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.

In 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-
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.

| Topic            | #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** |

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). 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 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 arguing 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.
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 $\gamma$ (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-
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; Preoțiu-Pietro et al., 2017; Preoțiu-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.\footnote{We are aware of the limitations and ethical risks that this simplification entails, as discussed in the Ethics section.} We acknowledge the sub-optimality of predicting binary gender labels and using self-reported training data with users having...
Table 3: Class distribution in training datasets

| 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% |

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 (Preotiuc-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 (Preotiuc-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.

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

| Gender (Class-Balanced Down) | LinSVM | LogReg | RForest | Base |
|------------------------------|--------|--------|---------|------|
| 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)    | LinSVM | LogReg | RForest | Base |
|------------------------------|--------|--------|---------|------|
| 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) | LinSVM | LogReg | RForest | Base |
|------------------------------|--------|--------|---------|------|
| 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 |

**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.
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.

| Cluster | Gender | Socio-demographics | Ideology |
|---------|--------|---------------------|----------|
| Violet (-0.457) | M: 64.1% | ≤ 30: 58.7% | Con: 52% |
| | F: 35.9% | ≤ 45: 19.7% | Mod: 0.6% |
| | | > 45: 21.6% | Lib: 47.5% |
| Green (1.05) | M: 25.5% | ≤ 30: 60.6% | Con: 25.1% |
| | F: 74.5% | ≤ 45: 21.2% | Mod: 6.6% |
| | | > 45: 18.2% | Lib: 68.3% |
| Yellow (0.635) | M: 53.5% | ≤ 30: 64.4% | Con: 44.3% |
| | F: 46.5% | ≤ 45: 22.6% | Mod: 4.4% |
| | | > 45: 13% | Lib: 55.3% |

Table 5: Comparison of predicted gender proportions

| Topic | Predicted Gender (M-F) | Regex Gender (M-F) | #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 |

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 intratopic 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 \( i \)th detected community \( c_{t,i} \), and \( \text{pred}_d(x) \) is a function yielding the distribution of
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)

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

The results in Table 6 indicate that values of separability and expansion that model the presence of an echo-chamber-reshuffling 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 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 Schnoebe- len. 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 Zerrella. 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 E 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 networked technologies. 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 Preotiuc-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 Preotiuc-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.
## Supplemental Material

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

| Topic                  | Subreddits                                                                 |
|------------------------|-----------------------------------------------------------------------------|
| Abortion               | ‘abortion’, ‘Abortiondebate’, ‘prochoice’, ‘prolife’, ‘trueprochoice’, ‘Insanepro-choice’, ‘ProLifeLibertarians’, ‘ThingsProChoicersSay’, ‘AskProchoice’, ‘insane-prolife’, ‘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’, ‘pro-gun’, ‘liberalgunowners’, ‘Firearms’ |
| Nuclear-Energy         | ‘nuclear’, ‘NuclearEnergy’, ‘NuclearPower’                                  |
| Veganism-Animalrights  | ‘AnimalRights’, ‘animalwelfare’, ‘VeganActivism’, ‘Veganism’, ‘Vegetarianism’, ‘Veganity’, ‘vegproblems’, ‘AntiVegan’, ‘DebateAVegan’, ‘debatemeataters’, ‘exvegans’ |

Table 7: The subreddits that were crawled to create the dataset from which the studied users, their posts and the interaction graphs were extracted
### A.2 Unigram Coefficients

Table 8: Gender svc-model coefficients for unigrams

|       | 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 | pleas        | -1.281| player     |
| 1.317 | friend       | -1.266| look       |
| 1.29  |              | -1.264| war        |
| 1.281 | 😊           | -1.263| problem    |
| 1.279 | thing        | -1.225| coronavirus|
| 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         |

84
| ≤ 30   | ≤ 45   | > 45   |
|--------|--------|--------|
| 1.712  | right  | 1.411  |
| 1.688  | movi   | 1.228  |
| 1.433  | mean   | 1.216  |
| 1.334  | excit  | 1.194  |
| 1.305  | tri    | 1.181  |
| 1.299  | fun    | 1.096  |
| 1.295  | or     | 1.04   |
| 1.284  | 🙏     | 1.01   |
| 1.238  | gonna  | 1.008  |
| 1.199  | life   | 1.006  |
| 1.129  | so     | 1.004  |
| 1.097  | like   | 0.989  |
| 1.078  | an     | 0.965  |
| 1.076  | class  | 0.964  |
| 1.055  | day    | 0.954  |
| 1.013  | becaus | 0.953  |
| 0.996  | 😐     | 0.925  |
| 0.977  | 😞     | 0.921  |
| 0.945  | y'all  | 0.903  |
| 0.944  | wanna  | 0.893  |
| 0.924  | :)     | 0.875  |
| 0.889  | pop    | 0.873  |
| 0.881  | <3     | 0.849  |
| 0.878  | 😂     | 0.847  |
| 0.876  | 😅     | 0.819  |
| 0.86   | 😒     | 0.816  |
| 0.85   | 😞     | 0.816  |
| 0.847  | punchcard | 0.809  |
| 0.836  | you'r  | 0.809  |
| 0.833  | can't  | 0.798  |
| 0.825  | –      | 0.798  |
| 0.825  | shop   | 0.798  |
| 0.82   | 1,000  | 0.795  |
| 0.818  | nigga  | 0.789  |
| 0.796  | dailylook | 0.789  |
| 0.796  | charact| 0.789  |
| 0.777  | no     | 0.779  |
| 0.777  | #cochlearimpl | 0.778  |
| 0.767  | berni  | 0.778  |
| 0.766  | chang  | 0.778  |

Table 9: Age group svc-model coefficients for unigrams
| Conservative      | Moderate      | Liberal      |
|-------------------|---------------|--------------|
| 1.405 polic       | 1.225 wow     | 1.347 🍊    |
| 1.154 leftist     | 1.107 back    | 1.206 omg    |
| 1.132 chat        | 1.087 by      | 1.177 prize  |
| 1.114 polit       | 1.075 😳      | 1.137 write  |
| 1.103 protest     | 1.07 money    | 1.112 save   |
| 1.046 kid         | 1.052 love    | 1.111 tweet  |
| 1.036 😱          | 0.995 stream  | 1.092 🙄     |
| 1.027 littl       | 0.97 can      | 1.044 serious|
| 1.009 they        | 0.963 🎉      | 1.025 We     |
| 1.005 it’         | 0.926 😃      | 0.987 still  |
| 1.004 blm         | 0.917 dot     | 0.977 episod |
| 0.98              | 0.909 😍      | 0.967 women  |
| 0.974 democrat     | 0.877 pack    | 0.938 so     |
| 0.972 call        | 0.857 Me      | 0.934 work   |
| 0.962 On          | 0.833 …       | 0.909 #voicesaveindia |
| 0.941 left        | 0.804 summer  | 0.903 fox    |
| 0.923 ✖           | 0.791 show    | 0.881 movi   |
| 0.904 kind        | 0.786 realli  | 0.872 chang  |
| 0.9 jesu          | 0.779 game    | 0.855 damn   |
| 0.889 state       | 0.773 time    | 0.853 pandem |
| 0.879 that'       | 0.755 play    | 0.851 spnwithlov |
| 0.876 know        | 0.749 enter   | 0.848 law    |
| 0.861 hillari      | 0.748 get     | 0.847 anyth  |
| 0.86 mani         | 0.748 #stevenunivers   | 0.844 he’    |
| 0.851 😘           | 0.744 need    | 0.84 again  |
| 0.85 school       | 0.737 when    | 0.828 right  |
| 0.845 also        | 0.733 😴       | 0.821 food   |
| 0.843 seem        | 0.731 befor   | 0.819 trump  |
| 0.842 legal       | 0.731 come    | 0.819 +      |
| 0.835 which       | 0.728 come    | 0.807 think  |
| 0.814 look        | 0.728 –>      | 0.792 today  |
| 0.81 china        | 0.724 🔥      | 0.774 l       |
| 0.808 hospit      | 0.717 school  | 0.773 white  |
| 0.794 rather      | 0.715 app     | 0.772 youtub |
| 0.787 viru         | 0.712 best    | 0.756 stay   |
| 0.774 these        | 0.696 free    | 0.753 &       |
| 0.771 down         | 0.685 coronaviru | 0.751 mean |
| 0.762 '             | 0.681 reach   | 0.749 protect |
| 0.759 own          | 0.679 awesom  | 0.748 gay     |
| 0.759 bill         | 0.666 learn   | 0.741 kat     |
| 0.756 ye           | 0.663 y       | 0.741 stori   |
| 0.743 clinton      | 0.657 reward  | 0.736 everi   |
| 0.737 illeg        | 0.65 join     | 0.727 dog     |
| 0.737 around       | 0.647 from    | 0.725 and     |
| 0.732 sens         | 0.646 $       | 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             | 0.024 | 2.024 | 10.443 | M: 0.641 | ≤ 30: 0.587 | Con: 0.52 |
|         |         | Std: 0.924         |       |      |      | F: 0.359 | ≤ 45: 0.197 | Mod: 0.006 |
|         |         | #Users: 36.0       |       |      |      |      | > 45: 0.216 | Lib: 0.475 |
| 1       | 797     | 1.185              | 0.027 | 0.687 | 15.764 | M: 0.255 | ≤ 30: 0.606 | Con: 0.251 |
|         |         | Std: 0.584         |       |      |      | F: 0.745 | ≤ 45: 0.212 | Mod: 0.066 |
|         |         | #Users: 33.0       |       |      |      |      | > 45: 0.182 | Lib: 0.683 |
| 2       | 1168    | 0.199              | 0.047 | 1.273 | 21.452 | M: 0.535 | ≤ 30: 0.644 | Con: 0.443 |
|         |         | Std: 1.051         |       |      |      | F: 0.465 | ≤ 45: 0.226 | Mod: 0.004 |
|         |         | #Users: 118.0      |       |      |      |      | > 45: 0.13  | 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.
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     | 0.018   | -0.761          | 0.018 | 0.421 | 20.289 | M: 0.824 | ≤ 30: 0.487 ≤ 45: 0.298 > 45: 0.215 Con: 0.499 Mod: 0.013 Lib: 0.488 |
| 1       | 1176    | 0.022   | -0.589          | 0.022 | 0.733 | 17.426 | M: 0.798 | ≤ 30: 0.367 ≤ 45: 0.355 > 45: 0.277 Con: 0.493 Mod: 0.019 Lib: 0.488 |
| 2       | 1168    | 0.023   | -0.604          | 0.023 | 0.757 | 17.997 | M: 0.782 | ≤ 30: 0.447 ≤ 45: 0.347 > 45: 0.206 Con: 0.427 Mod: 0.011 Lib: 0.562 |
| 3       | 331     | 0.035   | 0.538           | 0.035 | 3.78  | 1.511  | M: 0.353 | ≤ 30: 0.344 ≤ 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 | #Users | stance | weighted stance | Metrics | Sociodemographics |
|---------|--------|--------|----------------|---------|------------------|
| 0       | 1705   | ϕ: 0.472 | ϕ: 0.503 | d(c): 0.472 | Gender: M: 0.941 |
|         |        | Std: 0.572 | Std: 0.615 | Std: 0.572 | ≤ 30: 0.297 |
|         |        | #Users: 9.0 |                  |          | ≤ 45: 0.256 |
|         |        |                  |                  |          | > 45: 0.448 |
| 1       | 1574   | ϕ: 0.517 | ϕ: 0.362 | d(c): 0.517 | Gender: M: 0.95 |
|         |        | Std: 0.717 | Std: 0.636 | Std: 0.717 | ≤ 30: 0.229 |
|         |        | #Users: 12.0 |                  |          | ≤ 45: 0.276 |
|         |        |                  |                  |          | > 45: 0.496 |
| 2       | 1509   | ϕ: 0.435 | ϕ: 0.372 | d(c): 0.435 | Gender: M: 0.966 |
|         |        | Std: 0.548 | Std: 0.571 | Std: 0.548 | ≤ 30: 0.236 |
|         |        | #Users: 8.0 |                  |          | ≤ 45: 0.254 |
|         |        |                  |                  |          | > 45: 0.51 |
| 3       | 398    | ϕ: 0.333 | ϕ: 0.055 | d(c): 0.333 | Gender: M: 0.925 |
|         |        | Std: 0.333 | Std: 0.183 | Std: 0.333 | ≤ 30: 0.219 |
|         |        | #Users: 2.0 |                  |          | ≤ 45: 0.226 |
|         |        |                  |                  |          | > 45: 0.555 |

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 | #Users | stance weighted | Metrics | Sociodemographics |
|---------|--------|-----------------|---------|-------------------|
|         |        |                 |         |                   |
| 0       | 407    | $\rho$: 0.333  | $\rho$: 0.312 | Gender | 0.688 | $\leq 30$: 0.57 |
|         |        | Std: 1.247      | Std: 1.21 | F: 0.312 | $\leq 45$: 0.204 |
|         |        | #Users: 3.0     |         | $> 45$: 0.226 | Con: 0.386 |
| 1       | 248    | $\rho$: 0.843  | $\rho$: 1.067 | Age | 0.573 | $\leq 30$: 0.496 |
|         |        | Std: 1.083      | Std: 0.696 | F: 0.427 | $\leq 45$: 0.19 |
|         |        | #Users: 23.0    |         | $> 45$: 0.315 | Con: 0.298 |
| 2       | 338    | $\rho$: 1.667  | $\rho$: 1.667 | Ideology | 0.55 | $\leq 30$: 0.503 |
|         |        | Std: 0.471      | Std: 0.471 | F: 0.45 | $\leq 45$: 0.157 |
|         |        | #Users: 3.0     |         | $> 45$: 0.34 | Con: 0.337 |
| 3       | 471    | $\rho$: -1.167 | $\rho$: -1.167 | predicted socio-demographics | 0.724 | $\leq 30$: 0.473 |
|         |        | Std: 0.0        | Std: 0.0 | F: 0.276 | $\leq 45$: 0.187 |
|         |        | #Users: 1.0     |         | $> 45$: 0.34 | Con: 0.461 |

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.
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 | #Users | stance | weighted stance | d(c) | s(c) | e(c) | Gender | Age | Ideology |
|---------|--------|--------|----------------|------|------|------|--------|-----|----------|
| 0       | 1460   | ϕ: -0.73 | Std: 0.754 #Users: 27.0 |       |      |      | M: 0.942 F: 0.058 | ≤ 30: 0.14 | Con: 0.701 |
|         |        | std: -0.829 | Std: 0.555 | 0.049 | 1.343 | 26.633 | ≤ 45: 0.338 | Mod: 0.023 |
|         |        |       |       |       |       |       | > 45: 0.521 | Lib: 0.276 |
| 1       | 903    | ϕ: -0.53 | Std: 1.029 #Users: 35.0 |       |      |      | M: 0.947 F: 0.053 | ≤ 30: 0.175 | Con: 0.714 |
|         |        | std: -0.561 | Std: 1.047 | 0.049 | 0.499 | 44.104 | ≤ 45: 0.368 | Mod: 0.014 |
|         |        |       |       |       |       |       | > 45: 0.457 | Lib: 0.271 |
| 2       | 259    | ϕ: 0.0 | Std: 0.0 #Users: 0.0 |       |      |      | M: 0.919 F: 0.081 | ≤ 30: 0.154 | Con: 0.73 |
|         |        | std: 0.0 | Std: 0.0 | 0.046 | 0.17 | 35.062 | ≤ 45: 0.421 | Mod: 0.012 |
|         |        |       |       |       |       |       | > 45: 0.425 | Lib: 0.259 |
| 3       | 229    | ϕ: -0.853 | Std: 0.547 #Users: 5.0 |       |      |      | M: 0.891 F: 0.109 | ≤ 30: 0.288 | Con: 0.664 |
|         |        | std: -0.977 | Std: 0.212 | 0.028 | 0.102 | 30.825 | ≤ 45: 0.319 | Mod: 0.017 |
|         |        |       |       |       |       |       | > 45: 0.393 | Lib: 0.319 |
| Cluster | #Users | stance | weighted stance | d(e) | s(c) | e(c) | Gender | Age | Ideology |
|---------|--------|--------|-----------------|------|------|------|--------|-----|----------|
| 0       | 845    | ϕ: 0.0 | ϕ: 0.176       | 0.047| 0.813| 24.555| M: 0.931| F: 0.069| ≤ 30: 0.346 | ≤ 45: 0.452 | > 45: 0.202 | Con: 0.685 | Mod: 0.004 | Lib: 0.311 |
| 1       | 615    | ϕ: -0.295 | ϕ: -1.063     | 0.039| 0.345| 35.093| M: 0.914| F: 0.086| ≤ 30: 0.315 | ≤ 45: 0.506 | > 45: 0.179 | Con: 0.685 | Mod: 0.002 | Lib: 0.314 |
| 2       | 898    | ϕ: 0.14  | ϕ: 0.115       | 0.039| 0.684| 25.758| M: 0.93 | F: 0.07  | ≤ 30: 0.331 | ≤ 45: 0.453 | > 45: 0.216 | Con: 0.688 | Mod: 0.004 | Lib: 0.307 |
| 3       | 396    | ϕ: -0.688 | ϕ: -0.533      | 0.034| 0.903| 7.328 | M: 0.917| F: 0.083| ≤ 30: 0.338 | ≤ 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 | #Users | stance | weighted stance | d(e) | s(c) | e(e) | Gender | Age | Ideology |
|--------|--------|--------|-----------------|------|------|------|--------|-----|----------|
| 0      | 105    | \(\varphi: 1.5\) \(\text{Std: 0.5}\) #Users: 2.0 | \(\varphi: 1.211\) \(\text{Std: 0.408}\) | 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    | \(\varphi: 0.533\) \(\text{Std: 0.972}\) #Users: 10.0 | \(\varphi: 0.699\) \(\text{Std: 0.641}\) | 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    | \(\varphi: 0.7\) \(\text{Std: 0.4}\) #Users: 5.0 | \(\varphi: 0.76\) \(\text{Std: 0.425}\) | 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    | \(\varphi: 1.016\) \(\text{Std: 0.778}\) #Users: 22.0 | \(\varphi: 1.146\) \(\text{Std: 0.443}\) | 0.111 | 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.

94