Total Actors	1866	197	130	132	152	309	1052	155	245	728	1223	587	2945	982	478	487	153	102	171	168	1194	111	233	772	112
# Actors in AfD	684	71	50	58	56	129	399	50	98	303	428	218	974	353	165	193	59	42	63	60	442	45	86	259	40

M. Top 50 most uniquely attributable first names for each political party

Table VII showcases an alternative approach to providing a unique perspective on the Facebook audience of German political parties. For each party, we identify the top 50 first names that are most uniquely attributable to the party at hand. We calculate relative percentage share of all audience members with a certain first name and select the top 50 highest percentage first names for each party. First names with less than 100 individuals and party names are filtered out. For example with AfD in table VII(a), we can see in the leftmost column that 64% of all individuals with the first name *Ronny* interact with the *AfD* Facebook page, while only 9% of all *Ronnys* interact with *CDU* page. The total number of individuals named *Ronny* in our data set is 1834, of which 838 (64%) interact with *AfD* during the campaign. We further examine the most uniquely attributable first names for each party and describe our findings:

1) *AfD* VII(a): Most uniquely attributable names are "stereotypical" for the eastern part of Germany. Frequency distribution heavily skewed towards *AfD*.
2) *CDU* VII(b): Mainly Arabic first names, but overall very low level of uniqueness (percentages less than 40%), many shared with *SPD*.
3) *CSU* VII(c): Traditional German names, both male and female, with percentages between 40 and 50%.
4) *FDP* VII(d): German male first names.
5) *GREEN* VII(e): German female first names.
6) *LINKE* VII(f): First names with some Turkish background, most likely related to immigrant workers during the early days of German federal republic.
7) *SPD* VII(g): First-ranked *TC* means *Türkiye Cumhuriyeti* (Republic of Turkey), Turkish activists added TC in front of their name to signal their support of Turkey during a shit storm including *SPD*. Most unique names related to *LINKE*, but Arabic names in long tail as shown in VII(b).

V. Discussion

Due to space restrictions, we presented only a subset of the empirical findings resulting from the use of the Social Set Visualizer (SoSeVi) tool by researchers and practitioners in various fields such as Corporate Social Responsibility (CSR), Computational Social Sciences (CSS) and healthcare. These empirical findings demonstrate the analytical utility of our proposed set theoretical approach to big social data and our social set analysis implementation in the SoSeVi visual analytics dashboard.

A. Reflections on the IT-Artifact

Computational social science research has reached a point where social media activity is ubiquitous yet hard to collect and analyze in domain-specific ways (with the notable exception of epidemiology). In conjunction with complex event timelines as depicted by the 2017 German federal election, the data at hand presents numerous opportunities for attaining deep insights. In this context, visual analytics present the means of reaching those insights to many users with different backgrounds, both experts and novices alike. The novel implementation of the present Social Set Visualizer (SoSeVi) dashboard showcases that the creation of visual analytics software, which meets the high technical, analytical and user experience requirements of present-day computing, is viable (and can be achieved by an academic research group with limited resources). Furthermore, the developed IT artifact leverages open-source visual analytics frameworks to maximum extent in order to achieve a pure implementation of important concepts in visual analytics.

B. Reflections on the Set Theoretical Approach

The current paradigm in computational social science is dominated by a theoretical focus on relationships of actors and artifacts, and the mathematical modeling of those relationships as social networks based on graph theory.

This leads to the big social data triumvirate of relational sociology (as candidate social philosophy), graph theory (as candidate mathematical and formal model), and social network analysis (as candidate analytical framework). Our argument is not that relational sociology, graph theory, and social network analysis are invalid or ineffective. Social Network approaches have proven their analytical suitability and ability in diverse application domains ranging from epidemiology to organizational behavior. Instead, our argument is that other candidate sociological approaches, mathematical theories, and analysis techniques need to be explored to further advance the field of computational social science. After all, relational sociology is just one of the many competing and co-existing theories in sociology describing, explaining and predicting social phenomena; along with process, ethnomethodology, structuration, identity, structural functionalism, cognitive and cultural theories. Our paper's primary contribution to not only to offer an alternate holistic approach of social theory (associations),

(b) CDU

Party \ First name	محمد	MAHMOUD	ANAS	MUHAMMAD	MOHAMAD	AHMAD	KHALED	AHMED	ALWIN	MOHAMMED	JAMAL	HUSSEIN	MOHAMMAD	MECHTHILD	BENEDICT	SAID	LUDGER	ADNAN	OMAR	MOHAMED	HAMID	ALBRECHT	REZA	SAMIR	SAMUEL
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	5	13	15	10	11	12	10	14	35	14	11	12	9	16	19	7	20	9	17	14	12	22	11	9	21
CDU	36	34	32	31	30	29	27	27	26	25	25	24	24	24	24	21	21	21	20	20	19	19	19	19	18
CSU	9	5	11	7	8	7	8	10	26	3	8	6	7	27	18	5	30	7	7	8	4	34	8	9	18
FDP	6	3	6	3	0	3	2	4	19	2	4	3	4	13	33	2	20	5	5	6	4	18	4	9	21
GREEN	19	12	13	15	14	15	19	14	13	11	7	17	17	19	15	16	16	15	13	10	15	12	16	13	15
LINKE	11	18	8	14	17	14	19	13	12	17	24	18	12	9	17	20	15	23	16	18	19	16	19	19	15
SPD	38	34	40	51	48	44	38	47	19	46	40	37	50	27	22	50	22	45	41	47	45	17	48	32	23
# Total Actors	126	187	125	189	270	668	200	472	103	323	106	135	569	101	176	137	212	192	306	503	135	179	139	167	458
# Actors in CDU	37	53	32	45	64	155	44	99	18	68	22	28	112	18	28	24	31	32	52	81	22	24	21	28	64

Party \ First name	SAMI	GERO	MIC	HEI	ANDREE	ARND	JORGE	CLAUDIUS	AZIZ	BENEDIKT	ABDUL	MAGNUS	SÖNKE	TIMON	HERIBERT	HASSAN	FERDI	IRA	ANSGAR	MARTEN	TAREK	EDDA	GEORGE	FRIEDEL	CARL
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	13	20	32	44	40	26	35	26	10	18	8	20	20	20	35	9	22	22	20	26	17	19	31	33	25
CDU	18	18	18	18	18	18	17	17	17	17	17	17	17	17	17	17	16	16	16	16	16	16	16	16	16
CSU	9	15	28	34	22	27	15	31	11	29	9	19	14	13	39	10	32	20	17	17	10	28	18	37	24
FDP	8	18	19	14	13	27	12	23	7	30	2	32	26	26	11	4	19	16	18	26	7	5	10	16	29
GREEN	15	13	17	18	13	12	21	16	10	16	18	16	22	18	10	12	10	14	13	17	10	19	10	12	11
LINKE	20	18	19	19	15	13	15	9	23	13	19	19	14	21	9	23	16	15	20	22	21	14	23	15	15
SPD	40	27	26	22	14	18	20	15	38	20	49	16	21	25	20	43	27	23	15	22	40	31	18	24	14
# Total Actors	195	121	115	123	162	144	148	127	158	957	209	272	167	160	161	266	130	100	146	146	143	106	309	125	479
# Actors in CDU	29	17	13	13	21	18	19	16	23	114	29	33	21	19	19	37	15	13	20	16	19	13	39	13	56

(c) CSU

Party \ First name	ALOIS	WALTRAUD	SIEGLINDE	SEPP	HEIDEMARIE	GUENTER	ANNELIESE	ALFONS	ELFRIEDE	JOSEF	HUBERT	MARGA	GERHARD	EDITH	KARLHEINZ	BRUNHILDE	WINFRIED	MARGIT	EDELTRAUD	REINHOLD	KORBINIAN	GÜNTER	WERNER	ELFI	HANNELORE
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	39	25	28	43	26	30	24	38	26	38	32	26	37	26	38	36	31	26	22	34	18	40	38	31	26
CDU	10	10	11	10	8	9	12	14	7	12	9	12	10	10	8	9	10	7	12	12	11	12	12	11	11
CSU	51	49	47	47	46	45	44	43	43	43	43	43	43	43	42	42	42	42	42	42	41	41	41	41	41
FDP	10	8	8	7	9	15	9	14	2	13	15	7	14	12	11	6	16	10	9	14	24	13	15	10	11
GREEN	12	10	11	8	14	14	10	7	7	10	11	12	12	14	9	10	15	15	9	9	19	12	12	11	10
LINKE	10	10	12	9	20	14	13	10	12	12	13	8	15	14	14	13	14	15	15	17	9	17	17	14	17
SPD	11	25	18	8	22	18	24	17	20	14	17	24	18	20	17	19	24	17	21	16	18	20	19	20	22
# Total Actors	293	485	244	245	179	244	235	263	162	1803	707	113	2343	532	440	106	460	565	196	609	132	2059	3241	182	684
# Actors in CSU	104	174	85	87	57	76	76	80	60	552	218	37	672	162	135	33	128	180	63	177	39	546	859	54	202

Party \ First name	LUDWIG	ERWIN	ROSEMARIE	DIETER	WOLFGANG	RUDOLF	GÜNTHER	ALFRED	HERIBERT	FRANZ	HERBERT	IRMGARD	ANNEMARIE	ROSWITHA	LEONHARD	MANFRED	HARALD	GERLINDE	JÜRGEN	WALTER	BERNHARD	HELMUT	HILDEGARD	GISELA	HORST
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	31	41	25	39	38	38	40	41	35	37	38	26	22	28	20	41	40	25	42	39	34	39	25	26	40
CDU	11	9	10	12	12	11	9	12	17	11	11	11	12	7	6	10	10	8	11	10	11	11	11	10	12
CSU	40	40	40	40	40	40	40	40	39	39	39	39	39	39	39	39	38	38	38	38	38	38	38	38	38
FDP	15	10	11	14	18	15	13	10	11	14	13	8	12	8	20	15	16	9	16	14	15	14	11	9	14
GREEN	13	10	14	12	13	10	10	10	10	11	11	14	17	17	11	12	12	18	13	10	13	10	15	15	10
LINKE	15	11	17	18	19	14	16	14	9	15	15	10	13	15	15	18	17	14	18	15	14	18	12	15	18
SPD	18	21	23	21	21	16	18	20	20	15	18	27	19	21	11	20	20	20	20	20	16	20	22	24	19
# Total Actors	727	758	480	3754	6138	964	874	998	161	2379	1702	352	334	685	129	3769	2654	353	6435	1814	1755	2407	374	1241	2395
# Actors in CSU	206	214	139	956	1522	267	240	268	45	655	460	102	98	198	41	946	663	102	1558	474	472	609	107	342	600

(d) FDP

Party \ First name	JUSTUS	JULIUS	TILMAN	CONSTANTIN	NICOLAS	BENEDICT	CORNELIUS	MAGNUS	MAXIMILIAN	PHILIPP	NIKLAS	KILIAN	HENRIK	MORITZ	HUBERTUS	MATS	NIKOLAS	JANIK	FREDERIK	BENEDIKT	THORBEN	FREDERIC	HENDRIK	MORTEN	HAUKE
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	10	15	9	15	14	19	17	20	23	25	19	20	22	16	23	15	15	25	16	18	22	11	22	25	16
CDU	13	13	11	13	12	24	12	17	16	12	13	11	13	12	13	15	9	13	14	17	11	9	13	12	12
CSU	14	19	18	22	20	18	20	19	29	21	17	25	20	17	27	16	21	21	20	29	16	20	19	21	14
FDP	44	37	36	34	33	33	32	32	32	32	32	31	31	31	31	31	30	30	30	30	30	30	29	29	29
GREEN	14	16	23	15	14	15	17	16	14	15	14	14	14	19	11	16	14	8	17	16	12	16	12	13	17
LINKE	18	19	22	14	15	17	12	19	17	18	18	18	14	20	14	15	24	15	16	13	15	16	18	19	21
SPD	24	21	24	16	20	22	12	16	19	23	25	17	20	25	18	23	19	19	26	20	25	25	23	32	29
# Total Actors	304	843	177	394	696	176	170	272	2772	5028	2090	389	614	2005	168	123	185	137	542	957	234	266	1360	103	304
# Actors in FDP	97	225	45	105	180	39	45	63	598	1098	478	89	142	443	38	29	42	31	116	198	53	62	293	20	64

Party \ First name	DIETRICH	CARL	JULIAN	FELIX	HENNING	TILL	PHILIP	LORENZ	ARND	NICOLAI	THILO	LINUS	FABIAN	FYNN	LENNARD	JANNIK	LUKAS	NILS	MATHIS	CEDRIC	JANEK	SÖNKE	JONATHAN	TIMON	RUBEN
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	26	25	24	22	19	16	26	15	26	25	23	17	25	21	21	18	22	24	18	28	18	20	21	20	24
CDU	16	16	11	13	14	9	11	13	18	10	14	11	12	9	12	10	13	11	14	11	7	17	12	17	10
CSU	28	24	21	19	22	12	15	22	27	18	21	15	23	14	20	12	20	17	16	12	13	14	15	13	14
FDP	29	29	29	29	28	28	28	27	27	27	27	27	27	27	27	27	27	27	27	27	26	26	26	26	26
GREEN	19	11	15	16	17	17	14	21	12	13	16	24	15	14	17	14	16	15	23	11	21	22	19	18	12
LINKE	12	15	19	21	17	23	16	20	13	17	16	21	18	21	21	20	21	19	15	16	18	14	23	21	19
SPD	21	14	27	25	22	27	20	23	18	14	21	19	24	23	30	27	25	26	18	22	14	21	22	25	19
# Total Actors	202	479	3276	4205	854	605	1077	387	144	218	411	266	3643	139	163	644	3957	2138	103	250	102	167	876	160	212
# Actors in FDP	39	103	648	826	172	128	228	75	28	48	82	54	691	29	30	135	742	412	21	53	23	33	165	30	45

(e) GREEN

Party \ First name	RONJA	CLARA	HANNAH	FRIEDERIKE	AMELIE	NELE	JULE	MERLE	REBEKKA	PAULA	NORA	LINA	GRETA	LUCIA	KIRA	WIEBKE	LEONIE	SULTAN	HELEN	MIRI	JOHANNA	SOPHIE	TABEA	DEBORAH	LARA
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	17	13	11	9	11	14	19	11	21	22	19	21	18	20	14	13	14	3	21	11	14	17	17	23	19
CDU	5	6	6	8	5	5	5	6	6	6	6	5	4	9	6	10	8	5	8	4	9	7	6	6	6
CSU	5	4	6	13	5	7	7	2	14	11	9	7	11	13	6	10	9	5	12	15	19	8	7	7	7
FDP	5	9	9	14	9	8	7	9	9	9	7	8	11	11	7	12	12	3	7	11	10	10	9	6	12
GREEN	43	40	40	40	38	36	35	35	34	33	33	33	33	33	33	33	33	33	32	32	32	32	32	32	32
LINKE	19	19	17	14	18	22	23	20	19	13	17	19	16	16	16	14	16	26	14	15	18	20	11	15	19
SPD	22	26	30	28	24	24	25	28	25	23	27	28	25	17	32	33	30	51	21	25	25	27	27	23	29
# Total Actors	220	299	656	295	163	261	440	137	164	418	377	374	118	170	198	308	420	100	195	109	1152	815	179	111	679
# Actors in GREEN	83	102	220	93	56	81	130	43	43	119	106	103	33	47	57	81	113	26	55	31	287	213	52	31	175

Party \ First name	SOPHIA	LEA	ANNIKA	LENA	KARO	MAREIKE	IMKE	LOUISA	LAURA	HANNA	ESTHER	KLARA	LILLI	LEYLA	PAULINE	ELIF	CHARLOTTE	MAREN	MARLENE	PAULINA	REGINE	SUSANNA	MARA	ANNA	SVENJA
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	17	16	18	18	23	15	10	14	18	16	18	31	21	8	21	11	15	19	22	13	23	14	23	21	21
CDU	8	9	8	8	10	6	10	7	8	10	7	5	11	3	6	6	10	8	8	7	9	8	6	8	7
CSU	11	6	7	9	14	10	8	9	9	12	11	20	13	3	8	4	12	10	22	5	24	20	7	14	9
FDP	13	9	11	12	8	10	16	13	11	9	11	11	8	5	6	6	13	11	9	13	16	6	8	12	12
GREEN	32	31	31	31	31	31	30	30	30	30	30	29	29	29	29	29	29	29	29	29	28	28	28	28	28
LINKE	16	16	17	16	12	15	15	15	16	18	17	15	23	26	22	18	18	14	15	19	14	18	16	17	14
SPD	20	31	29	29	23	29	25	29	30	27	24	28	14	41	29	38	25	25	23	31	20	27	25	26	29
# Total Actors	502	909	806	1582	132	376	121	177	2097	578	324	190	157	111	238	115	508	458	368	130	232	129	232	3702	613
# Actors in GREEN	136	240	206	399	34	99	32	46	509	140	81	40	39	28	58	30	120	113	83	32	49	30	58	830	144

(f) LINKE

Party \ First name	CEMAL	AZAD	YILMAZ	FIRAT	UMUT	JONA	GEORGIOS	ÖZGÜR	JANNES	DIMA	SERHAT	HÜSEYIN	CEM	CHRISTOS	KEMAL	FERHAT	BARIS	ERKAN	CIHAN	SERGEJ	EREN	MERLIN	LU	BÜLENT	DORIAN
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	8	8	9	15	17	14	24	9	16	40	15	5	13	12	9	12	15	18	12	39	15	23	31	12	21
CDU	3	7	6	4	6	9	8	2	11	3	4	5	5	6	4	2	4	3	7	3	5	10	7	4	11
CSU	4	6	9	6	4	13	17	5	10	5	5	6	5	21	6	8	7	5	6	14	9	11	19	8	17
FDP	5	6	5	6	8	17	7	9	19	7	9	5	10	10	2	6	4	7	8	9	10	19	14	5	26
GREEN	24	18	25	28	15	22	18	20	10	16	12	25	13	7	18	15	8	11	15	7	7	23	22	19	16
LINKE	44	43	42	40	38	37	37	37	35	35	35	34	34	34	34	33	33	33	33	32	32	32	31	31	31
SPD	39	41	45	44	37	27	26	40	30	9	40	47	45	27	43	50	44	42	41	12	48	18	17	44	27
# Total Actors	100	144	138	103	111	128	135	120	105	105	114	333	262	104	154	144	115	149	130	312	110	163	180	140	121
# Actors in LINKE	35	48	41	29	34	34	36	36	28	32	33	88	71	30	45	38	33	41	35	86	28	38	40	35	25

Party \ First name	CAN	MO	DENIZ	SINAN	HASAN	NI	JACOB	ALI	EL	NIKOLA	MAXIM	KAT	ERCAN	NIKOLAI	JAY	SAN	MOE	ARON	KRIS	FILIZ	SALIH	MEHMET	DI	ORHAN	ARI
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	17	23	15	13	10	34	17	12	37	18	30	27	13	19	33	34	21	29	29	8	15	11	36	13	24
CDU	8	6	6	8	8	9	9	12	8	9	10	8	1	8	8	9	5	7	12	7	8	5	9	6	6
CSU	8	16	9	7	8	22	12	8	18	14	7	16	8	23	15	16	20	10	22	8	3	7	24	6	19
FDP	13	16	9	4	7	15	15	6	8	9	19	10	8	19	13	12	20	13	17	5	7	8	12	7	9
GREEN	15	19	15	16	22	20	22	19	13	12	11	20	11	9	14	18	17	15	16	26	16	24	11	22	19
LINKE	31	30	30	30	30	30	30	30	29	29	29	29	29	29	29	29	28	28	28	28	28	28	28	28	28
SPD	43	23	41	46	45	32	25	43	21	23	19	22	50	19	17	21	23	22	22	47	40	47	30	46	26
# Total Actors	350	425	560	225	494	206	263	1740	553	114	169	187	120	187	248	106	123	102	189	110	101	755	207	139	104
# Actors in LINKE	80	98	136	55	115	38	60	402	119	29	39	41	29	43	55	22	26	23	36	24	24	161	38	30	22

(g) SPD

Party \ First name	TC	FATIH	RAMAZAN	MUHAMMED	GÖKHAN	MERT	METIN	SALIM	SERKAN	AYHAN	AHMET	HATICE	OSMAN	KENAN	SULTAN	YASIN	MUHAMMAD	BURAK	FERHAT	AYSE	FATMA	MOHAMMAD	SAID	UGUR	ERCAN
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AfD	8	16	13	14	14	19	6	17	14	14	11	4	17	9	3	22	10	19	12	6	7	9	7	11	13
CDU	4	4	5	8	7	3	6	9	4	2	7	4	7	10	5	7	31	5	2	2	6	24	21	6	1
CSU	5	10	11	8	8	4	6	10	5	14	10	6	7	6	5	10	7	7	8	8	3	7	5	8	8
FDP	5	11	4	6	8	7	8	6	7	8	6	6	5	6	3	7	3	7	6	4	5	4	2	9	8
GREEN	10	16	24	12	11	15	21	15	14	12	15	26	14	13	33	15	15	10	15	19	26	17	16	16	11
LINKE	15	14	17	10	22	18	27	16	22	17	25	18	21	24	26	17	14	19	33	25	19	12	20	25	29
SPD	63	58	57	56	56	55	55	54	54	53	52	52	52	51	51	51	51	51	50	50	50	50	50	50	50
# Total Actors	583	262	132	167	150	143	205	103	198	123	434	120	220	135	100	218	189	202	144	123	175	569	137	125	120
# Actors in SPD	336	118	58	82	67	65	87	44	89	55	180	54	93	58	41	86	73	86	57	54	75	230	56	50	50

Party \ First name	FARID	ISMAIL	HALIL	MURAT	ABDUL	MALIK	ENES	MOHAMAD	KADIR	CENGIZ	REZA	ÖMER	TOLGA	EREN	ABDULLAH	HÜSEYIN	HAKAN	YUSUF	AHMED	MUSTAFA	MOHAMED	KAAN	MEHMET	FILIZ	MESUT
Rank	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
AfD	6	8	11	10	8	15	22	11	15	14	11	12	17	15	10	5	15	10	14	11	14	21	11	8	8
CDU	14	9	6	5	17	15	7	30	5	2	19	7	6	5	14	5	7	7	27	11	20	3	5	7	3
CSU	10	8	9	9	9	8	5	8	9	11	8	11	6	9	7	6	11	7	10	9	8	5	7	8	7
FDP	6	4	7	5	2	10	6	0	11	9	4	7	6	10	5	5	9	8	4	6	6	10	8	5	6
GREEN	18	18	16	20	18	12	14	14	17	13	16	17	14	7	15	25	20	18	14	19	10	13	24	26	24
LINKE	25	25	18	26	19	16	11	17	21	23	19	18	23	32	22	34	25	26	13	25	18	23	28	28	26
SPD	49	49	49	49	49	48	48	48	48	48	48	48	48	48	48	47	47	47	47	47	47	47	47	47	46
# Total Actors	107	305	164	492	209	110	134	270	164	120	139	216	134	110	268	333	337	286	472	728	503	112	755	110	120
# Actors in SPD	41	123	69	191	84	43	57	102	63	48	54	86	53	42	105	123	119	109	172	269	194	43	271	40	46

mathematics (set theory), and analytics (social set analysis) but also to demonstrate its technical viability, suitability and utility by designing, developing and evaluating an IT-artifact, the Social Set Visualizer (SoSeVi). In other words, we postulated and - hopefully - illustrated that Set Theory in general is better suited from a mathematical standpoint to model human social associations than network theory or graph theory. Beyond the immediate social network and particularly on large scale social media platforms such as Facebook, Twitter and Tencent QQ, we believe, and hope, that this fundamental change in the foundational mathematical logic of the formal model from graphs to sets will allow for new insights.

C. Limitations

One of this paper's limitations is that we do not present domain-specific empirical findings in terms of political sciences and social media management. That said, such domain-specific empirical findings of the set theoretical approach can be found in [28], [29]. A second limitation is the lack of exposition of the full range of set theoretical approaches beyond the classical "crisp sets" discussed in the paper (for example: fuzzy sets, rough sets, random sets, Bayesian sets). A third and final limitation is the limited space devoted to the technical aspects of the IT-artifact. Also, the data set is only for 2017 and does not contain previous years of political discourse on Facebook.

D. Future Research

Current and planned future work in our Center for Business Data Analytics is addressing some of the theoretical limitations identified above in terms of developing formal models and analytical methods for fuzzy, rough and random sets. Furthermore, more advanced modeling of political social media discourse needs to be performed through machine learning. Our focus is on data visualization, and merging these capabilities with innovative methods of extracting meaningful insights from the social media data at hand. We suggest future work on the 2017 German federal election also takes into account not only the party Facebook pages, but also the Facebook pages of each individual member of parliament. This would enable analysis of further grass-roots political activity and discourse.

REFERENCES

[1] W. S. Cleveland, "Data science: an action plan for expanding the technical areas of the field of statistics," *International Statistical Review*, vol. 69, no. 1, pp. 21–26, 2001. [Online]. Available: http://dx.doi.org/10.1111/j.1751-5823.2001.tb00477.x 1

[2] M. Loukides, *What Is Data Science?* O'Reilly Media, 2012. 1

[3] N. Ohsumi, "From data analysis to data science," in *Data Analysis, Classification, and Related Methods*. Springer Berlin Heidelberg, 2000, pp. 329–334. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-59789-3_52 1

[4] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabsi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne, "Computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009. 1

[5] R. Vatrapu, "Understanding social business." in *Emerging Dimensions of Technology Management*. Springer, 2013, pp. 147–158. 1

[6] J. Sterne, *Social media metrics: How to measure and optimize your marketing investment*. John Wiley & Sons, 2010. 1

[7] M. Sponder, *Social media analytics: effective tools for building, intrepreting, and using metrics*. McGraw-Hill, 2012. 1

[8] Z. Tufekci, "Big questions for social media big data: Representativeness, validity and other methodological pitfalls," *arXiv preprint arXiv:1403.7400*, 2014. 1

[9] J. L. Gross and J. Yellen, *Graph theory and its applications*. CRC press, 2005. 1

[10] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323(5916), pp. 892–895, 2009. 1

[11] M. Emirbayer, "Manifesto for a relational sociology," *The American Journal of Sociology*, vol. 103(2), pp. 281–317, 1997. 1

[12] R. R. Mukkamala, A. Hussain, and R. Vatrapu, "Towards a formal model of social data," IT University of Copenhagen, Denmark, IT University Technical Report Series TR-2013-169, November 2013. 1

[13] ——, "Towards a set theoretical approach to big data analytics," in *3rd International Congress on Big Data (IEEE BigData 2014)*, June 2014. 1

[14] M. S. Mizruchi, "Social network analysis: Recent achievements and current controversies," *Acta sociologica*, vol. 37, no. 4, pp. 329–343, 1994. 1

[15] R. Vatrapu, R. R. Mukkamala, and A. Hussain, "A set theoretical approach to big social data analytics: Concepts, methods, tools, and findings," in *ECCS Satellite Workshop 2014*, 2014, pp. 22–24. 1, 2

[16] B. Flesch, R. Vatrapu, R. R. Mukkamala, and A. Hussain, "Social set visualizer: A set theoretical approach to big social data analytics of real-world events," in *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2418–2427. 1, 2

[17] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister, "Upset: visualization of intersecting sets," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 12, pp. 1983–1992, 2014. 2, 3

[18] M. Berry, I. Garcia-Blanco, and K. Moore, "Press coverage of the refugee and migrant crisis in the eu: a content analysis of five european countries," 2016. 2

[19] S. A. Testa, "Financial (in) stability, banking crisis and policy implications: An empirical analysis on the eu countries," B.S. thesis, Università Ca'Foscari Venezia, 2017. 2

[20] J. Fritsch, "Frauen und führung in deutschland: Analyse der chancen und risiken der gesetzlichen frauenquote," 2016. 2

[21] A. Hussain and R. Vatrapu, "Social data analytics tool (sodato)," in *DESRIST 2014*, ser. Lecture Notes in Computer Science (LNCS). Springer, vol. 8463, 2014, pp. 368–372. 3

[22] ——, "Social data analytics tool: Design, development, and demonstrative case studies," in *Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International*, Sept 2014, pp. 414–417. 3

[23] A. Hussain, R. Vatrapu, D. Hardt, and Z. Jaffari, "Social data analytics tool: A demonstrative case study of methodology and software." in *Analysing Social Media Data and Web Networks*. Palgrave Macmillan, 2014. 3

[24] H. G. Miller and P. Mork, "From data to decisions: a value chain for big data," *IT Professional*, vol. 15, no. 1, pp. 57–59, 2013. 3

[25] Statista, "Mitgliederzahlen der politischen parteien in deutschland am 31. dezember 2016," Webpage, 2016, https://de.statista.com/statistik/daten/studie/1339/umfrage/mitgliederzahlen-der-politischen-parteien-deutschlands/. 5

[26] B. fr Politische Bildung, "Die soziale zusammensetzung der parteimitgliederschaften," Webpage, 2016, https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/zahlen-und-fakten/140358/soziale-zusammensetzung. 6

[27] J. Mueller and G. Stumme, "Gender inference using statistical name characteristics in twitter," *arXiv preprint arXiv:1606.05467*, 2016. 6

[28] R. R. Mukkamala, J. I. Srensen, A. Hussain, and R. Vatrapu, "Detecting corporate social media crises on facebook using social set analysis," in *Proceedings of IEEE Bigdata Congress*, 2015. 10

[29] ——, "Social set analysis of corporate social media crises on facebook," in *Proceedings of IEEE 19th International Enterprise Distributed Object Computing Conference (EDOC)*, 2015. 10