OK Boomer: Probing the socio-demographic Divide in Echo Chambers

Henri-Jacques Geiss[†] **Flora Sakketou**[‡] **Lucie Flek**[‡]

[†]Department of Computer Science, Technical University of Darmstadt
henri-jacques.geiss@stud.tu-darmstadt.de

[‡]Conversational AI and Social Analytics (CAISA) Lab
Department of Mathematics and Computer Science,
University of Marburg, Germany
{flora.sakketou, lucie.flek}@uni-marburg.de

Abstract

Social media platforms such as Twitter or Reddit have become an integral part in political opinion formation and discussions, accompanied by potential echo chamber forming. In this paper, we examine the relationships between the interaction patterns, the opinion polarity, and the socio-demographic characteristics in discussion communities on Reddit. On a dataset of over 2 million posts coming from over 20k users, we combine network community detection algorithms, reliable stance polarity annotations, and NLP-based socio-demographic estimations, to identify echo chambers and understand their properties at scale. We show that the separability of the interaction communities is more strongly correlated to the relative socio-demographic divide, rather than the stance polarity gap size. We further demonstrate that the socio-demographic classifiers have a strong topical bias and should be used with caution, merely for the relative community difference comparisons within a topic, rather than for any absolute labeling.

1 Introduction

Social media platforms such as Twitter or Reddit have become an integral part in political opinion formation and discussions, as users exchange opinions on numerous polarising topics such as gun control, abortion or healthcare. This process is accompanied by the forming of echo chambers, i.e. clusters formed by users with a homogeneous content production and diffusion (Cota et al., 2019), where users mostly see posts reinforcing their preexisting belief (DiFranzo and Gloria-Garcia, 2017; Barberá, 2015). At the same time, humans tend to be *homophile* in their social connections and interactions in general, implying that socio-demographic similarities in categories such as age, gender, ethnicity, religion or political ideology significantly increase the chance of a connection between two individuals in social networks (McPherson et al., 2001; Li et al., 2015; Himelboim et al., 2013).
So while we do know that similarities foster connections and *common identities* (Ren et al., 2007; McPherson et al., 2001) as well as that online communities can become echo chambers, by combining network-based community modeling, stance annotations, and socio-demographic projections from natural language of self-identified authors, it is possible to coarsely estimate the extent to which these phenomena inter-play such that this socio-demographic clustering is intensified in online echo chambers.
Based on that, we explore the following hypotheses in this paper: (i) key societal topics on Reddit shape network interaction communities indicating the echo chamber phenomenon, (ii) stance polarity mean values are further apart in more separated network communities, (iii) a distinct socio-demographic divide exists between groups of interacting users with diverse stance polarities showing echo chamber characteristics, (iv) automated socio-demographic profiling tools suffer from a strong topical bias, which hinders their ability to characterize the communities.

This paper provides the following contributions:

- We create a Reddit dataset of over 20k users (over 2M posts) within 8 current societal topics (Sec. 3), aligned with manual stance polarity annotations of 640 users (Sec. 4).
- We quantify the presence and extent of echo chambers in these discussions, employing network-based community detection metrics, such as separability and expansion, and the stance polarity annotations (Sec. 5).

- We develop classification models for socio-demographic variable estimates (age, gender, ideology) and find a strong topical bias, validating their use only for relative comparison of differences between communities rather than absolute labels (Sec. 6).
- By applying our socio-demographic classifiers on the detected and quantified network communities, we assess the echo-chamber phenomenon by identifying correlations between the relative difference in socio-demographic variables, the stance polarity differences, and the separability as well as expansion scores of the communities (Sec. 7).

2 Related Work

Recent works have either studied social-media data with regards to their graph-theoretical properties to detect echo-chamber-like phenomena from the user interactions (Barberá et al., 2015; Colleoni et al., 2014; Conover et al., 2011; Duseja and Jhamtani, 2019; Garimella et al., 2018) or estimated socio-demographic properties of social media users with NLP methods in isolation (Wiegmann et al., 2019; Wood-Doughty et al., 2018; Volkova and Bachrach, 2016; Burger et al., 2011). However, the combination of these two procedures in order to study a potential socio-demographic divide in such user groups at scale has, to the best of our knowledge, not been investigated so far.

In the broader context of analysing the political orientation of users in combination with their demographics, Barberá (2015) studies how Twitter users cluster with respect to different political leanings and shows that women tend to be on average slightly more liberal than men. A similar study demonstrated that there are differences in the average political leaning depending on gender, age, marital status and possession of a college degree (Bond and Messing, 2015), and observed that stronger ties between friends lead to a stronger correlation between their ideologies, which inspired us to the hypotheses explored in this paper.

On a similar note, Bamman et al. (2014) showed that mutual @-connections are more likely to appear between same-gender individuals. Comparable clustering effects were found for age as well as ideology (Li et al., 2015; Himelboim et al., 2013). Furthermore, Bastos et al. (2018) studied the relationship between echo chambers concerning the Brexit referendum on Twitter and the geographic location of its members, while Ebrahimi et al. (2016) found clear differences in the predicted stance towards Donald Trump between users from different US states, both embodying the idea of extracting social media stance-wise user groups and analysing their characteristics.

Regarding content-based models for the prediction of stance and socio-demographic properties, Durmus and Cardie (2018) studied discriminating tokens in the joined prediction of gender and stance towards abortion, finding that these correlate to the two labels differently, hinting towards our hypothesized topical bias for tokens that correlate more with stance than gender.

In the proposed work, we combine the users' political orientation, their estimated socio-demographic properties and their social media network, while previous works combine only a subset of these concepts, as shown in Table 1. Through a multifaceted analysis of the communities formed, we provide a more spherical insight into the relative differences of their members, in an attempt to analyze political opinion formation.

Authors	Demographics	Networks	Political stance
Barberá (2015)	✓		✓
Bond and Messing (2015)	✓		✓
Durmus and Cardie (2018)	✓		✓
Bamman et al. (2014)	✓	✓	
Li et al. (2015)	✓	✓	
Himelboim et al. (2013)		✓	✓
Bastos et al. (2018)		✓	✓
Ebrahimi et al. (2016)		✓	✓
Proposed method	✓	✓	✓

Table 1: Concepts covered by the related works

3 Dataset Characteristics

We used the API of Reddit to build our dataset. We chose Reddit as our source of data since it provides (i) rich content, due to the fact that there is no word limit, and (ii) a clear relationship between the text and the target topic, since users post within a subreddit. Previous work (Matthes et al., 2018) showed that the controversiality of the topic is one of the main drivers of opinion formation. Therefore, we manually compiled a set of contemporary discussion topics together with subreddits devoted to them (Table 7 in supplementary material for the 8 topics that were included). We crawled threads from these subreddits between November 2019 and June 2021 and periodically extended a database of posts and authors, preserving also the thread hier-

Topic	Reddit #Users	#Posts	#Posts/User
Abortion	3,747	631,177	168.4
Brexit	2,857	423,294	148.2
Capitalism	2,757	418,476	151.8
Climate change	1,117	269,032	240.9
Feminism	3,613	510,768	141.4
Gun control	5,192	667,477	128.6
Veganism	1,467	277,786	189.4
Nuclear-Energy	535	157,082	293.6
Total	**20,571**	**2,716,998**	**132**

Table 2: Amount of users, posts and posts per user for the studied topics. A user can be present in multiple topics, as we study in-topic interactions only.

archy.

For the study at hand, we selected the 8 most active topics, and for each of those, we extracted all users with at least 10 posts. The final dataset statistics are provided in Table 2.

4 Stance Polarity and Intensity Labels

As the opinion of the users towards the investigated topics is a central dimension when studying the echo chamber phenomenon, we discarded our automated stance classification efforts (F1-score around 60% on three classes) and utilized the human labels of the SPINOS dataset instead (Sakketou et al., 2022)[1]. Since the SPINOS dataset resembles a proper subset of the data that will be studied in this paper, it was deemed a feasible source for human labeled user stance samples. The stance labels and their corresponding numeric values are: *strongly against* (-2), *moderately against* (-1), *stance not inferrable* (0), *moderately in favor* (1) and *strongly in favor* (2).

The dataset consists of 3526 manually annotated posts from 640 users, which fully overlap with our Reddit data. We analyzed the annotated stances of each user and verified that most users consistently persist on a particular stance polarity. There are a few users with vacillating stances, who seem to mostly persist on one pole (either in favor or against) and express strong stance intensity only for that pole. We therefore compute each user's average stance based on the individual stances of their posts.

Figure 1 shows the distribution of the averaged user stances for each topic. We note that, based on

[1]https://github.com/caisa-lab/SPINOS-dataset

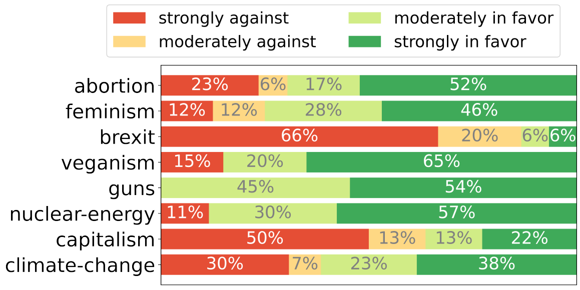


Figure 1: Average user stance distribution per topic

the annotation guidelines, a positive stance in the topic of feminism means being *in favor of equal rights for all genders*, a positive stance against climate change means believing that *climate change is caused by humans and constitutes a potential threat on survival* and a positive stance in the case of gun control means argumenting *in favor of the public availability of guns*.

5 Identifying Echo Chambers

Apart from the stance, a central aspect in the analysis of network structures, and especially echo chambers, in social media datasets is the definition of the interaction itself. Researchers have used retweets (Barberá et al., 2015; Conover et al., 2011) or follows (Colleoni et al., 2014; Duseja and Jhamtani, 2019; Garimella et al., 2018) to represent edges between user nodes. These however do not involve an explicit effort of content production (Cota et al., 2019), which is why we, following the work of Trabelsi and Zaiane (2018), focus on replies in the downward subtree of the post to extract the social network topologies. Figure 2 shows an example of how the post-reply tree of a social media post is transformed into the connecting edges of a user interaction network.

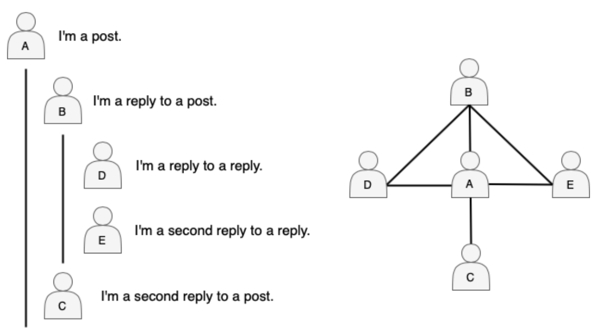


Figure 2: Examples of user interaction under a post and the resulting interaction network edges.

Detection Approach Similar to the studies of Conover et al. (2011) or Duseja and Jhamtani (2019), we search for interconnected user groups to detect echo-chamber-resembling structures in the extracted interaction topology. We experimented with the Louvain algorithm (Cota et al., 2019), the label propagation algorithm from Conover et al. (2011), and the Fluid algorithm (Parés et al., 2017). The latter yielded the best qualitative results.

To choose the optimal amount of communities, we encompass the fluid community detection with a meta-algorithm based on the notion of *modularity* of the detected communities. The meta algorithm runs the fluid community detection on the extracted network graphs from 2 up until 7 communities, and keeps track of the *modularity* of the created partitions, which is defined as:

$$\text{modularity} = m(p) = \sum_{c=1}^{n} \left[\frac{i_c}{m} - \gamma \left(\frac{k_c}{2m} \right)^2 \right]$$

where for a certain community c in a graph with m total edges, its number of internal edges is defined by i_c, the sum of degrees of the nodes in the community by k_c and the resolution parameter γ (Clauset et al., 2004; Hagberg et al., 2008).

In the end, the partition with the highest modularity score is returned. We chose *modularity* since it measures the division of the network into communities, i.e., whether there are only a few connections between the communities, while the nodes within them are densely connected (Clauset et al., 2004), which aligns with our goal of finding echo chambers. To ensure the consistency of our results even despite the elements of randomness in the community detection, 30 runs of the detection function are performed for each community amount, while still maximizing for *modularity*.

Based on these created partitions, we capture three graph community metrics (Yang and Leskovec, 2015), in order to measure the degree to which each distinct community represents an echo chamber in the network topology on variables that are not explicitly optimized during the community detection. We utilize *separability*, *expansion* and *density*, defined as follows for a community c:

$$\text{separability} = s(c) = \frac{i_c}{o_c}$$

$$\text{expansion} = e(c) = \frac{o_c}{n_c}$$

$$\text{density} = d(c) = \frac{i_c}{n_c(n_c - 1) \times 0.5}$$

here, the number of community-internal edges is defined by i_c, outbound edges by o_c, and community node count by n_c

The higher the values for *separability* and *density* of a detected community are, the more the interactions of its users are segregated from the rest of the network and rather take place with people from the same "bubble" and therefore represent an echo chamber. *Expansion* resembles echo chamber effects anti-proportionally as it is increased, when users of a community have more interactions with members from the other groups. To visualize the network nodes, we use the Fruchterman-Reingold force-directed placement (Fruchterman and Reingold, 1991).

We further provide the average manually annotated user stance, as described in section 4, for each of the detected interaction communities to explore if these also represent distinct stance clusters. Additionally, a weighted average stance for a community is determined by weighting the sampled stance values by the node degree of the users that contribute them.

Echo Chamber Identification Results We observe that the Reddit discussions take place in different network topology shapes, not all of them representing the echo chamber phenomenon. Rather, we distinguish three typical shapes:

1. Characteristic properties of the first structure, represented by the topic of nuclear-energy in the studied topics, are low values for *separability* ($\leq$ 0.5), a high minimum *expansion* (around 20), a rather uniform distribution of the sampled stances, as well as no visually separated cluster of user-nodes. In discussions of this type, there are no separated communities with opposing stances, rather all discussion participants acting as one community. Such is the case for nuclear energy (Figure 3a). This can be further validated by the manual stance annotations on this topic, where the majority of the average user stances is 'in favor of the use of nuclear energy' (Figure 1).
2. The second and most frequent structure is characterized by average *separability* and *expansion* values, presence of at least one cluster with a rather neutral stance, and visually clearly distinct communities that are spatially close to each other. While for the topics of gun control and Brexit all communities show a similar average sampled stance, in the cases of capitalism and abortion a cluster with strongly parti-

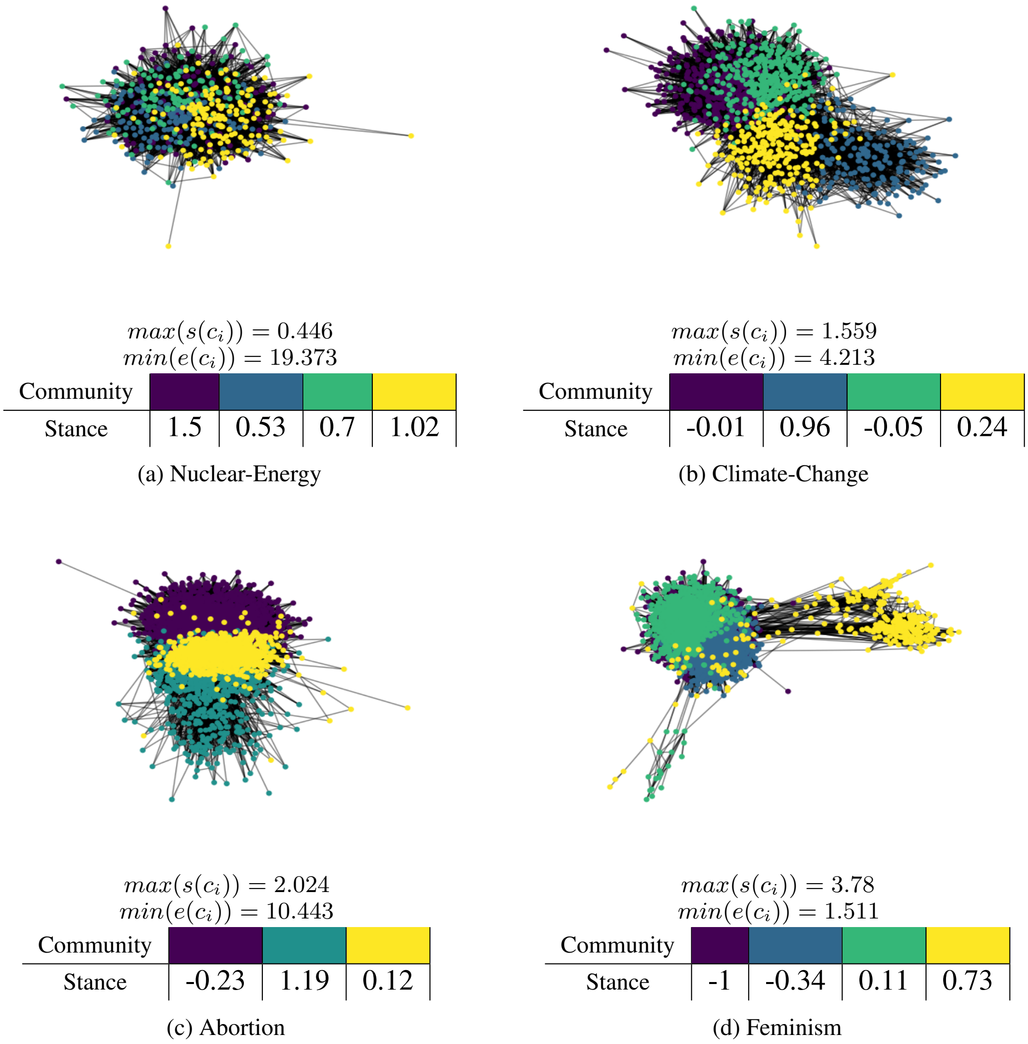


Figure 3: Four examples of topic interaction topologies and their detected communities on Reddit. $max(s(c_i))$ and $min(s(e_i))$ are the maximum *separability* and minimum *expansion* among the detected communities

san stances can be observed, hinting towards an echo-chamber-like "homogeneous content production and diffusion"(Cota et al., 2019).

3. The third structure type in the discussion topology, embodied by the topic of feminism in our analyzed data (Figure 3d), resembles the echo chamber phenomenon the most. In this case there is at least one detected community with a high *separability* score (≥ 4), a low minimum *expansion* (≤ 2) and at least one cluster with a clearly partisan average stance. The detected communities are also more spatially separated than in the other structure types. Here, like-minded individuals interact segregated from the rest of the network; an echo chamber is formed.

6 Socio-Demographic Prediction

To study how these community structures relate to their participants' socio-demographic traits we train interpretable supervised classifiers on datasets from previous social-media (Twitter) studies on gender, age, and political ideology (Preotiuc-Pietro et al., 2016; Preoţiuc-Pietro et al., 2017; Preoţiuc-Pietro and Ungar, 2018). For all three of these dimensions, it has been previously shown that social media users cluster along their labels (Bamman et al., 2014; Li et al., 2015; Himelboim et al., 2013). Following previous studies, and considering the available data volumes, we approach these tasks as classification.[2] We acknowledge the suboptimality of predicting binary gender labels and using self-reported training data with users having

[2]We are aware of the limitations and ethical risks that this simplification entails, as discussed in the Ethics section.

Label / Data	Tw1	Tw2	Tw3	Reddit
Male	34.3%	38.1%	34.8%	55.2%
Female	65.7%	61.9%	65.1%	44.8%
≤ 30	38.9%	39.9%	54.3%	37.3%
≤ 45	41.2%	43.7%	32.2%	31.2%
> 45	19.9%	16.4%	13.4%	31.5%
Liberal	-	-	50.3%	28.2%
Moderate	-	-	26.8%	-
Conservative	-	-	22.9%	71.8%

Table 3: Class distribution in training datasets

only binary option (Larson, 2017). We interpret the predictions in line with (Bem, 1974), examining if the discussion communities differ in constituent features around the class modes. The self-reported labels were obtained through the survey platforms Qualtrics and Amazon Mechanical Turk. Note that the actual posts of the twitter timelines we retrieved for each user might differ from the previous studies. We predict the user's: (i) self-reported age (three classes: below 30, between 30-45, and 46+), (ii) self-reported gender (male/female), and (iii) political ideology (conservative, moderate, or liberal).

Our training data from Twitter for age and gender consists of 3960 users. In order to directly include also users from reddit in the training data, we employ an automatic annotation generation for the dimensions of gender and age group based on regex-matching of 'I am'-statements in the user posts (Welch et al., 2020). For instance we annotate gender by searching for statements such as 'I am a guy/girl' or age with phrases such as 'I am X years old' or 'My grandson/granddaughter'. Users with multiple contradicting 'I am'-statements are excluded from the dataset. This way, we enhance our training data with 966 users from reddit, annotated for gender and 289 users for age. We then enhance the ideology training data with 1223 users from subreddits *r/Liberal* and *r/Conservative*, excluding users with less than 5 posts.

Feature settings (1) `TF-IDF`: We use the Porter Stemmer together with the TF-IDF weighting scheme (Manning et al., 2008).
(2) `Unigrams`: A user vector is calculated by summing up the appearances of every token, used by at least one percent of the training user base, across all the posts of one user and normalizing these values with the number of posts (Preoţiuc-Pietro and Ungar, 2018).
(3) `word2vec`: Spectral clustering of word embeddings creating a feature vector for a given set of posts from a user by calculating the proportion of tokens that belong to each of the topic clusters (Preoţiuc-Pietro and Ungar, 2018). These however didn't outperform the unigram and TF-IDF results.

We intentionally apply only easily interpretable classification models; linear SVM, logistic regression, and random forest. The best-performing setup for each of the user traits, which is used further in paper, is highlighted in Table 4.

	LinSVM	LogReg	RForest	Base
	Gender (Class-Balanced Down)			
tf-idf	0.735	0.693	0.769	0.5
word2vec	0.696	0.659	0.728	0.5
unigrams	**0.786**	0.7652	0.756	0.5
	Age (Class-Balanced Down)			
tf-idf	0.549	0.51	0.542	0.33
word2vec	0.516	0.492	0.546	0.33
unigrams	**0.577**	0.56	0.564	0.33
	Ideology (Class-Balanced Up)			
tf-idf	**0.587**	0.574	0.506	0.33
word2vec	0.563	0.557	0.524	0.33
unigrams	0.585	0.6	0.516	0.33

Table 4: Socio-demographic predictor accuracies with 5-fold cross-validation on balanced data

Socio-Demographic Prediction Analysis In the cases of gender and age group, the best-performing predictor(LinSVM) uses unigram-based user vectors, with accuracies of 79% and 58% respectively. For the prediction of political ideology, tf-idf features perform the best with 59%, more than 20% above the random prediction baseline.

We then analyze the predictive unigrams, extracting the feature score for each class from the LinSVM coefficient vector as per (Guyon and Elisseeff, 2003; Guyon et al., 2002). The results (Appendix) align with previous work, e.g. self-identified female users referring more to emotions (Burger et al., 2011; Carpenter et al., 2017).

Furthermore, Table 5 compares the predicted gender distribution of users participating in each topic with the more accurate, but sparser information detected by the regular expressions. We see that our content-based predictor tends to generally over-estimate the percentage of male users for most political topics. The two predictors are in more agreement on the three topics with the lowest amount of male participants, namely abortion, veganism-animalrights and feminism.

Cluster	Socio-demographics		
	Gender	**Age**	**Ideology**
Violet (-0.457)	M: 64.1% F: 35.9%	$\leq$ 30: 58.7% $\leq$ 45: 19.7% $>$ 45: 21.6%	Con: 52% Mod: 0.6% Lib: 47.5%
Green (1.05)	M: 25.5% F: 74.5%	$\leq$ 30: 60.6% $\leq$ 45: 21.2% $>$ 45: 18.2%	Con: 25.1% Mod: 6.6% Lib: 68.3%
Yellow (0.635)	M: 53.5% F: 46.5%	$\leq$ 30: 64.4% $\leq$ 45: 22.6% $>$ 45: 13%	Con: 44.3% Mod: 0.4% Lib: 55.3%

Cluster	Socio-demographics		
	Gender	**Age**	**Ideology**
Violet (0.503)	M: 94.1% F: 5.9%	$\leq$ 30: 29.7% $\leq$ 45: 25.6% $>$ 45: 44.8%	Con: 79.8% Mod: 0.3% Lib: 19.9%
Blue (0.362)	M: 95. % F: 5%	$\leq$ 30: 22.9% $\leq$ 45: 27.6% $>$ 45: 49.6%	Con: 79.9% Mod: 0.4% Lib: 19.8%
Green (0.372)	M: 96.6% F: 3.4%	$\leq$ 30: 23.6% $\leq$ 45: 25.4% $>$ 45: 51%	Con: 81.6% Mod: 0.4% Lib: 18%
Yellow (0.055)	M: 92.5% F: 7.5%	$\leq$ 30: 21.9% $\leq$ 45: 22.6% $>$ 45: 55.5%	Con: 77.1% Mod: 1% Lib: 21.9%

Figure 4: Predicted socio-demographic distributions of the detected communities in the discussion about **abortion** (left) and **gun control** (right) on Reddit. The clusters' degree-weighted average sampled stance is given in brackets.

Topic	**Predicted Gender** (M-F)	**Regex Gender** (M-F)	**Regex #Users**
abortion	53%-47%	39%-61%	222
climate-change	91%-9%	64%-36%	14
feminism	76%-24%	59%-41%	301
gun control	95%-5%	8-%2%	49
veganism	65%-35%	47%-53%	47
Brexit	94%-6%	71%-29%	24
capitalism	92%-8%	82%-18%	39
nuclear-energy	95%-5%	100%-0%	5

Table 5: Comparison of predicted gender proportions

7 Result of Combining the Studies

Labeling the posts of each user yields a percentage distribution for socio-demographic labels in the communities we extract from the interaction graphs of each topic. Figure 4 shows two examples of the determined socio-demographic distributions of a topic's communities and visually explains how we combined the systems derived in sections 5 and 6 for the following analyses (see Appendix for results on the rest of the topics).
Generally, we observe that diverse topics show diverse socio-demographic community profiles. For Abortion, the violet and green communities have opposing stances and large differences in the predicted gender and ideology distributions. In contrast, for Gun Control, all socio-demographic labels only differ by a small margin.

To formalize our hypothesis that the *relative* socio-demographic differences between the intra-topic community groups grow with the groups becoming more resembling to an echo chamber, we propose to measure, across all 8 topics, the correlation between the *separability* and *expansion* values of each community and the average RMSE (Equation 2) of each of the socio-demographic variables (Equation 1) of the detected clusters from the topic's baseline (i.e. the distribution for all users in the topic). A positive correlation in this case means that the more the communities of one topic resemble an echo chamber, the more they also differ in their socio-demographics. In Equation 1, d is the analyzed socio-demographic label, t is a certain topic with the corresponding full user base for this topic b_t and the ith detected community $c_{t,i}$, and $\text{pred}_d(x)$ is a function yielding the distribution of

	Stance	Stance σ	Gender	Age	Ideology
Separability	0.483	0.317	0.630	0.110	0.498
Expansion	-0.549	-0.090	-0.403	-0.170	-0.585

Table 6: Pearson's correlation of the *separability* and *expansion* values of the detected communities to (a) their mean stance, (b) their stance standard deviation (σ) and (c) their deviations from the full in-topic socio-demographics (Equation 1)

label d in a user group:

$$\text{sd-value}_d(b_t, c_{t,i}) = \text{rmse}(\text{pred}_d(b_t) - \text{pred}_d(c_{t,i})) \quad (1)$$

$$\text{rmse}(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2}, \quad x \in \mathbb{R}^N \quad (2)$$

Additionally, we measure the correlation of the two community metrics to the community's difference in stance from the complete topic's contributor average as well as the absolute values of standard deviation (σ). In the four stance-related correlations values, only communities with at least 5 stance-user-samples were considered.

Correlation Results The results in Table 6 indicate that values of *separability* and *expansion* that model the presence of an echo-chamber-resembling interaction network structure (high *separability* and low *expansion*) correlate with a larger separation of a sub-community in terms of gender and ideology of the topic's user average. Hence hypothesis (iii) holds - a distinct socio-demographic divide exists between groups of interacting users with diverse stance polarities showing echo chamber characteristics.

Furthermore, increased *separability* and decreased *expansion* also correlate with a stronger stance-wise segregation, confirming our hypothesis (ii) that stance polarity mean values are further apart in more separated network communities. That being said, these communities also show an increased standard deviation of stances, indicating that at least some variance in the opinion of contributing users is present, while more uniform network structures also tend to have more uniform stances.

8 Discussion and Limitations

The topic- and platform-specific environment underlines the limits of text-based user studies such as ours, indicating a lexical issue in the predictors used, confirming our hypothesis (iv) that the automated socio-demographic profiling tools suffer from a strong topical bias. While words such as *problem*, *understand*, or *politics* tend to be in general statistically more often used by self-identified men (Table 8), this does not hold when comparing discussions within a given topic. Similarly, while words like *women*, *mom* or *girl* are in general strong lexical cues for an author being female (compare Table 8), they tend to be used frequently by both genders just as a part of discussion about abortion or feminism. Similar issues occur with age models, leading to prediction biases. However, note that comparing *relative* differences (gaps) in estimated demographics between communities within one topic, as we did in Equation 1, is possible, as the bias merely shifts the distribution. In line with (Bem, 1974), we can still examine if the communities differ in constituent features around the class modes.

9 Summary and Conclusions

We explore the social media phenomenon of echo chambers with regards to its socio-demographic implications. To quantify the forming of these structures, we employed an interaction graph-based algorithm, exploring the *separability*, *density* and *expansion* of the detected communities. For the network topologies of abortion, capitalism, and feminism, we found a moderate to high resemblance of the echo-chamber phenomenon. Bridging the gap between theory and practice, these algorithm and measures could also be used by actual social-media platforms to track where its communities are structurally 'echo-chambered' and potential counter-measures are needed.

To capture the socio-demographic distributions of the detected communities, we trained interpretable socio-demographic estimation models, scrutinized by keyphrase-based approaches. By merging the network and content information, we found that more 'echo-chambered' topic communities also show an increased separation in their stance and gender and ideology profiles. These results reinforce the call for incorporating socio-demographic and network information into data sets and models for tasks like sentiment analysis, text generation and stance prediction (Hovy, 2015; Hovy and Yang, 2021), while keeping in mind that a lexical topic-related bias can be a source of misinterpretation in domain-specific user modeling.

Ethical Considerations

We acknowledge the suboptimality of predicting binary gender labels and using self-reported training data with users having only binary option (Larson, 2017). The topic- and platform-specific environment underlines the limits of such user studies. Any user-augmented classification efforts risk invoking stereotyping and essentialism, which can cause harm even if they are accurate on average differences (Rudman and Glick, 2008), and can be emphasized by the semblance of objectivity created by the use of a computer algorithm (Koolen and van Cranenburgh, 2017). It is important to be mindful of these effects when interpreting the model results in its application context. Use of any user data for socio-demographic estimates shall be transparent, and limited to the given aggregated purpose (Williams et al., 2017), no individual posts shall be republished and the study authors were advised to take account of users' privacy expectations (Williams et al., 2017; Shilton and Sayles, 2016; Townsend and Wallace, 2016) when collecting online data for user-based predictions. In our case, we utilize publicly available Reddit data in a purely observational, community-aggregated, and non-intrusive manner (Norval and Henderson, 2017) and restrain from any verbatim citations of the post contents.

Acknowledgements

This work has been supported by the German Federal Ministry of Education and Research (BMBF) as a part of the Junior AI Scientists program under the reference 01-S20060.

References

David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18(2):135–160.

Pablo Barberá. 2015. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political analysis*, 23(1):76–91.

Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. 2015. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542.

Marco Bastos, Dan Mercea, and Andrea Baronchelli. 2018. The geographic embedding of online echo chambers: Evidence from the brexit campaign. *PloS one*, 13(11):e0206841.

Sandra Bem. 1974. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2):155–162.

Robert Bond and Solomon Messing. 2015. Quantifying social media's political space: Estimating ideology from publicly revealed preferences on facebook. *American Political Science Review*, 109(1):62–78.

John D Burger, John Henderson, George Kim, and Guido Zarrella. 2011. Discriminating gender on twitter. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1301–1309.

Jordan Carpenter, Daniel Preotiuc-Pietro, Lucie Flekova, Salvatore Giorgi, Courtney Hagan, Margaret L Kern, Anneke EK Buffone, Lyle Ungar, and Martin EP Seligman. 2017. Real men don't say "cute" using automatic language analysis to isolate inaccurate aspects of stereotypes. *Social Psychological and Personality Science*, 8(3):310–322.

Aaron Clauset, Mark EJ Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Physical review E*, 70(6):066111.

Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. 2014. Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of communication*, 64(2):317–332.

Michael D Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. 2011. Predicting the political alignment of twitter users. In *2011 IEEE third international conference on privacy, security, risk and trust*, pages 192–199. IEEE.

Wesley Cota, Silvio C Ferreira, Romualdo Pastor-Satorras, and Michele Starnini. 2019. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science*, 8(1):35.

Dominic DiFranzo and Kristine Gloria-Garcia. 2017. Filter bubbles and fake news. *XRDS: Crossroads, The ACM Magazine for Students*, 23(3):32–35.

Esin Durmus and Claire Cardie. 2018. Understanding the effect of gender and stance in opinion expression in debates on "abortion". In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 69–75.

Nikita Duseja and Harsh Jhamtani. 2019. A sociolinguistic study of online echo chambers on twitter. In *Proceedings of the Third Workshop on Natural Language Processing and Computational Social Science*, pages 78–83.

Javid Ebrahimi, Dejing Dou, and Daniel Lowd. 2016. Weakly supervised tweet stance classification by relational bootstrapping. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 1012–1017.

Thomas MJ Fruchterman and Edward M Reingold. 1991. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164.

Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference*, pages 913–922.

Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182.

Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3):389–422.

Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

Itai Himelboim, Stephen McCreery, and Marc Smith. 2013. Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on twitter. *Journal of computer-mediated communication*, 18(2):154–174.

Dirk Hovy. 2015. Demographic factors improve classification performance. In *Proceedings of the 53rd annual meeting of the Association for Computational Linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*, pages 752–762.

Dirk Hovy and Diyi Yang. 2021. The importance of modeling social factors of language: Theory and practice. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 588–602, Online. Association for Computational Linguistics.

Corina Koolen and Andreas van Cranenburgh. 2017. These are not the stereotypes you are looking for: Bias and fairness in authorial gender attribution. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pages 12–22, Valencia, Spain. Association for Computational Linguistics.

Brian N Larson. 2017. Gender as a variable in natural-language processing: Ethical considerations. *EACL 2017*, page 1.

Peter Li, Jiejun Xu, and Tsai-Ching Lu. 2015. Leveraging homophily to infer demographic attributes: Inferring the age of twitter users using label propagation. In *Proceedings of Workshop on Information In Networks (WIN15)*.

Christopher D Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to information retrieval*. Cambridge university press.

Jörg Matthes, Johannes Knoll, and Christian von Sikorski. 2018. The "spiral of silence" revisited: A meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research*, 45(1):3–33.

Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.

Chris Norval and Tristan Henderson. 2017. Contextual consent: Ethical mining of social media for health research. *CoRR*, abs/1701.07765.

Ferran Parés, Dario Garcia Gasulla, Armand Vilalta, Jonatan Moreno, Eduard Ayguadé, Jesús Labarta, Ulises Cortés, and Toyotaro Suzumura. 2017. Fluid communities: A competitive, scalable and diverse community detection algorithm. In *International Conference on Complex Networks and their Applications*, pages 229–240. Springer.

Daniel Preotiuc-Pietro, Jordan Carpenter, Salvatore Giorgi, and Lyle Ungar. 2016. Studying the dark triad of personality through twitter behavior. In *Proceedings of the 25th ACM international on conference on information and knowledge management*, pages 761–770.

Daniel Preoţiuc-Pietro, Ye Liu, Daniel Hopkins, and Lyle Ungar. 2017. Beyond binary labels: political ideology prediction of twitter users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pages 729–740.

Daniel Preoţiuc-Pietro and Lyle Ungar. 2018. User-level race and ethnicity predictors from twitter text. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1534–1545.

Yuqing Ren, Robert Kraut, and Sara Kiesler. 2007. Applying common identity and bond theory to design of online communities. *Organization studies*, 28(3):377–408.

Laurie A Rudman and Peter Glick. 2008. The social psychology of gender: How power and intimacy shape gender relations.

Flora Sakketou, Allison Lahnala, Liane Vogel, and Lucie Flek. 2022. Investigating user radicalization: A novel dataset for identifying fine-grained temporal shifts in opinion.

Katie Shilton and Sheridan Sayles. 2016. " we aren't all going to be on the same page about ethics": Ethical practices and challenges in research on digital and social media. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 1909–1918. IEEE.

Leanne Townsend and Claire Wallace. 2016. Social media research: A guide to ethics. *University of Aberdeen*, 1:16.

Amine Trabelsi and Osmar Zaiane. 2018. Unsupervised model for topic viewpoint discovery in online debates leveraging author interactions. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12.

Svitlana Volkova and Yoram Bachrach. 2016. Inferring perceived demographics from user emotional tone and user-environment emotional contrast. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 1567–1578.

Charles Welch, Jonathan K. Kummerfeld, Verónica Pérez-Rosas, and Rada Mihalcea. 2020. Compositional demographic word embeddings. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*.

Matti Wiegmann, Benno Stein, and Martin Potthast. 2019. Celebrity profiling. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2611–2618.

Matthew L Williams, Pete Burnap, and Luke Sloan. 2017. Towards an ethical framework for publishing twitter data in social research: Taking into account users' views, online context and algorithmic estimation. *Sociology*, 51(6):1149–1168.

Zach Wood-Doughty, Nicholas Andrews, Rebecca Marvin, and Mark Dredze. 2018. Predicting twitter user demographics from names alone. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 105–111.

Jaewon Yang and Jure Leskovec. 2015. Defining and evaluating network communities based on ground-truth. *Knowledge and Information Systems*, 42(1):181–213.

A Supplemental Material

A.1 Annotated List of Subreddits for each of the Studied Topics

Topic	Subreddits
Abortion	‘abortion’, ‘Abortiondebate’, ‘prochoice’, ‘prolife’, ‘trueprochoice’, ‘Insaneprochoice’, ‘ProLifeLibertarians’, ‘ThingsProChoicersSay’, ‘AskProchoice’, ‘insaneprolife’, ‘abortionopinions’
Brexit	‘brexit’, ‘brealism’
Capitalism	‘CapitalismVSocialism’, ‘DebateCommunism’, ‘SocialismVCapitalism’, ‘occupywallstreet’, ‘Capitalism’, ‘communism’
Climate Change	‘climate’, ‘climatechange’, ‘climateskeptics’, ‘GlobalClimateChange’, ‘FridaysForFuture’
Feminism	‘DebateFeminism’, ‘feminisms’, ‘feministtheory’, ‘GenderCritical’, ‘RadicalFeminism’, ‘INeedFeminismBecause’, ‘meToo’, ‘masculinism’, ‘Egalitarianism’, ‘masculism’, ‘MensRights’, ‘MRActivism’, ‘MenGetRapedToo’, ‘LeftistsForMen’, ‘feminismformen’, ‘mensrightslinks’, ‘antifeminists’, ‘Feminism’, ‘Radical_Feminists’, ‘RadicalFeminismUSA’
Gun control	‘guncontrol’, ‘GunDebates’, ‘gunpolitics’, ‘GunResearch’, ‘GunsAreCool’, ‘progun’, ‘liberalgunowners’, ‘Firearms’
Nuclear-Energy	‘nuclear’, ‘NuclearEnergy’, ‘NuclearPower’
Veganism-Animalrights	‘AnimalRights’, ‘animalwelfare’, ‘VeganActivism’, ‘Veganism’, ‘Vegetarianism’, ‘Veganity’, ‘vegproblems’, ‘AntiVegan’, ‘DebateAVegan’, ‘debatemeateaters’, ‘exvegans’

Table 7: The subreddits that were crawled to creat the dataset from which the studied users, their posts and the interaction graphs were extracted

A.2 Unigram Coefficients

Female		Male	
2.805	girl	-2.597	game
2.723	love	-2.224	men
2.572	♀	-2.088	wife
1.781	book	-2.05	♂
1.653	bodi	-1.89	man
1.611	so	-1.751	good
1.611	about	-1.627	bro
1.606	woman	-1.509	some
1.583	omg	-1.506	back
1.482	women	-1.481	#x200b
1.469	no	-1.439	🔥
1.442	oh	-1.404	guy
1.41	senat	-1.349	beat
1.338	cute	-1.287	doe
1.321	💜	-1.281	player
1.317	pleas	-1.266	look
1.29	friend	-1.264	war
1.281	😂	-1.263	problem
1.279	thing	-1.225	coronaviru
1.27	mom	-1.215	enjoy
1.267	:)	-1.195	year
1.263	hous	-1.167	en
1.239	are	-1.159	3
1.218	birthday	-1.143	mplusreward
1.208	husband	-1.13	harm
1.198	ad	-1.126	should
1.193	excit	-1.11	great
1.179	sticker	-1.097	shit
1.178	color	-1.09	check
1.175	ye	-1.07	time
1.156	stop	-1.049	much
1.137	he	-1.048	comic
1.135	didn't	-1.043	If
1.118	okay	-1.039	understand
1.089	public	-1.03	valu
1.08	cooki	-1.023	#es161
1.074	serious	-1.02	complet
1.062	danc	-1.013	down
1.061	mental	-0.996	against
1.061	heart	-0.986	youtub
1.061	night	-0.981	mpoint
1.06	text	-0.978	app
1.055	tweet	-0.971	hi

Table 8: Gender svc-model coefficients for unigrams

≤ 30		≤ 45		> 45	
1.712	be	2.037	right	1.411	she
1.688	actual	1.278	movi	1.228	enter
1.433	is	1.269	mean	1.216	have
1.334	my	1.173	excit	1.194	pleas
1.305	i'm	1.166	tri	1.181	those
1.299	Me	1.139	fun	1.096	via
1.295	it'	1.124	or	1.04	thank
1.284	gonna	1.087	🙏	1.01	he
1.238	life	1.057	aquariu	1.008	well
1.199	so	1.056	awesom	1.006	hi
1.129		1.022	ago	1.004	ani
1.097	like	1.016	teacher	0.989	great
1.078	an	0.989	white	0.973	#ifb
1.076	class	0.969	wait	0.965	by
1.055	day	0.953	kid	0.964	must
1.013	becaus	0.949	babi	0.954	read
0.996	😭	0.941	leo	0.953	#photographi
0.977	🌼	0.935	bad	0.925	they
0.945	y'all	0.933	product	0.921	veri
0.944	wanna	0.921	man	0.903	place
0.924	:)	0.887	some	0.893	most
0.889	pop	0.886	...	0.875	die
0.881	3	0.884	chat	0.873	100
0.878	🏻	0.825	year	0.849	💜
0.876	<3	0.817	idea	0.847	scorpio
0.876	i	0.816	exactli	0.819	video
0.86	okay	0.806	free	0.816	there
0.85	give	0.796	episod	0.809	oia
0.847	punchcard	0.791	#winterofzombi	0.798	trump
0.836	you'r	0.788	odd	0.795	daughter
0.833	can't	0.782	#saveforev	0.789	see
0.825	–	0.768	week	0.779	use
0.825	shop	0.761	narcissist	0.778	safe
0.82	1,000	0.759	great	0.776	would
0.818	nigga	0.756	mayb	0.759	your
0.796	dailylook	0.753	total	0.746	happi
0.796	charact	0.748	#debatenight	0.744	were
0.777	no	0.745	then	0.741	:)
0.777	#cochlearimpl	0.743	guess	0.74	stay
0.767	berni	0.738	can't	0.739	now
0.766	chang	0.737	wow	0.734	here

Table 9: Age group svc-model coefficients for unigrams

Conservative	
1.405	polic
1.154	leftist
1.132	chat
1.114	polit
1.103	protest
1.046	kid
1.036	👱🏼
1.027	littl
1.009	they
1.005	it’
1.004	blm
0.98	😘
0.974	democrat
0.972	call
0.962	On
0.941	left
0.923	❤️
0.904	kind
0.9	jesu
0.889	state
0.879	that’
0.876	know
0.861	hillari
0.86	mani
0.851	😺
0.85	school
0.845	also
0.843	seem
0.842	legal
0.835	which
0.814	look
0.81	china
0.808	hospit
0.794	rather
0.787	viru
0.774	these
0.771	down
0.762	’
0.759	own
0.759	bill
0.756	ye
0.743	clinton
0.737	illeg
0.737	around
0.732	sens

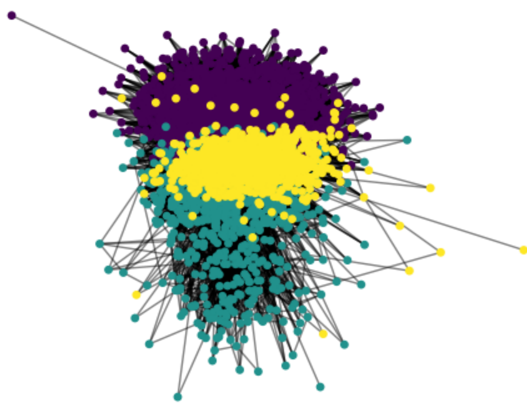
Moderate	
1.225	wow
1.107	back
1.087	by
1.075	😂
1.07	money
1.052	love
0.995	stream
0.97	can
0.963	🙏
0.926	🥳
0.917	dot
0.909	😍
0.877	pack
0.857	Me
0.833	...
0.804	summer
0.791	show
0.786	realli
0.779	game
0.773	time
0.755	play
0.749	enter
0.748	get
0.748	#stevenunivers
0.744	need
0.737	when
0.733	👀
0.731	befor
0.728	come
0.728	->
0.724	🔥
0.717	school
0.715	app
0.712	best
0.696	free
0.685	coronaviru
0.681	reach
0.679	awesom
0.666	learn
0.663	y
0.657	reward
0.65	join
0.647	from
0.646	$

Liberal	
1.347	💚
1.206	omg
1.177	prize
1.137	write
1.112	save
1.111	tweet
1.092	📷
1.044	serious
1.025	We
0.987	still
0.977	episod
0.967	women
0.938	so
0.934	work
0.909	#voicesaveindia
0.903	fox
0.881	movi
0.872	chang
0.855	damn
0.853	pandem
0.851	spnwithlov
0.848	law
0.847	anyth
0.844	he’
0.84	again
0.828	right
0.821	food
0.819	trump
0.819	+
0.807	think
0.792	today
0.774	|
0.773	white
0.772	youtub
0.756	stay
0.753	&
0.751	mean
0.749	protect
0.748	gay
0.741	kat
0.741	stori
0.736	everi
0.727	dog
0.725	and
0.724	beauti

Table 10: Ideology svc-model coefficients for unigrams

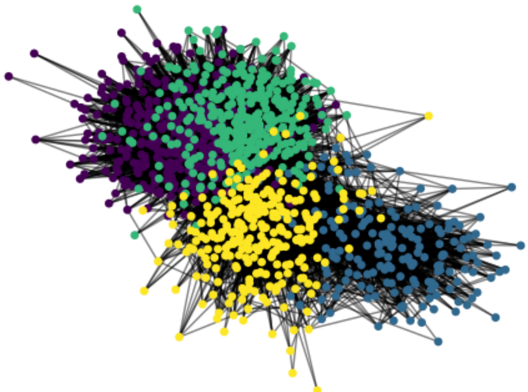
A.3 Full results of community detection and socio-demographic prediction

The following are the complete study results for all eight topics. They include the interaction graph with its detected communities as well as a table presenting each communities' user count, weighted and unweighted annotated stance, graph community metrics and predicted socio-demographic distributions. All correlations and analyses in the paper were based on these results.



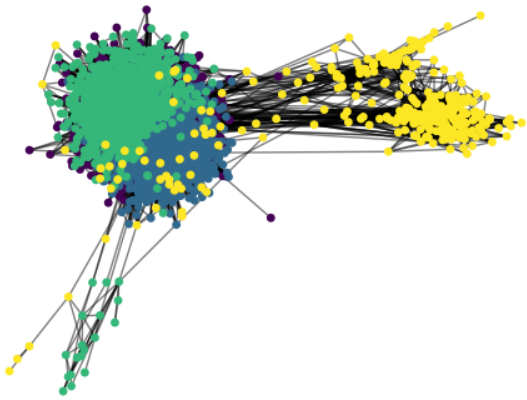
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	1776	∅: -0.231 Std: 0.924 #Users: 36.0	∅: -0.457 Std: 0.706	0.024	2.024	10.443	M: 0.641 F: 0.359	$\leq$ 30: 0.587 $\leq$ 45: 0.197 $>$ 45: 0.216	Con: 0.52 Mod: 0.006 Lib: 0.475
1	797	∅: 1.185 Std: 0.584 #Users: 33.0	∅: 1.05 Std: 0.693	0.027	0.687	15.764	M: 0.255 F: 0.745	$\leq$ 30: 0.606 $\leq$ 45: 0.212 $>$ 45: 0.182	Con: 0.251 Mod: 0.066 Lib: 0.683
2	1168	∅: 0.199 Std: 1.051 #Users: 118.0	∅: 0.635 Std: 0.947	0.047	1.273	21.452	M: 0.535 F: 0.465	$\leq$ 30: 0.644 $\leq$ 45: 0.226 $>$ 45: 0.13	Con: 0.443 Mod: 0.004 Lib: 0.553

Figure 5: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **abortion** on Reddit. The weighted stance is calculated based on the user's degree in the graph



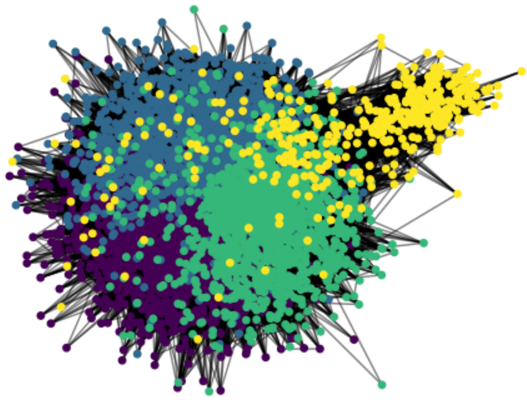
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	337	⌀: -0.007 Std: 1.042 #Users: 8.0	⌀: -0.372 Std: 0.896	0.059	0.764	12.896	M: 0.899 F: 0.101	$\leq$ 30: 0.291 $\leq$ 45: 0.237 $>$ 45: 0.472	Con: 0.656 Mod: 0.009 Lib: 0.335
1	178	⌀: 0.963 Std: 0.129 #Users: 4.0	⌀: 0.97 Std: 0.101	0.074	1.559	4.213	M: 0.933 F: 0.067	$\leq$ 30: 0.152 $\leq$ 45: 0.309 $>$ 45: 0.539	Con: 0.438 Mod: 0.011 Lib: 0.551
2	336	⌀: -0.047 Std: 0.751 #Users: 6.0	⌀: -0.281 Std: 0.455	0.068	0.829	13.688	M: 0.905 F: 0.095	$\leq$ 30: 0.321 $\leq$ 45: 0.196 $>$ 45: 0.482	Con: 0.634 Mod: 0.009 Lib: 0.357
3	262	⌀: 0.243 Std: 1.226 #Users: 15.0	⌀: 0.521 Std: 1.002	0.065	0.986	8.553	M: 0.927 F: 0.073	$\leq$ 30: 0.248 $\leq$ 45: 0.279 $>$ 45: 0.473	Con: 0.531 Mod: 0.031 Lib: 0.439

Figure 6: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **climate-change** on Reddit. The weighted stance is calculated based on the user's degree in the graph



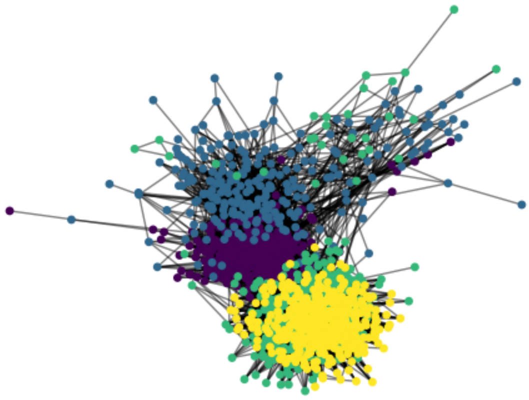
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	930	⌀: -1.0 Std: 0.4 #Users: 2.0	⌀: -0.761 Std: 0.321	0.018	0.421	20.289	M: 0.824 F: 0.176	$\leq$ 30: 0.487 $\leq$ 45: 0.298 $>$ 45: 0.215	Con: 0.499 Mod: 0.013 Lib: 0.488
1	1176	⌀: -0.34 Std: 0.694 #Users: 16.0	⌀: -0.589 Std: 0.42	0.022	0.733	17.426	M: 0.798 F: 0.202	$\leq$ 30: 0.367 $\leq$ 45: 0.355 $>$ 45: 0.277	Con: 0.493 Mod: 0.019 Lib: 0.488
2	1168	⌀: 0.113 Std: 1.012 #Users: 26.0	⌀: -0.604 Std: 0.796	0.023	0.757	17.997	M: 0.782 F: 0.218	$\leq$ 30: 0.447 $\leq$ 45: 0.347 $>$ 45: 0.206	Con: 0.427 Mod: 0.011 Lib: 0.562
3	331	⌀: 0.728 Std: 0.9 #Users: 32.0	⌀: 0.538 Std: 0.944	0.035	3.78	1.511	M: 0.353 F: 0.647	$\leq$ 30: 0.344 $\leq$ 45: 0.369 $>$ 45: 0.287	Con: 0.269 Mod: 0.012 Lib: 0.719

Figure 7: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **feminism** on Reddit. The weighted stance is calculated based on the user's degree in the graph



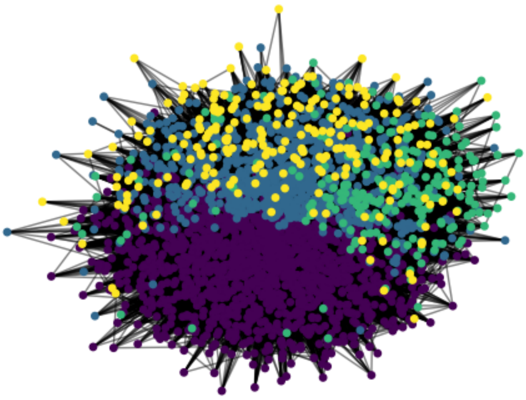
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	1705	⌀: 0.472 Std: 0.572 #Users: 9.0	⌀: 0.503 Std: 0.615	0.018	0.597	26.266	M: 0.941 F: 0.059	$\leq$ 30: 0.297 $\leq$ 45: 0.256 $>$ 45: 0.448	Con: 0.798 Mod: 0.003 Lib: 0.199
1	1574	⌀: 0.517 Std: 0.717 #Users: 12.0	⌀: 0.362 Std: 0.636	0.02	0.535	28.745	M: 0.95 F: 0.05	$\leq$ 30: 0.229 $\leq$ 45: 0.276 $>$ 45: 0.496	Con: 0.799 Mod: 0.004 Lib: 0.198
2	1509	⌀: 0.435 Std: 0.548 #Users: 8.0	⌀: 0.372 Std: 0.571	0.023	0.549	31.557	M: 0.966 F: 0.034	$\leq$ 30: 0.236 $\leq$ 45: 0.254 $>$ 45: 0.51	Con: 0.816 Mod: 0.004 Lib: 0.18
3	398	⌀: 0.333 Std: 0.333 #Users: 2.0	⌀: 0.055 Std: 0.183	0.043	0.867	9.872	M: 0.925 F: 0.075	$\leq$ 30: 0.219 $\leq$ 45: 0.226 $>$ 45: 0.555	Con: 0.771 Mod: 0.01 Lib: 0.219

Figure 8: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **guncontrol** on Reddit. The weighted stance is calculated based on the user's degree in the graph



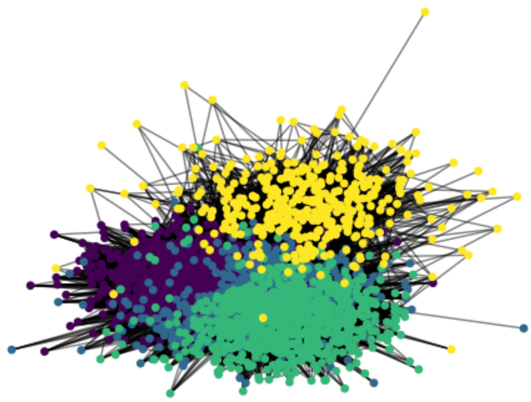
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	407	⌀: 0.333 Std: 1.247 #Users: 3.0	⌀: 0.312 Std: 1.21	0.045	2.405	3.828	M: 0.688 F: 0.312	$\leq$ 30: 0.57 $\leq$ 45: 0.204 $>$ 45: 0.226	Con: 0.386 Mod: 0.01 Lib: 0.604
1	248	⌀: 0.843 Std: 1.083 #Users: 23.0	⌀: 1.067 Std: 0.696	0.027	1.354	2.472	M: 0.573 F: 0.427	$\leq$ 30: 0.496 $\leq$ 45: 0.19 $>$ 45: 0.315	Con: 0.298 Mod: 0.056 Lib: 0.645
2	338	⌀: 1.667 Std: 0.471 #Users: 3.0	⌀: 1.667 Std: 0.471	0.051	0.737	11.577	M: 0.55 F: 0.45	$\leq$ 30: 0.503 $\leq$ 45: 0.157 $>$ 45: 0.34	Con: 0.337 Mod: 0.041 Lib: 0.621
3	471	⌀: -1.167 Std: 0.0 #Users: 1.0	⌀: -1.167 Std: 0.0	0.063	1.734	8.548	M: 0.724 F: 0.276	$\leq$ 30: 0.473 $\leq$ 45: 0.187 $>$ 45: 0.34	Con: 0.461 Mod: 0.013 Lib: 0.527

Figure 9: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **veganism-animalrights** on Reddit. The weighted stance is calculated based on the user's degree in the graph



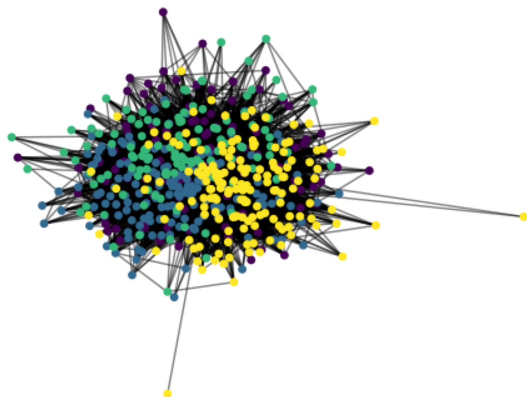
Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	1460	⌀: -0.73 Std: 0.754 #Users: 27.0	⌀: -0.829 Std: 0.555	0.049	1.343	26.633	M: 0.942 F: 0.058	$\leq$ 30: 0.14 $\leq$ 45: 0.338 $>$ 45: 0.521	Con: 0.701 Mod: 0.023 Lib: 0.276
1	903	⌀: -0.53 Std: 1.029 #Users: 35.0	⌀: -0.561 Std: 1.047	0.049	0.499	44.104	M: 0.947 F: 0.053	$\leq$ 30: 0.175 $\leq$ 45: 0.368 $>$ 45: 0.457	Con: 0.714 Mod: 0.014 Lib: 0.271
2	259	⌀: 0.0 Std: 0.0 #Users: 0.0	⌀: 0.0 Std: 0.0	0.046	0.17	35.062	M: 0.919 F: 0.081	$\leq$ 30: 0.154 $\leq$ 45: 0.421 $>$ 45: 0.425	Con: 0.73 Mod: 0.012 Lib: 0.259
3	229	⌀: -0.853 Std: 0.547 #Users: 5.0	⌀: -0.977 Std: 0.212	0.028	0.102	30.825	M: 0.891 F: 0.109	$\leq$ 30: 0.288 $\leq$ 45: 0.319 $>$ 45: 0.393	Con: 0.664 Mod: 0.017 Lib: 0.319

Figure 10: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **brexit** on Reddit. The weighted stance is calculated based on the user's degree in the graph



Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	845	⌀: 0.0 Std: 0.816 #Users: 3.0	⌀: 0.176 Std: 0.472	0.047	0.813	24.555	M: 0.931 F: 0.069	$\leq$ 30: 0.346 $\leq$ 45: 0.452 $>$ 45: 0.202	Con: 0.685 Mod: 0.004 Lib: 0.311
1	615	⌀: -0.295 Std: 1.212 #Users: 3.0	⌀: -1.063 Std: 1.263	0.039	0.345	35.093	M: 0.914 F: 0.086	$\leq$ 30: 0.315 $\leq$ 45: 0.506 $>$ 45: 0.179	Con: 0.685 Mod: 0.002 Lib: 0.314
2	898	⌀: 0.14 Std: 0.648 #Users: 37.0	⌀: 0.115 Std: 0.701	0.039	0.684	25.758	M: 0.93 F: 0.07	$\leq$ 30: 0.331 $\leq$ 45: 0.453 $>$ 45: 0.216	Con: 0.688 Mod: 0.004 Lib: 0.307
3	396	⌀: -0.688 Std: 0.872 #Users: 18.0	⌀: -0.533 Std: 0.858	0.034	0.903	7.328	M: 0.917 F: 0.083	$\leq$ 30: 0.338 $\leq$ 45: 0.359 $>$ 45: 0.303	Con: 0.614 Mod: 0.008 Lib: 0.379

Figure 11: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **capitalism** on Reddit. The weighted stance is calculated based on the user's degree in the graph



Cluster				Metrics			Sociodemographics		
	#Users	**stance**	**weighted stance**	**d(c)**	**s(c)**	**e(c)**	**Gender**	**Age**	**Ideology**
0	105	⌀: 1.5 Std: 0.5 #Users: 2.0	⌀: 1.211 Std: 0.408	0.107	0.28	19.819	M: 0.971 F: 0.029	$\leq$ 30: 0.276 $\leq$ 45: 0.276 $>$ 45: 0.448	Con: 0.467 Mod: 0.067 Lib: 0.467
1	147	⌀: 0.533 Std: 0.972 #Users: 10.0	⌀: 0.699 Std: 0.641	0.117	0.427	19.98	M: 0.952 F: 0.048	$\leq$ 30: 0.34 $\leq$ 45: 0.299 $>$ 45: 0.361	Con: 0.49 Mod: 0.041 Lib: 0.469
2	123	⌀: 0.7 Std: 0.4 #Users: 5.0	⌀: 0.76 Std: 0.425	0.109	0.316	21.146	M: 0.927 F: 0.073	$\leq$ 30: 0.382 $\leq$ 45: 0.333 $>$ 45: 0.285	Con: 0.341 Mod: 0.057 Lib: 0.602
3	158	⌀: 1.016 Std: 0.778 #Users: 22.0	⌀: 1.146 Std: 0.443	0.11	0.446	19.373	M: 0.956 F: 0.044	$\leq$ 30: 0.31 $\leq$ 45: 0.285 $>$ 45: 0.405	Con: 0.456 Mod: 0.044 Lib: 0.5

Figure 12: Sampled stance (unweighted and weighted average), *separability* s(c), *density* d(c), *expansion* e(c) and predicted socio-demographics of the detected communities in the discussion around **nuclear-energy** on Reddit. The weighted stance is calculated based on the user's degree in the graph