

DEbateNet-mig15: Tracing the 2015 Immigration Debate in Germany Over Time

**Gabriella Lapesa¹, André Blessing¹, Nico Blokker², Erenay Dayanik¹,
Sebastian Haunss², Jonas Kuhn¹, Sebastian Padó¹**

Institute for Natural Language Processing¹, University of Stuttgart (Germany)

Research Center for Inequality and Social Policy², University of Bremen (Germany)

name.surname@ims.uni-stuttgart.de, name.surname@uni-bremen.de

Abstract

DEbateNet-mig15 is a manually annotated dataset for German which covers the public debate on immigration in 2015. The building block of our annotation is the political science notion of a claim, i.e., a statement made by a political actor (a politician, a party, or a group of citizens) that a specific action should be taken (e.g., vacant flats should be assigned to refugees). We identify claims in newspaper articles, assign them to actors and fine-grained categories and annotate their polarity and date. The aim of this paper is two-fold: first, we release the full *DEbateNet-mig15* corpus and document it by means of a quantitative and qualitative analysis; second, we demonstrate its application in a discourse network analysis framework, which enables us to capture the temporal dynamics of the political debate.

Keywords: computational social science, discourse network analysis, political science, annotation, interdisciplinarity

1. Introduction

The last decade has witnessed a dramatic increase in the number of refugees attempting to enter Europe from Africa and the Middle East. This rise in numbers has had a huge impact on society, thereby strongly resonating in the public debate. Is it right to open the borders to immigrants, or should migration be more strictly regulated? To what extent should host countries be responsible for accommodating refugees? What are the implications for internal security and economy? This is just a small sample of the questions which characterize the public debate on immigration, both at the level of single individuals (the citizens) and public institutions (politicians).

From a political science perspective, a crucial aspect of the discourse concerning societally relevant topics is represented by its dynamics. For example, one highly influential political actor may change her/his mind about a specific issue within the domain of interest (e.g., should we establish a quota for refugees?), and initiate a change in the opinion of other discourse participants who, in turn, may converge on the same position or take a strong stand against it. Such dynamics lead to the emergence of new political coalitions and to the disappearance of others. Empirically, political discourse is represented in terms of a network of actors and their statements concerning relevant aspects of the discourse domain (e.g., Angela Merkel, actor, declares that the borders should stay open for refugees; a far-right political party, actor, states exactly the opposite). This abstract representation allows for an empirical investigation of the unfolding of the debate over time, which can be captured in terms of the quantitative properties of the network (Haunss and Kohlmorgen, 2009; Leifeld and Haunss, 2012; Haunss et al., 2013; Leifeld, 2016). This is precisely the target of our work: investigating the dynamic of the debate on immigration with the support of discourse network analysis.

The first step towards a discourse network analysis of a political debate is the annotation of relevant texts (in our case, newspaper articles). The dataset described in this paper, *DEbateNet-mig15*, provides a fine-grained picture

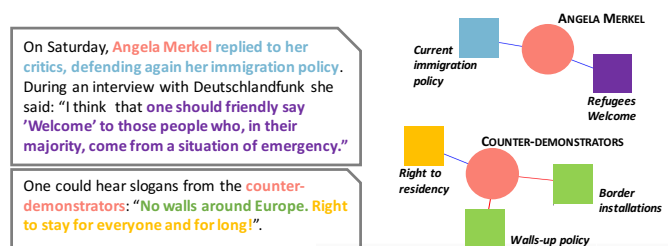


Figure 1: Annotation and corresponding network representation for the immigration debate, October 3rd 2015

of the public discourse concerning the domestic debate on immigration in Germany in 2015, the crucial year for the "refugee crisis" in Europe.

Figure 1 exemplifies our annotation and the discourse network that can be created from it. It is a real example from *DEbateNet-mig15* and it is based on the annotated documents for October 3rd, 2015. From this national holiday, and thus a slow news day, only two immigration-related claims are reported in the newspaper published the following day. In one article, Angela Merkel is reported to have replied to those who criticised her immigration policy, and a direct quotation from her speech is reported, stating the need for a welcoming attitude towards refugees. The second set of claims are attributed to a group of counter-demonstrators, who showed up during an official ceremony in Saxony (a state with a conservative government): claiming the right of residency for refugees, the demonstrators also made two claims against the isolation of Europe and the construction of border installations as a solution to the immigration problem. The claims are highlighted in colors in the text, and give rise to the corresponding parts of the network representation to the right. The actors are represented by red squares in the discourse network. Blue edges indicate support towards a claim category (Merkel supports the "Refugees Welcome" claim), red edges indicate opposition to it (the demonstrators stand against the claim "Controlling migration with border

installations").

Methodologically, this paper follows our previous work (Padó et al., 2019; Blessing et al., 2019). The novel contribution of this paper are: (a) the release and the documentation of the complete dataset, including a quantitative and qualitative analysis with corpus linguistic tools such as keywords and collocations (Baker et al., 2008); and (b) a quantitative/qualitative analysis of the concrete discourse network structures that arise from the annotation as well as the temporal dynamics of these structures.

In Section 2, we provide background on the discourse network analysis framework from political science that we build on, and on the role of and challenges for NLP in such a study. In Section 3, we describe our corpus and annotation guidelines and provide an empirical characterization of our corpus at the level of annotation¹. Section 4 proceeds to investigate and analyze the discourse networks that arise from the annotation. Section 5 concludes by picking up the methodological considerations raised in the paper and discusses potential research developments and ongoing work.

2. Background: Claims Analysis, Discourse Networks, and the Role of NLP

Understanding the structure and evolution of political debates is essential for understanding democratic decision making, and is therefore of central interest to political science (de Wilde, 2011; Zürn, 2014; Haunss and Hofmann, 2015).

Democratic decision making can broadly follow two logics, one of which is a “technocratic” mode, where decisions are taken by administrative staff and field-specific experts. We focus on the second type of decision making, the “politicized” mode, which proceeds through programmatic statements (Schmidt and Radaelli, 2004) and political debates. While there is no general theory about mechanisms driving political discourse, there seems to be at least general agreement that the formation and evolution of discourse coalitions is a core mechanism (Hajer, 1993; Sabatier and Weible, 2007) and that change in these coalitions is influenced by external events and by the discourse itself (Leifeld, 2016).

One promising way to gain insight into such discourse dynamics in an empirically robust fashion, based on widely available newspaper corpora, combines political claims analysis (Koopmans and Statham, 1999) with *discourse network analysis* (Leifeld and Haunss, 2012). The unit of analysis is the *claim*, that is, a demand, proposal, or criticism that is *supported* or *rejected* by an *actor* (a person or a group of persons) and can be *categorized* with regard to its contribution to the debate at hand. Crucially, not all statements concerning the topic are to be considered a claim, but only those which target a specific action to be taken (e.g., giving empty flat to refugees). Claims and the actors who make them are represented together in a bipartite *affiliation network*. A discourse coalition is then the projection of the

¹The language of our resource is German. In this paper, we provide examples in English (with or without the German counterpart, depending on space constraints).

affiliation network on the actor side, while the projection on the concept side yields the argumentative clusters present in the debate.

Clearly, manual annotation of such claims and claim-actor relations is a resource intensive process. It is therefore natural to ask if Natural Language Processing can help: What are the potentials, limitations, and the practical issues of applying NLP to the automatic construction of discourse networks?

At a general level, the NLP take on debate modeling can build on the insights from argumentation mining and subjectivity analysis (Peldszus and Stede, 2013; Ceron et al., 2014; Swanson et al., 2015; Stab and Gurevych, 2017; Vilares and He, 2017). An ideal NLP tool would automatically identify the actors and their contributions to the debate, and analyze such contributions at a structural level (identifying argumentative structure in their statements), at a semantic level (classifying statements into relevant categories), and at a pragmatic level (detecting the polarity of the statements). In our concrete experience (Padó et al., 2019), however, this task cannot be completely automated, at least not if the target is to acquire representations at the level of granularity and at the level of quality which are required for the political science analysis. What proved successful is instead the integration of manual annotation and NLP methods (Blessing et al., 2019) into a semi-automatic procedure that speeds up the manual work by providing intelligent proposals, efficient annotation interfaces, and adding automatic “pre-annotation” that is clearly labeled as such. We find that this approach scales up with manageable loss in fine-grainedness and quality and can serve as an example of successful “mixed methods” that are becoming more prominent at the intersection where big data meets humanities and social sciences (Kuhn, 2019).

3. DEbateNet-mig15

In this section we introduce *DEbateNet-mig15* and characterize it both quantitatively and qualitatively. The dataset is available in a .json format as a CLARIN resource, at the PID <http://hdl.handle.net/11022/1007-0000-0007-DB07-B>, along with a bundle of R utilities employed for the analysis presented the following sections.

3.1. Source Corpus

Our source corpus is *Die Tageszeitung (taz)*, a major national German quality newspaper. It is perceived as the most left-oriented major German newspaper, but can still be assumed to portray both sides of the relevant policy issues. We decided to focus on newspaper texts for a number of reasons, both practical and theoretical: (a) to build our discourse networks we need identifiable and trackable actors (e.g., politicians, political parties, specific groups of protesters) – this would not be straightforward, for example, in social media texts; (b) our fine-grained, multi-level annotation is better carried out relying on large textual spans like the ones provided by newspaper articles, which obviously offer a better support for the interpretation of the annotators; (c) quality newspapers represent an elite discourse which plays an important role in political decision making (Schneider

Corpus	Articles	Tokens	Keywords
Taz 2015	38,414	16,955,430	Flüchtling (<i>refugee</i>), Syrien (<i>Syria</i>), syrisch (<i>syrian</i>), Lageso (<i>Agency for Health and Social Affairs</i>), Flüchtlingsunterkunft (<i>refugee accommodation</i>), Unterkunft (<i>accommodation</i>), Syriza, Syrer (<i>Syrian</i>), Mensch (<i>man</i>), Paris
DEbateNet-mig15	959	556,185	Flüchtling (<i>refugee</i>), Ungarn (<i>Hungary</i>), Grenze (<i>boundary</i>), Deutschland (<i>Germany</i>), Asylbewerber (<i>asylum seeker</i>), Merkel, Asyl (<i>asylum</i>), Land (<i>country</i>), EU, Migranten (<i>immigrants</i>)
DEbateNet-mig15 (claim)	430	239,351	Flüchtling (<i>refugee</i>), Deutschland (<i>Germany</i>), Asylbewerber (<i>asylum seeker</i>), Merkel, Asyl (<i>asylum</i>), Grenze (<i>boundary</i>), Ungarn (<i>Hungary</i>), Maizière (interior minister, CDU), Land (<i>country</i>), Herkunftsstaat (<i>country of origin</i>).

Table 1: Sampling *DEbateNet-mig15* from *Die Tageszeitung (taz)*: keywords

et al., 2007); and (d) newspaper articles are more likely to represent both pro and contra positions. Note also that our approach builds on the assumption that the political discourse in the public sphere is not only a mere snapshot of the real world debate, but it stands in a causal relation to it.

3.2. Article Selection

As it was infeasible for the annotators to read all the *Tageszeitung* articles from 2015, the first step was to select only articles related to our topic. For more details on this step, refer to Blessing et al. (2019). The subcorpora involved in the subsequent steps of article selection are listed in Table 1, along with their size in number of documents (*Docs*) and tokens (*Tokens*). *Taz 2015* is the starting point of our sampling procedure; *DebateNet-migr15* is the whole released dataset, and *DebateNet-migr15 (claim)* is the subcorpus of *DebateNet-migr15* which contains the articles in which the annotators have found at least one claim. The amount of articles containing claims corresponds to slightly less than a half of the entire dataset. Recall that we are only interested in claims which relate to the topic of immigration in *Germany*: many articles could very well be about immigration, but target a non-domestic aspect of it (e.g., actors commenting on which policies other countries should or should not adopt). Such articles, however, are still a valuable part of *DebateNet-migr15*: given that they have been inspected by the annotators, the absence of annotation can be exploited as a source of negative examples for the training of claim identifier models (Padó et al., 2019).²

To give our readers a fingerprint of our three corpora, we have conducted a keyword analysis by employing standard methods (log-likelihood with a frequency threshold of 3 in both focus and reference corpus) and varying the reference corpora.³ The table reports the top 10 keywords for each corpus. For each row, focus corpus in the keyword extraction is the specified corpus, while reference corpus is different from case to case.

For *Taz 2015*, the reference corpus is *Taz 2010* (43,827 articles, 17,722,489 tokens): the inspection of the keyword list highlights the central role played by the debate on immigration in 2015, as the majority of the keywords is in a direct

or indirect relation to the immigration discourse.

For the *DebateNet-migr15* and *DebateNet-migr15 (claim)* subcorpora, the reference corpus is *Taz 2015*. Unsurprisingly, the keywords of *DebateNet-migr15* are more specific to the immigration discourse: proper names (actors, i.e., Merkel, EU, Ungarn) as well as common nouns such as *Asylbewerber*, *Asyl* (asylum seeker and asylum) *Grenze* (border) and *Land* (country), which correspond to central (and controversial) concepts within the political discourse. The "keyness" of *Migranten* alongside with *Flüchtling* is a clear cue to debate concerning the distinction between (economic) migrants and refugees, which has been a hot-topic at least in some stages of the so called "refugee crisis" (see Section 4. for more details on this aspect).

The keywords of *DebateNet-migr15 (claim)* are largely overlapping with those of the whole dataset, albeit with a slightly different ranking. Interestingly, however, two new keywords show an increase in specificity: (*Thomas de*) *Maizière* (interior minister) is a crucial actor in the political scene (and in the immigration debate), while *Herkunftsstaat* is a technical term which refers to the notion of "country of origin" of the refugees.

3.3. Annotation

The annotation has been carried out with the support of the *MARDY* annotation environment, developed for this project and described in detail in (Blessing et al., 2019).

Our manual annotation targets the following levels:

1. **Claim identification:** identification of the textual spans containing claims. Recall from our examples in figure 1 that claim-bearing textual spans do not necessarily coincide with a sentence: they can be a subpart of a sentence, or span beyond sentence boundary.
2. **Claim classification:** assignment of theoretically-motivated claim categories to the textual spans. Note that a textual span can be assigned more than one claim category.
3. **Actor identification & claim attribution:** identification of the strings corresponding to actor mentions (e.g., "Angela Merkel", "Die Kanzlerin", "Frau Merkel") and linking of the previously identified textual spans to the relevant actor. Note that a single claim can be attributed to more than one actor, and actors can be mentioned inside or outside the textual span.

²A previous version of the corpus, with 764 marked spans, was published as part of Padó et al. (2019).

³For an in-depth discussion of methodological issues of keyword analysis, refer to Baker (2004).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Type
Span length	1.000	11.50	16.00	17.97	22.00	63.00	n.a
No. of claims per span	1.000	1.000	1.000	1.213	1.000	6.000	n.a
No. of actors per span	1.000	1.000	1.000	1.01	1.000	4.000	n.a
Fine-grained claims	1.00	6.00	14.00	22.7	25.00	170.00	97
Claim bigrams	1.000	1.000	1.000	2.038	2.000	25.000	208

Table 2: Span- and claim-level statistics

Code	Claim label	f	%
1xx	Steuerung von Migration (Controlling Migration)	493	20.7
2xx	Aufenthalt (Residency)	412	17.3
3xx	Integration (Integration)	195	8.2
4xx	Innere Sicherheit (Domestic Security)	86	3.6
5xx	Aussenpolitik (Foreign Policy)	433	18.2
6xx	Ökonomie, Arbeitsmarkt (Economy, Labor Market)	96	4.0
7xx	Gesellschaft (Society)	266	11.2
8xx	Verfahren (Procedures)	393	16.5

Table 3: High level categories: Code, Label, absolute frequency and percentage

104: Walls-up policy

- **Explanation:** This claim refers to demands that deal with (European) isolation from the refugee problem. These include, for example, unspecified demands for isolation or a stop to immigration.
- **Example:** "Europe cannot [reading: must not] cordon itself off."

207: Deportation

- **Explanation:** This claim refers to demands for the deportation of refugees. This claim often occurs together with a request for an expedited procedure (Fast/accelerated procedure), but is also suitable for requests for a general increase in deportations.
- **Example:** "In addition, it should be possible to deport rejected asylum seekers even if they are not life-threateningly ill."

504: Safe legal status for country of origin

- **Explanation:** This claim refers to demands for the extension of the legal status of one or more countries of origin of refugees as "safe".
- **Example:** "In addition, the legislator must declare more Balkan countries safe third countries 'to which we can then deport more quickly'."

Table 4: Fine-grained categories: codebook examples and annotation guidelines

4. **Date assignment:** the claim is assigned a date, which is by default the day preceding the publication of the article. It is the annotator's task to reconstruct the claim date, based on textual information.
5. **Polarity:** does the actor support or reject the categorized claim?

Before moving on to the quantitative properties of the dataset, let us elaborate upon our classification schema (*codebook*, in political science terminology). The annotation codebook⁴ specifies 8 higher-level categories, further subdivided into 97 fine-grained categories. In the codebook, each claim category is associated to annotation guidelines (which usually also make explicit reference to the relations between annotation categories) and one or more textual spans which are considered representative of the targeted category.

Table 3 lists the 8 high-level categories (along with their frequencies, to which we will come back in the next section) and Table 4 lists three hand-picked instances of finer-grained categories, along with the corresponding annotation guidelines and examples.

Throughout the annotation process (which took roughly a year) the codebook has evolved, displaying the "hermeneutic cycle" which is typical of Digital Humanities projects. Once the covered sample enlarged, new categories have been added, or existing categories have been redesigned. In this process, a crucial role has been played also by the interaction between the experts and the trained annotators, as a claim category which is problematic for all annotators is likely to be ill-defined, or even ill-motivated.

3.4. Analysis

DebateNet-migr15 contains 1815 observations (textual spans), corresponding to 2274 distinct claims (recall that one textual span may contain more than one claim).

Observation-level statistics. Table 2 reports descriptive statistics aggregated at the level of textual spans, namely length in words (*Span length*), *No. of claims per span* and *No. of actors per span*.

Correlation analysis (Spearman ρ , ρ) revealed extremely weak positive correlations between *Span length* and *No. of claims per span* ($\rho = .13$) and between *Span length* and *No. of actors per span* ($\rho = .07$), showing that there is no influence of the amount of conceptual content being encoded and the amount of textual material, at least as far as the reporting style in the newspaper articles is concerned. Moreover, we found no correlation between *No. of claims per span* and *No. of actors per span*.

During the annotation, we also marked whether the actor would be mentioned inside or outside the textual span identified as a claim. Getting back to the examples in Figure 1, the two claims by Merkel are, respectively, and example of internal actor mention ("Angela Merkel replied to her critics, defending her immigration policy") and of an external actor mention in reported speech ("I think that..."). This is an

⁴https://github.com/mardy-spp/mardy_acl2019/blob/master/codebook.pdf

Actor	Freq	Class	Role
Thomas de Maizière	90	per	Interior minister (CDU)
Angela Merkel	87	per	German Chancellor (CDU)
CSU	56	org	Conservative party (in government)
Bundesregierung	49	org	German cabinet (Chancellor and Ministers)
SPD	41	org	Left-center party (in government)
Horst Seehofer	39	per	President of Bavaria (CSU)
Jean-Claude Juncker	26	per	President of EU commission
CDU	23	org	Conservative party (in government)
EU-Kommission	21	org	EU commission
Deutschland	21	org	Germany
Sigmar Gabriel	21	per	Vice-chancellor (SPD)
EU	18	org	European Union

Table 5: Most frequent actors

important contrast, and one that can support the interpretation of the performance of automatic classification methods trained on the dataset. An actor mention is a very powerful cue to identify a claim, but probably not the most reliable one (for example because the mentioned actors change over time). In our dataset, the majority (66%) of textual spans includes the mention of the actor.

Actor-level statistics. The split between person actors (e.g., "Angela Merkel", "the counter-demonstrators") and organization actors (e.g., "CDU", "EU", "Hungary") is in favor of the first ones, (62% PER vs. 38% ORG). Table 5 reports the actors with the highest frequency (number of claims), along with their role (institutional and/or politic).⁵ We observe a bias towards government and EU actors who were crucial interlocutors in the domestic debate on migration.

Claim-level statistics. We now look into the claim annotation, starting from the high-level categories. The bottom panel of Table 2 illustrates the breakdown of high-level claim categories in our dataset. We report the category codes, labels (in German and English), the absolute frequency and the percentage over the total number of claims in the dataset. We notice a predominance of the more "concrete" categories, related to "Controlling Migration" (concrete actions to be made to regulate the amount of incoming refugees), followed by "Foreign Policy" (how to interact with other countries, who are in control (or lack of it) of the immigration flows), "Residency" (how to deal with the refugees that are already in Germany), and "Procedures" (concrete actions targeting, for example, protection of minors and women, but also rules for deportation of immigrants). A weaker role is played by "Society" and "Immigration", which tend to be more "ab-

⁵ Actor mentions are mapped to canonical names by applying the method described in Blessing et al. (2019); for example, we employ knowledge bases to detect that mentions of "Kanzlerin" are, in fact, to be mapped to "Angela Merkel".

Code	Freq	Glob	Claim Category
POSITIVE			
501	152	193	EU solution (quotas for refugees)
812	97	103	Fast/Accelerated Procedure
504	93	124	Safe country of origin
805	78	82	Additional Financing
207	70	82	Deportation
102	69	80	Ceiling/Upper Limit
105	59	73	Border Controls
309	55	76	Care (medical, financial)
705	54	60	Refugees Welcome
108	46	59	Immigration law
NEGATIVE			
703	45	54	Xenophobia
501	41	193	EU solution (quotas for refugees)
190	36	51	Current Migration Policy
504	31	124	Safe Country of Origin
709	24	25	Right-wing radicalism
203	24	64	Centralized Accommodation
104	21	52	Walls-Up Policy
309	21	76	Care (medical, financial)
110	17	41	Asylum Right
202	17	34	Refugee Accommodation

Table 6: Top 10 categories by polarity

stract" in that they deal, for example, with human rights). Weak is also the role of and "Economy/Labor market" and "Domestic security", which are far from abstract but apparently less likely to surface when the debate is in a "crisis" mode and targets concrete solutions for more immediate problems.

At the level of the fine-grained actual annotation, the frequency distribution is shown in the bottom panel of Table 2 (row *Fine-grained claims*). The frequency distribution of fine-grained categories shows very clearly that our annotation, albeit accurate, could not possibly provide the basis for reliable automatic classification (too many classes, too few items per class, overall).

Table 6 reports the 10 most frequent claim categories per each polarity (positive vs. negative), along with the overall frequency of the category (*Glob*, the sum of positive and negative instances for a category). Given that the positive claims are in the vast majority, it would have made no sense to report the global ranking (irrespective of polarity), because it is almost identical to the one of the positive claims. A comparison of the two lists provides input for a number of observations, whose common denominator is the clear separation between discourse coalitions on the left and right sides of the political spectrum. First, the claims "EU solution" and "Safe country of origin" dominate both rankings (first and third in the positive ranking, second and fourth in the negative ranking), albeit with a largest share for the positive polarity; this indicates that these two claims could be strong indicator for the identification of coalitions. Second, the societal claims appearing in both lists show a clear left-wing nuance: almost complete negative predominance towards "Xenophobia" and "Right-wing extremism", almost complete positive predominance towards "Refugees Welcome". Third, still on the left vs. right divide, we observe the positive predominance towards "Deportation", "Ceil-

Bigram	f>1	s-ll	Categories
207-812	25	57.5	Deportation & Fast/Accelerated Procedure
109-508	6	17.8	Contrasting smuggling activities & Military Intervention
504-812	9	14.5	Safe country of origin & Fast/Accelerated Procedure
706-714	3	10.7	Fundamental Rights & Leading culture
110-801	2	10.2	Asylum Right & Legal Principles
190-705	5	9.66	Current Migration Policy & Refugees Welcome
209-212	4	9.37	Restricted Residency Obligation & In-kind Contributions
701-703	8	9.3	Populism & Xenophobia
303-706	3	9.3	Forced Integration & Fundamental Rights
703-709	5	8.9	Xenophobia & Right-wing Radicalism

Table 7: Claim pairs, ranked by simple-log likelihood (s-ll)

ing/Upper Limit", as well as "Border Control", which can be straightforwardly be classified as right-dominated. A trickier case, on the second position of the positive ranking, is "Fast/accelerated procedure": positive predominance towards this claim is "ambiguous" on the left vs. right scale, as quicker procedures could be those acknowledging the refugee status (and thus the right to stay in the country), but also lead to quicker deportations. Overall, the fact that, albeit the left-wing coloring of our source newspaper, we still find a balance of left and right-wing claim/polarity combinations confirms (a) the fact that our annotation schema is indeed capable to produce the multi-faceted representation of the policy debate we are interested in and (b) that *taz* is indeed a balanced source, and can thus be employed for future annotation projects.

We now turn to a collocation analysis to quantify the association strength between claim categories. In the bottom panel of Table 2 we observe the frequency distribution of claim bigrams (i.e., pairs of claims co-occurring together in the same textual span): in total, we observe 208 different claim combinations, with a mean frequency of 2 each. Table 7 reports the most strongly associated claim pairs, ranked by simple-log likelihood (Evert, 2008).⁶ In some cases, the relation between the two components is of an episodic kind (e.g. Contrasting smuggling activities and military intervention), and it also allows to "disambiguate" some claim categories on the left- vs. right: the strongly association between "Fast/Accelerated procedure" on the one hand, and "Deportation" and "Safe country of origin" on the other confers to this claim a pretty strong right-wing nuance. In other cases, we observe strong conceptual ties ("Populism and Xenophobia", "Xenophobia and Right-wing radicalism") which indicate the possibility of merging categories (thus

⁶To collect frequencies of claim pairs, we considered all the cases in which a given span was annotated for multiple claims. If more than two claims were present, we stored all the possible claim combinations. This procedure resulted in a list of bigram pairs along with their co-occurrence frequencies; from the same list, we also calculated marginal frequencies and sample size.

Month	Obs.	C_token	C_type	Actors	degree
Jan	141	189	47	77	2.44
Feb	66	83	30	57	1.77
Mar	41	60	27	31	1.72
Apr	79	89	37	49	1.84
May	78	106	32	41	2.41
Jun	89	100	41	56	1.86
Jul	140	177	46	90	2.13
Aug	207	264	52	109	2.73
Sep	411	541	66	168	3.38
Oct	211	285	55	116	2.69
Nov	186	267	58	82	2.91
Dec	166	213	54	78	2.39

Table 8: DEbateNet over time

reducing the sparsity of the network).

4. Discourse Networks in *DEbateNet-mig15*

The fine-grained, multi-level annotation contained in *DEbateNet-mig15* lends itself well to a discourse network analysis of the dynamics of the underlying debate. Our annotation allows for aggregation at several levels, the most straightforward obviously being the level of actors, claims, and, crucially, time.

Actor-level aggregation We first discuss aggregation over time at the actor level, using Angela Merkel as an example actor. Figure 2 displays her claims, aggregated over 3 time-spans of 4 months each. The most obvious observation is the increase in complexity: this topic was not too much on the agenda in early 2015, but it became a critical part of policy over the year. We observe a balance between pro/con in spring, and again in winter. In summer, we only observe pro claims, arguably "position statements". Summer was a time of (frantic) activity, with little room for discussion. This is actually a recurring criticism of Merkel's style, which is often characterized by her use of the word *alternativlos*, "without alternatives". Only in autumn did a proper debate take place. The conceptualization also appears to have changed from the spring one, in which we observe a majority of 7xx claims ("society"), indicating a more "foundational" debate. In summer, 1xx ("controlling migration") and 5xx ("foreign policy") claims enter the discourse, as a result of the need of coming to grips with the influx of migrants, on practical terms. In autumn/winter we observe all major categories.

Claim-level aggregation We now turn to claim-level aggregation over time. Figure 3 illustrates the discourse network centered on the claim 707, whose topic is the establishment of a legal distinction between economic immigrants and refugees⁷. The spring network displays the standard German discourse coalition configuration with respect to this topic: compassion vs. pragmatism. It mostly consists of conservative actors and pits the Catholic church (Marx) against the mainstream conservative government parties

⁷In Figure 3, the number of visible edges corresponds to claim frequency on distinct days. To reduce clutter, in Figure 4 we add up the frequency, unless the polarity differs.

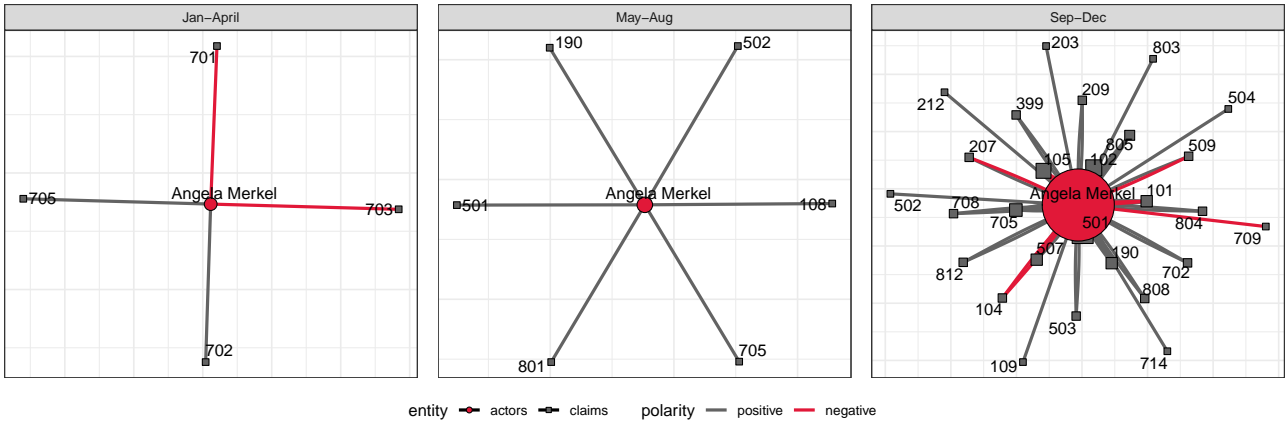


Figure 2: Angela Merkel network: (a) Jan-Apr; (b) May-Aug; (c) Sep-Dec

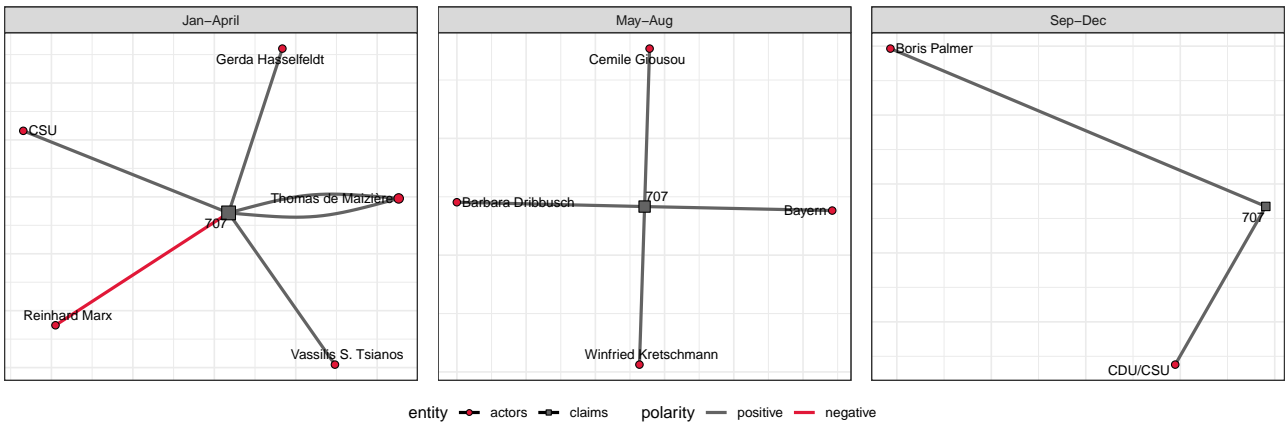


Figure 3: Claim 707 network: (a) Jan-Apr; (b) May-Aug; (c) Sep-Dec

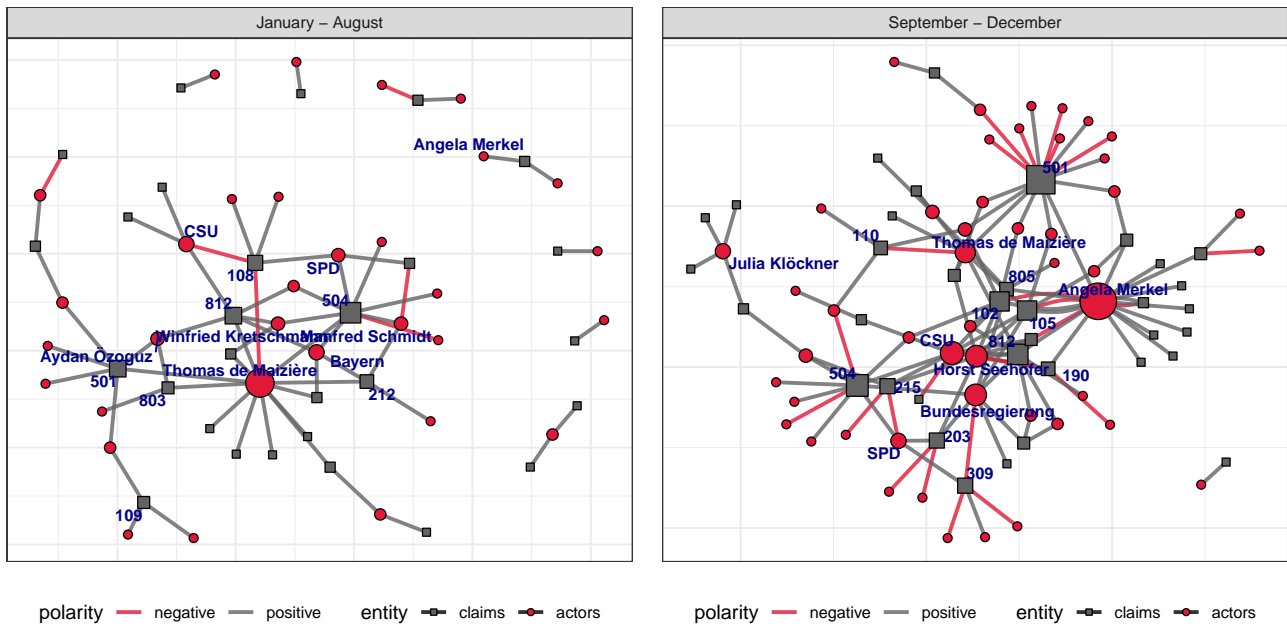


Figure 4: (Core) Discourse network: (a) Jan-Aug; (b) Sep-Dec

(Hasselfeldt, De Maizière, CSU). The summer network only contains voices in favor of 707: this can be interpreted as a reaction to the events as the increasing number of refugees resulted in many practical problems to be addressed. This gives rise to a more unusual discourse coalition consisting of a Green minister president (Kretschmann, who is admittedly known for being very conservative within his party), a journalist for a leftist newspaper (Dribbusch), and conservative players such as Bavaria (among the most conservative states and most hit by the immigrant wave) and Giousou (a conservative representative). The autumn network shows that this issue has taken a back seat compared to others, as few actors talk about 707.

While conclusions from such small samples need to be taken with a grain of salt, they illustrate the role that such networks, even built on relatively limited data, can play in formulating research hypotheses and empirically exploring them.

Discourse coalitions over time In the previous sections we have taken the perspective of a specific actor/claim and exploited *DEbateNet-mig15* to create a "snapshot" of the discourse centered on them. Yet, what a discourse network analysis ultimately aims to do is modeling the development of the discourse through an analysis of affiliation networks, as discussed in section 2.

Table 8 displays aggregated statistics of the whole *DEbateNet-mig15* network, on a monthly basis. For each month, we report number of observations (*Obs.*), number of claims (*C-token*), number of distinct claims *C-type*, number of distinct actors *Actors*, as well as the average *degree* centrality in the network, measured as the number of incident connections (edges) to each node (Wasserman and Faust, 1994, p. 100).

We observe that September is the month with the highest average degree. This indicates an increase in intensity of the discourse. Actors support more claims on average and certain claims receive enhanced support.

We conclude this illustration of the potential of discourse network analysis by showing the core discourse network for the months January/August and September/December, in figure 4. The core captures the central tendencies and structure within the network by only displaying claims which have been raised on at least two different days by the corresponding actor in the specified time frame (de Nooy et al., 2005, p. 109). The node-size corresponds to the prominence of actors and claims in terms of degree-centrality; additionally, these nodes tend to occupy more central positions in the network. We provide labels only for the nodes with a well above average degree-centrality, with the exception of Angela Merkel who, for the sake of comparison, is explicitly labelled in the left panel despite its low centrality.

Substantially, we can confirm the increase in actor participation (red nodes) and observe a change in the discourse dynamic. We exemplify this on two instances. Firstly: Although, chancellor Angela Merkel is no central figure in the earlier time period, she becomes by far the leading actor as the discourse progresses. Secondly: While the claim 501 (EU-Solution) has undisputed support in the core network from January to August, its support starts to crumble from September to December (as indicated by the red edges).

5. Conclusions

In this paper, we have introduced *DEbateNet-mig15*, an annotated dataset for the analysis of political debates, targeting the public discourse during the domestic debate on immigration in Germany in 2015. We have shown, by means of concrete examples, how our annotation framework can be exploited (in combination with discourse network analysis) to explore political science hypotheses concerning the dynamics of a debate.

A question we deliberately left aside is the potential of NLP methods to support the creation of such discourse networks. Our experience shows that semi-automatic, NLP-supported annotation is the right avenue for this synergy: while a fully automatic 97-class classification is not feasible (in particular given the amount of data that we have available), prediction of high-level categories works fairly well, and so does claim identification (Padó et al., 2019). In a recently concluded experiment (under review) we have tested the potential and limits of such semi-automatic classification with the MARDY annotation environment (Blessing et al., 2019). Our results show that, while semi-automatic annotation does not speed up the annotation, it has a positive impact on the inter-annotator agreement. Moreover, and crucially, we found that the combination of manual annotation on some annotation levels (e.g., fine-grained claim classification) and classifier predictions optimised for recall on some other levels (e.g., claim identification) yields excellent results if the goal is to identify the "core" (2-slice) network, thanks to the redundancy in the data.

Another issue we have disregarded in this paper has to do with the position of our work in the broader Argument Mining context. Our definition of claim corresponds to a subset of what is traditionally defined as an argument in NLP: the crucial distinctive feature is the "being targeted at an action". As far as the argumentative structure is concerned, the claim/justification relation from Argument Mining can be mapped into the claim/frame relation in political science. We have conducted some pilot annotation to investigate the potential of the framing annotation in a discourse network setting. While the results are promising, there is still much to be understood concerning the claim/actor dynamics, and we plan to do more steps in this direction before moving to a way more complex tripartite network, involving claims, actors and frames.

While the annotation contained in this release of *DEbateNet-mig15* has been conducted fully manually, ongoing work is currently exploiting semi-automatic support for the annotation of the immigration debate in different years (2005 and 2010) and on a different textual type (political parties manifestos). Besides, our investigation is also targeting a new topic, the domestic debate about pension policies in 2005, 2010 and 2015. In this connection, we are keeping constant the textual type (newspaper texts), the annotation framework and environment, but we obviously employ a different codebook and rely on different hypotheses concerning the debate dynamics.

6. Acknowledgments

We acknowledge funding by Deutsche Forschungsgemeinschaft (DFG) through MARDY (*Modeling Argumentation*

Dynamics) within SPP RATIO and by Bundesministerium für Bildung und Forschung (BMBF) through *Center for Reflected Text Analytics* (CRETA).

7. Bibliographical References

- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., and Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & society*, 19(3):273–306.
- Baker, P. (2004). Querying keywords: Questions of difference, frequency, and sense in keywords analysis. *Journal of English Linguistics*, 32(4):346–359.
- Blessing, A., Blokker, N., Haunss, S., Kuhn, J., Lapesa, G., and Padó, S. (2019). An Environment for the Relational Annotation of Political Debates. In *Proceedings of ACL System Demonstrations*, pages 105–110, Florence, Italy.
- Ceron, A., Curini, L., Iacus, S. M., and Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, 16(2):340–358.
- de Nooy, W., Mrvar, A., and Batagelj, V. (2005). *Exploratory Social Network Analysis with Pajek*. Cambridge University Press, Cambridge.
- de Wilde, P. (2011). No Polity for Old Politics? A Framework for Analyzing the Politicization of European Integration. *Journal of European Integration*, 33(5):559–575.
- Evert, S. (2008). Corpora and collocations. *Corpus linguistics. An international handbook*, 2:1212–1248.
- Hajer, M. A. (1993). Discourse Coalitions and the Institutionalization of Practice: The Case of Acid Rain in Britain. In *The Argumentative Turn in Policy Analysis and Planning*, pages 43–76. Duke University Press.
- Haunss, S. and Hofmann, J. (2015). Entstehung von Politikfeldern – Bedingungen einer Anomalie. *dms – der moderne staat*, 8(1):29–49.
- Haunss, S. and Kohlmorgen, L. (2009). Lobbying or Politics? Political Claims-Making in IP Conflicts. In Sebastian Haunss et al., editors, *Politics of Intellectual Property*, pages 107–128. Edward Elgar Publishing.
- Haunss, S., Dietz, M., and Nullmeier, F. (2013). Der Ausstieg aus der Atomenergie. Diskursnetzwerkanalyse als Beitrag zur Erklärung einer radikalen Politikwende. *Zeitschrift für Diskursforschung*, 1(3):288–316.
- Koopmans, R. and Statham, P. (1999). Political Claims Analysis: Integrating Protest Event And Political Discourse Approaches. *Mobilization*, 4(2):203–221.
- Kuhn, J. (2019). Computational text analysis within the humanities: How to combine working practices from the contributing fields? *Language Resources and Evaluation*, 53:565–602.
- Leifeld, P. and Haunss, S. (2012). Political Discourse Networks and the Conflict over Software Patents in Europe. *European Journal of Political Research*, 51(3):382–409.
- Leifeld, P. (2016). Discourse Network Analysis: Policy Debates as Dynamic Networks. In *The Oxford Handbook of Political Networks*. Oxford University Press.
- Padó, S., Blessing, A., Blokker, N., Dayanik, E., Haunss, S., and Kuhn, J. (2019). Who Sides With Whom? Towards Computational Construction of Discourse Networks for Political Debates. In *Proceedings of ACL*, pages 2841–2847, Florence, Italy.
- Peldszus, A. and Stede, M. (2013). From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence*, 7(1):1–31.
- Sabatier, P. A. and Weible, C. M. (2007). The Advocacy Coalition Framework: Innovations and Clarifications. In *Theories of the Policy Process*, pages 189–220. Westview Press.
- Schmidt, V. A. and Radaelli, C. M. (2004). Policy Change and Discourse in Europe: Conceptual and Methodological Issues. *West European Politics*, 27(2):183.
- Schneider, S., Nullmeier, F., and Hurrelmann, A. (2007). Exploring the communicative dimension of legitimacy: Text analytical approaches. In Achim Hurrelmann, et al., editors, *Legitimacy in an Age of Global Politics*, pages 126–155. Palgrave Macmillan.
- Stab, C. and Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.
- Swanson, R., Ecker, B., and Walker, M. (2015). Argument mining: Extracting arguments from online dialogue. In *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 217–226, Prague, Czech Republic.
- Vilares, D. and He, Y. (2017). Detecting perspectives in political debates. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1573–1582, Copenhagen, Denmark.
- Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Zürn, M. (2014). The politicization of world politics and its effects: Eight propositions. *European Political Science Review*, 6(01):47–71.