



FEBRUARY, 2024

ALARMISM VS ACTIVISM

MAPPING CLIMATE CHANGE
DISCOURSE IN SOCIAL NETWORKS

INTRODUCTION

In today's digital landscape, climate change discourse has shifted to social media platforms, challenging traditional media's role. Our project, "Alarmism or Activism: Mapping Climate Change Discourse in Social Networks," aims to analyze discourse patterns surrounding climate change, distinguishing between activist and alarmist perspectives. Through social network analysis, we had mapped interactions and linguistic dynamics within online communities. Understanding these dynamics is crucial due to the urgent need to address climate change and the significant impact of digital conversations on public awareness and action.

RESEARCH QUESTION

"What are the prevalent discourse patterns surrounding the theme of 'climate change' in social media networks, distinguishing between instances where the discussion serves an activist perspective advocating for action and instances where it propagates excessive alarmism?"



THEORETICAL FRAMEWORKS

- **Social Network Analysis (SNA)**

This framework allows us to examine the relationships between actors (users, groups, organizations) in a social network. We can map interactions and analyze the structure of connections between users discussing "climate change" on social media platforms. SNA provides insights into the flow of information, influential nodes, and community structures.

THEORETICAL FRAMEWORKS

- **Media Framing Theory**

This theory investigates how media (including social media) shapes public opinion by framing issues in particular ways. In our case, it helps identify how "climate change" discourse is framed—whether as a tangible, imminent risk needing action or as alarmist rhetoric—by analyzing language, tone, and the context in which the term is used in social media discussions

METHODOLOGY

1

Data Collection
and cleaning

2

Initial analysis

3

Topic
detection

4

Sentiment
analysis

DATA COLLECTION AND NETWORK DEFINITION

Data acquisition

Data downloaded from Reddit through Reddit API. Almost 4000 posts and more than 80000 comments published during the last year

Network definition

A bipartite graph having words and documents as nodes is created. In case of posts only the title is kept

DOCUMENT CLEANING

Documents with null or removed body are discarded. Three networks have been defined:

POSTS

Only posts are kept

POSTS AND BEST COMMENTS

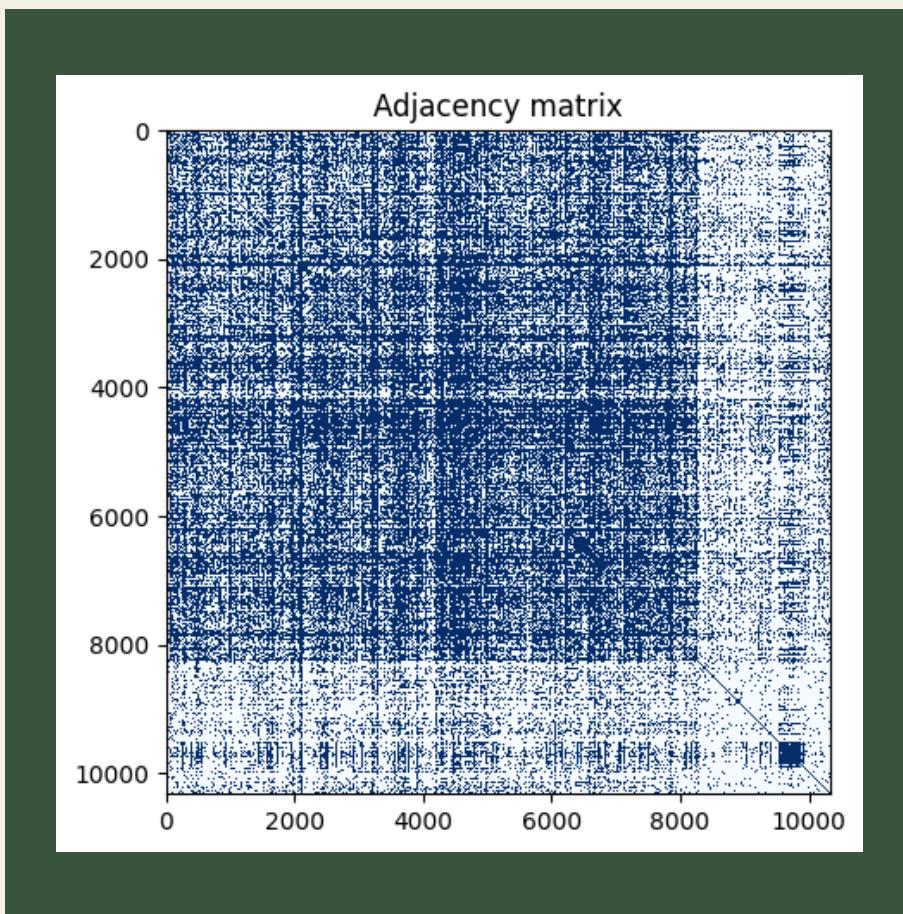
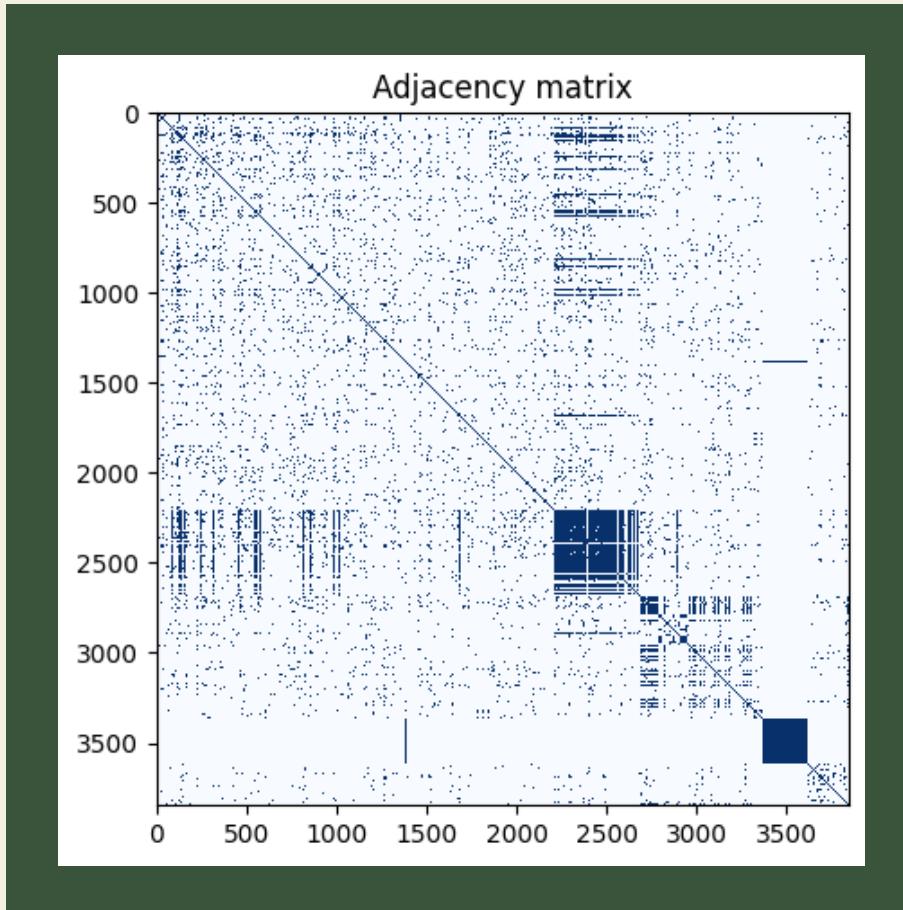
All posts are kept while comments are filtered based on their length and ups

ALL

All documents are kept

INITIAL ANALYSIS

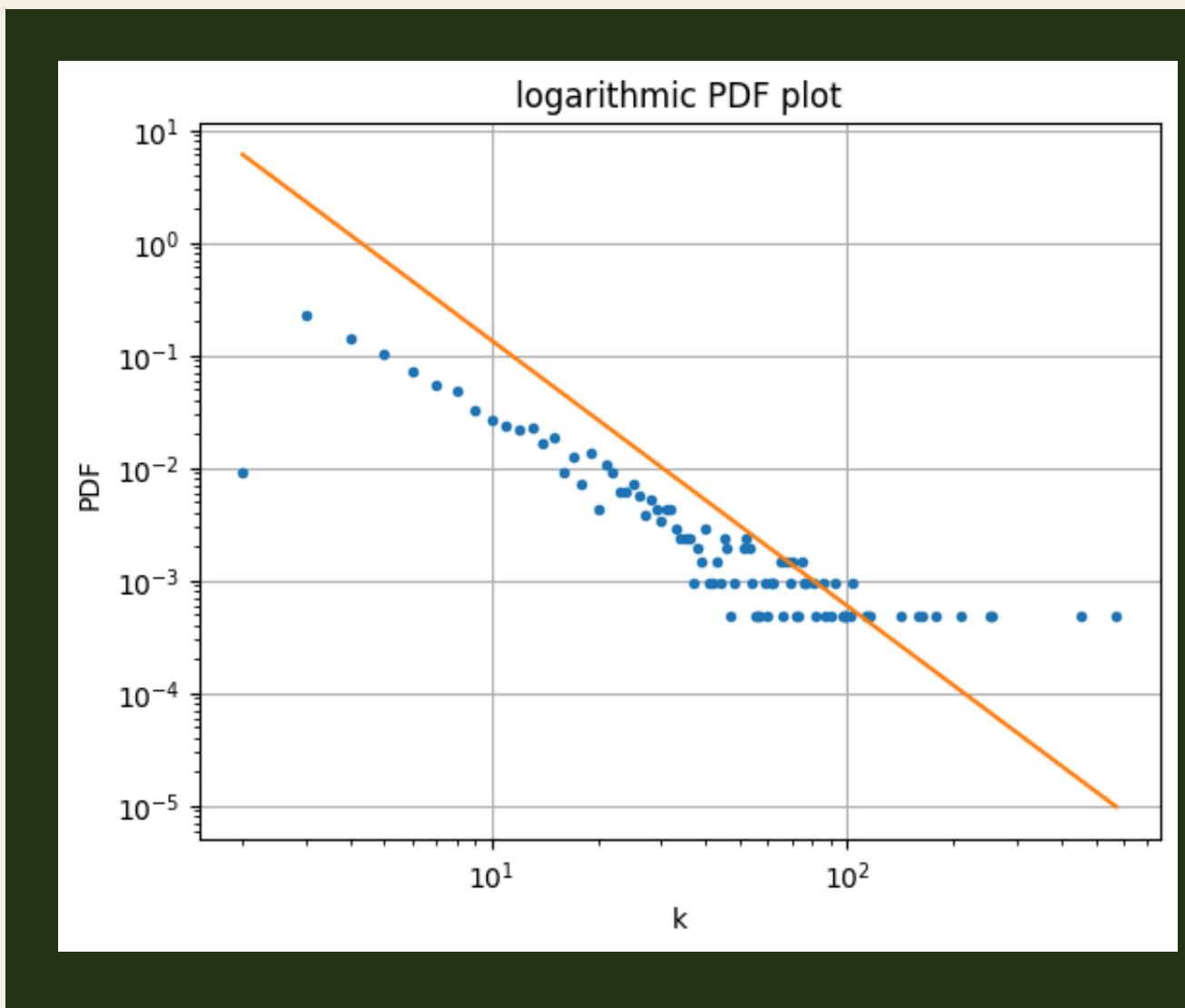
PROJECTION ON WORDS



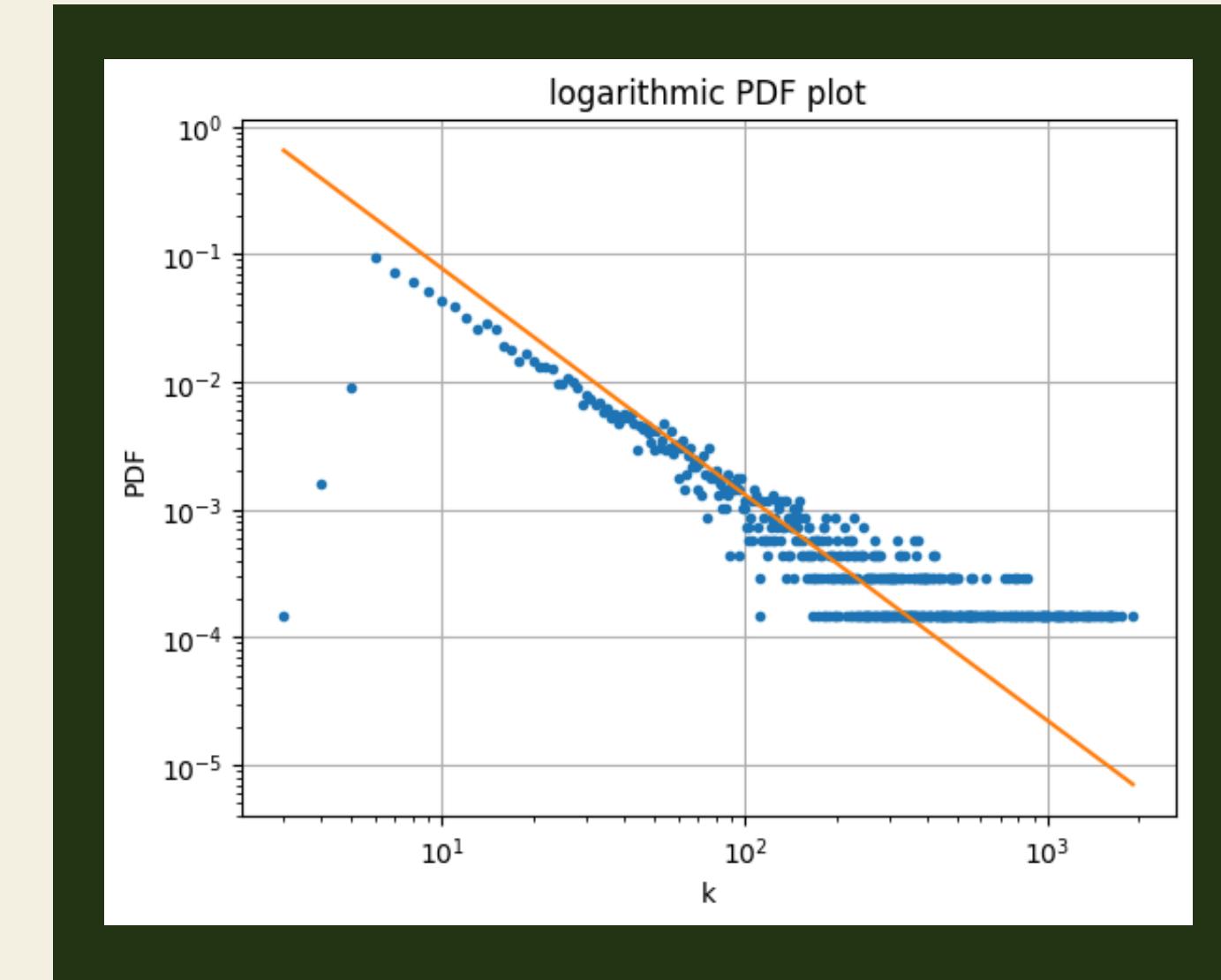
	POSTS	POSTS AND COMMENTS
NUMBER OF NODES	77182	6868
NUMBER OF EDGES	2082	3855186
AVERAGE SHORTEST PATH	2.13	1.85
DIAMETER	4	3

SCALE FREE PROPERTY

Both the networks are heavy-tailed



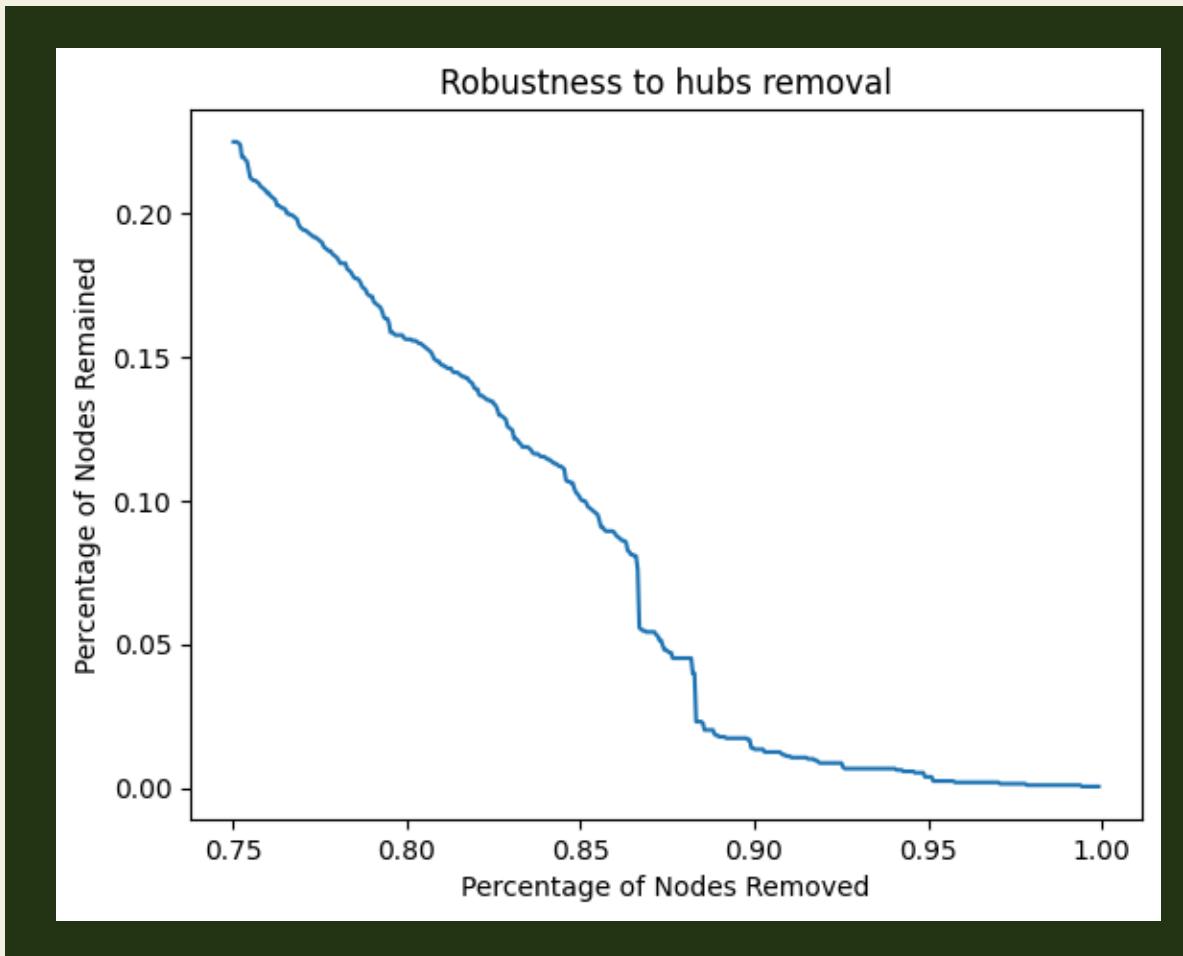
Estimated γ : 2.32
SCALE FREE REGIME



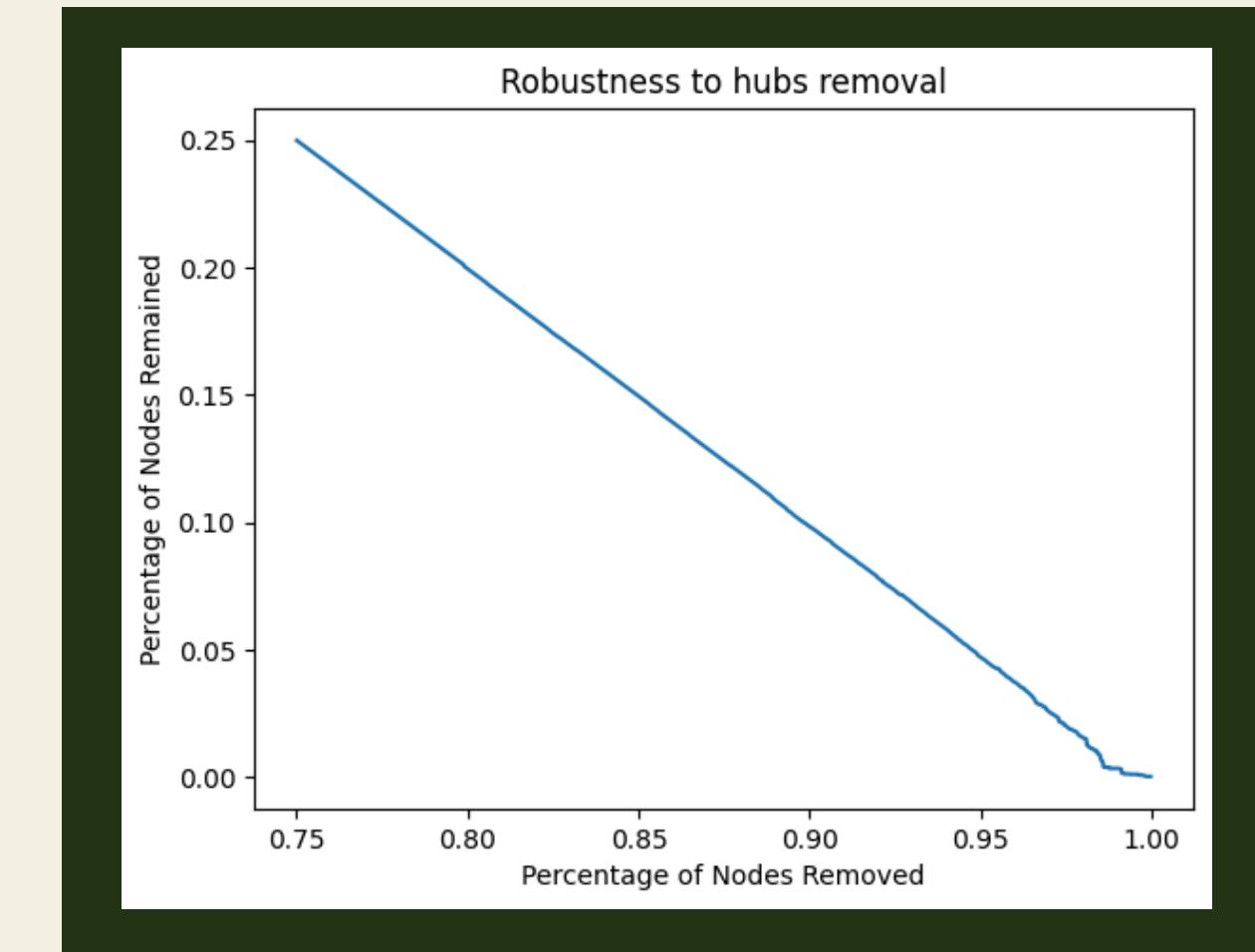
Estimated γ : 1.77
ANOMALOUS REGIME

ROBUSTNESS

Remove nodes starting from the ones with higher degree. Breaking point is very high suggesting that the networks are very connected



Posts network: almost 85% of nodes need to be removed to break the network



Posts and comments network: almost 95% of nodes need to be removed to break the network



TOPIC DETECTION

Communities identified on the projection on documents.

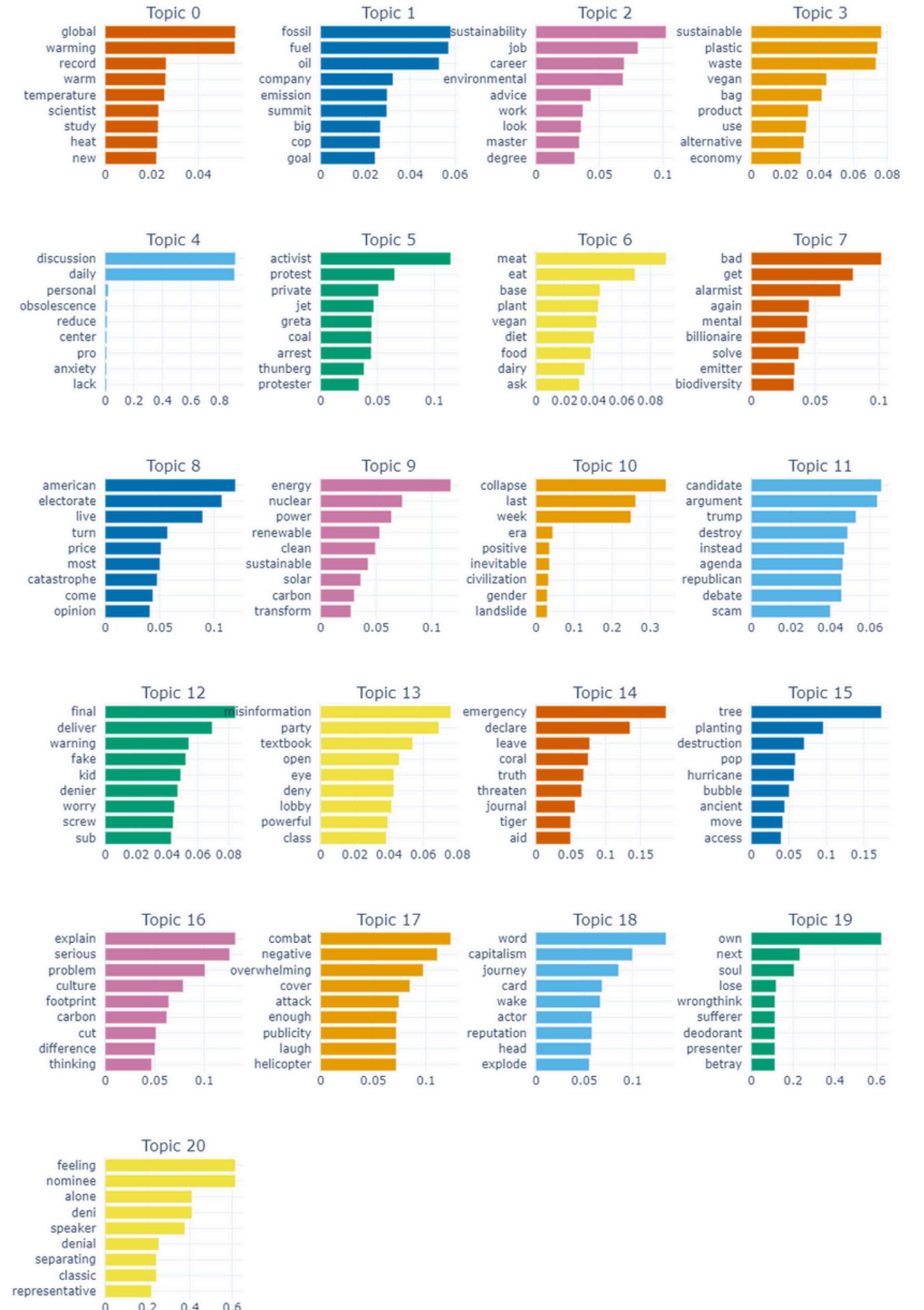
Three different algorithms:

Louvain algorithm

BERTopic

bigCLAM

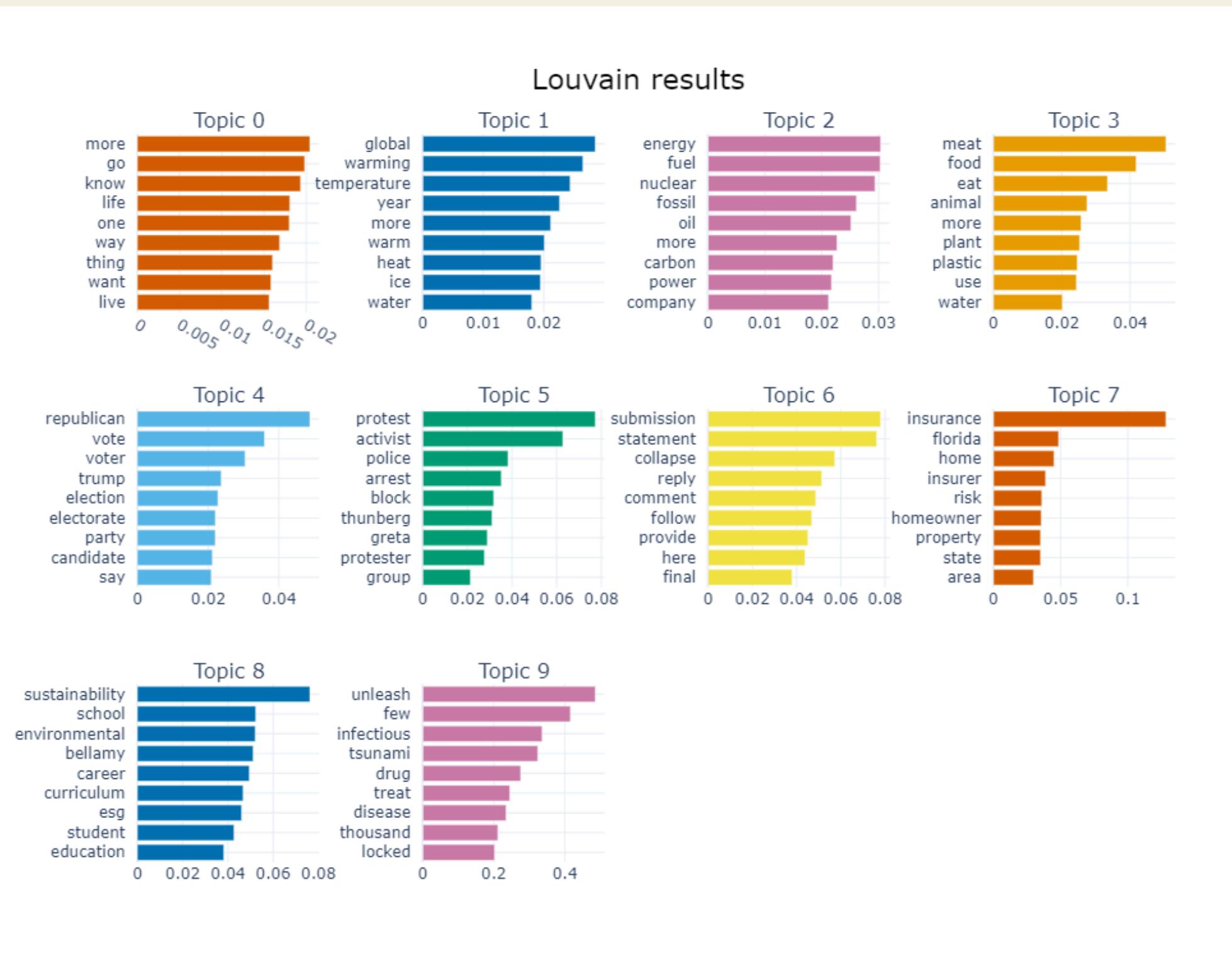
Louvain results



RESULTS ON POSTS

- Found between 15 and 25 communities through Louvain
- Most of the communities are very small and specific (most of the documents belong to the first four topics)
- Typically there are local communities: climate is local and related to national politics

LOUVAIN ON POSTS AND COMMENTS



- Many attempts to have a good community assignment: uncleaned data don't have good results
- Found between 7 and 12 communities
- First community is very big and generic because nodes are too connected

COMPARISON

Conclusion: Louvain works better than BerTopic and the dataset of post has better-defined topics

	LOUVAIN ON POSTS	BERTOPIC ON POSTS	LOUVAIN ON POSTS AND COMMENTS
NMI	0.57	0.53	0.25
NCUT	0.41	0.77	0.76
MODULARITY	0.54	0.23	0.15
INFOMAP	-0.017	0.04	0.06

BIGCLAM

- Using BigCLAM algorithm in case of **overlapping** and **non-overlapping** communities to compare with *Louvain* and *BerTopic*.
- Defined **15 fixed communities** to adapt to the number of the other algorithms.

RESULTS:

non-overlapping case

- Occurrency matrix built keeping only the *maximum* probability for each column (document).
- One community for one document.

VISUALIZATION OF RESULTS

Visualization of words in those 15 communities keeping **all the words** (*w4c.json*) contained in each document.

some communities were
relevant

```
"c8": [  
    "alarmism NOUN",  
    "blow VERB",  
    "datum NOUN",  
    "gape VERB",  
    "hole NOUN",  
    "nasa PROPN",  
    "new PROPN"  
,
```

```
"c14": [  
    "biden PROPN",  
    "degree NOUN",  
    "next ADJ",  
    "nuclear ADJ",  
    "say VERB",  
    "scary ADJ",  
    "topping ADJ",  
    "war NOUN",  
    "year NOUN"  
,
```

```
"c1": [  
    "surprised ADJ"  
,
```

```
"c7": [  
    "world NOUN"  
,
```

(*w4c.json*)

RESULTS:

overlapping case

- Occurrency matrix built keeping only probabilities higher than 95% *of max probability.*
- More than one community for each document.

VISUALIZATION OF RESULTS

Visualization of words in those 15 communities keeping all the words contained in each document.

```
"c5": [  
    "comment NOUN",  
    "delusional ADJ",  
    "detail NOUN",  
    "economic ADJ",  
    "environmental ADJ",  
    "europe PROPN",  
    "extremist ADJ",  
    "growth NOUN",  
    "more ADV",  
    "new ADJ",  
    "policy NOUN",  
    "report NOUN",  
    "see VERB",  
    "states PROPN",  
    "threaten VERB",  
    "united PROPN",  
    "well ADV"  
,
```

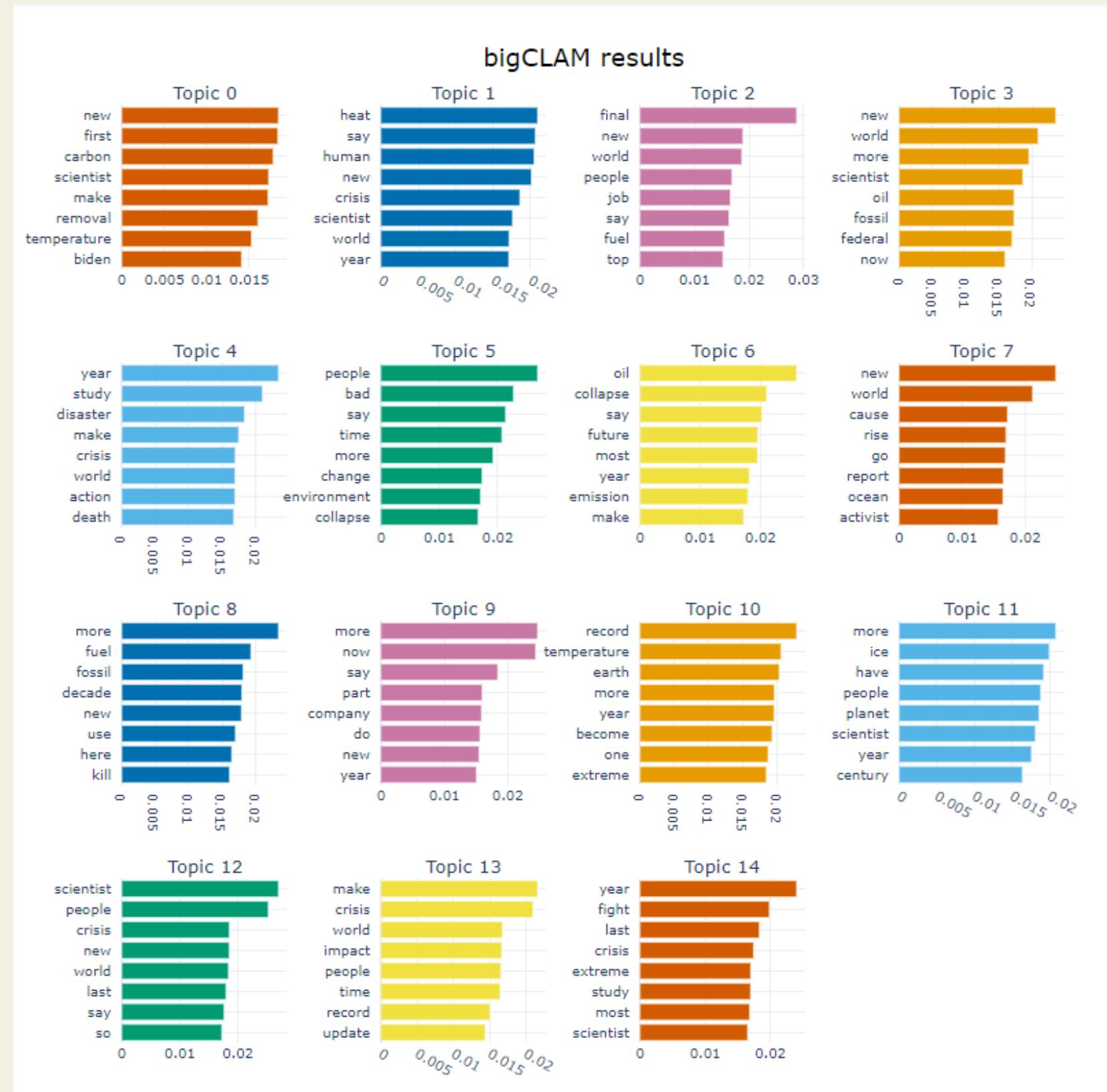
some
communities
were relevant

some others were not

```
"c12": [  
    "real ADJ",  
    "time NOUN"  
,
```

(w4c_overlapping.json)

VISUALIZATION USING BERTOPIC TOOLS



- Communities less relevant than the ones found in Louvain and BerTopic.
- Still more relevant than keeping all the words for the documents.

COMPARISON OF METRICS

	LOUVAIN	BERTOPIC	BigCLAM	BigCLAM overlapping
NMI	0.53	0.51	0.03	0.015
NCUT	0.61	0.74	0.81	0.85
MODULARITY	0.34	0.21	0.12	0.08
INFOMAP	0.0056	0.04	0.10	0.17

This proves that Louvain (and Bertopic) is better than BigCLAM

SENTIMENT ANALYSIS

BERTAGENT

- Compute agency of documents
- Define per-topic agency
- Echo chamber effect

ROBUSTNESS

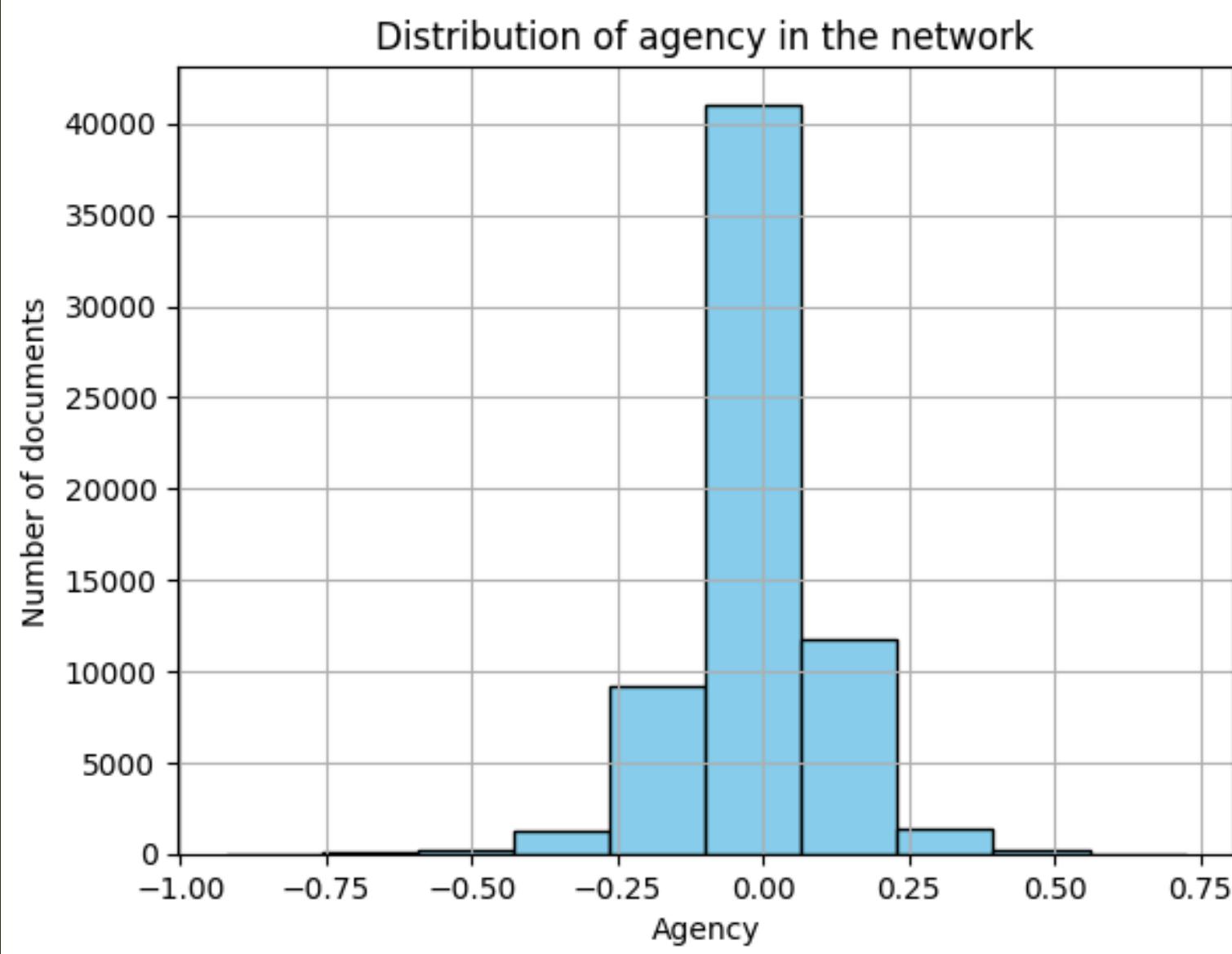
- Robustness on the network to positive and negative nodes removal
- Robustness of most relevant communities to negative words removal

LIWC

- Compute sentiments of documents
- Define the average value per-topic
- Polarization

All computations related to sentiment analysis have been done on the network of posts

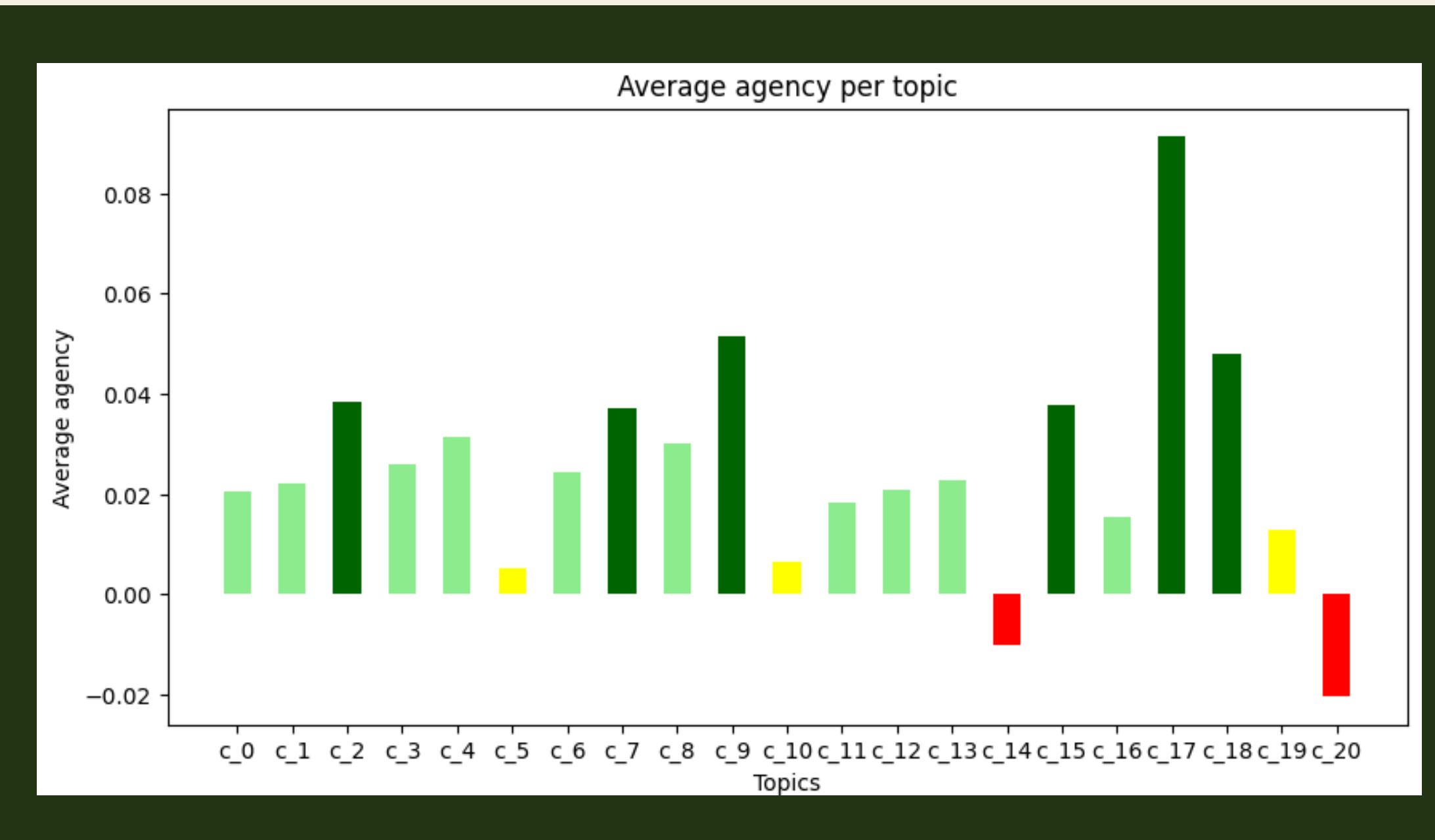
BERTAGENT



- BertAgent: Detect linguistic agency in text using large language model (pretrained transformers architecture)
- Agency value associated to each document in the network of posts
- Normalized histogram of agency values (peak in the origin)

PER-COMMUNITY AGENCY

Conclusion: There are not significant differences in terms of agency but the result is coherent with the correspondent community assignment



Community 10:

Climate Shocks Are Making Parts of America Uninsurable. It Just Got Worse.
Last Week in Collapse: March 19-25, 2023
Humans Are Now Functionally Extinct
Mountains are collapsing

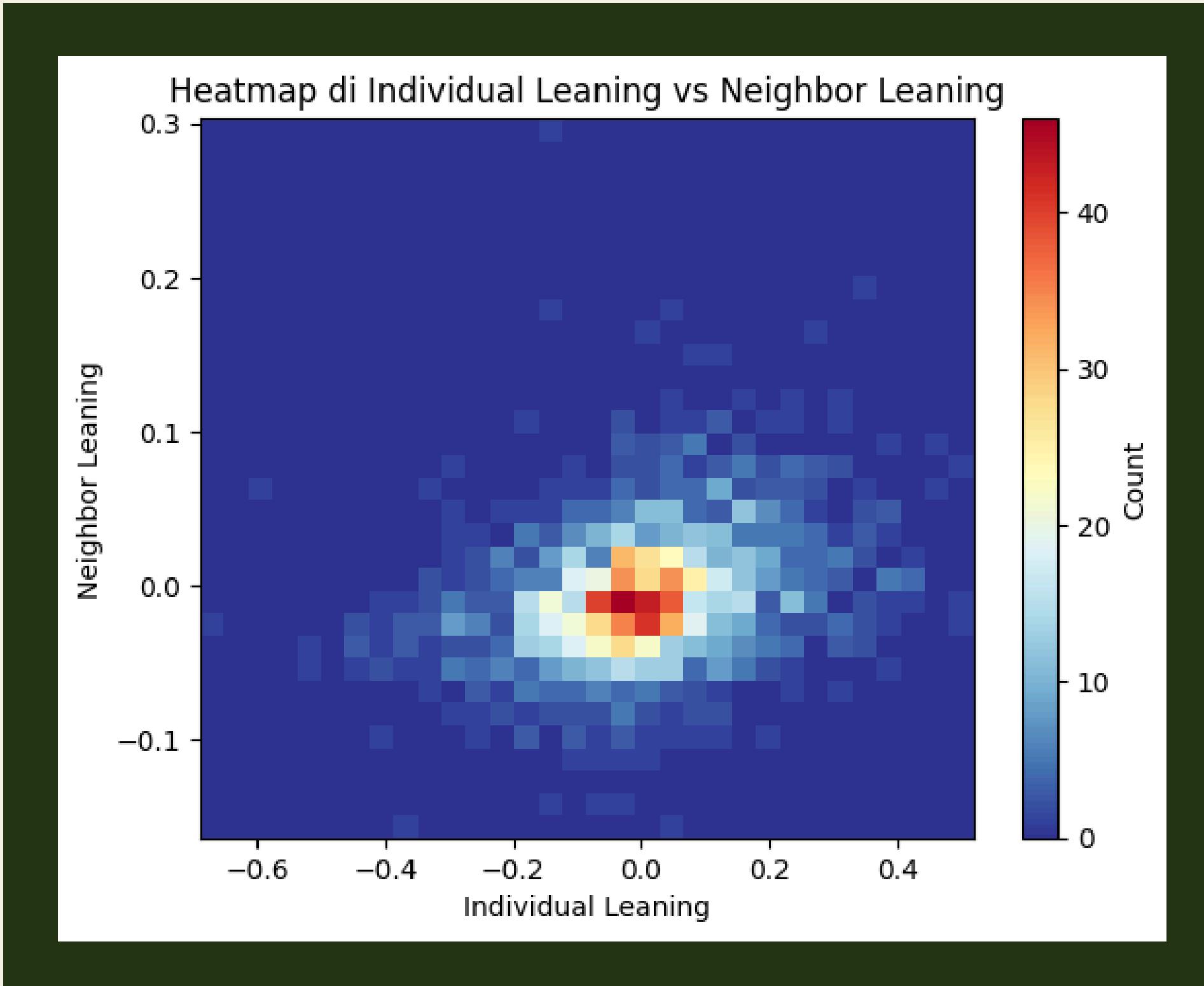
Community 17:

Scientists deliver 'final warning' on climate crisis: act now or it's too late
U.S. oil production hits all-time high, conflicting with efforts to curb climate change
Global Citizen launches 'power our planet' action program on climate change

Community 20:

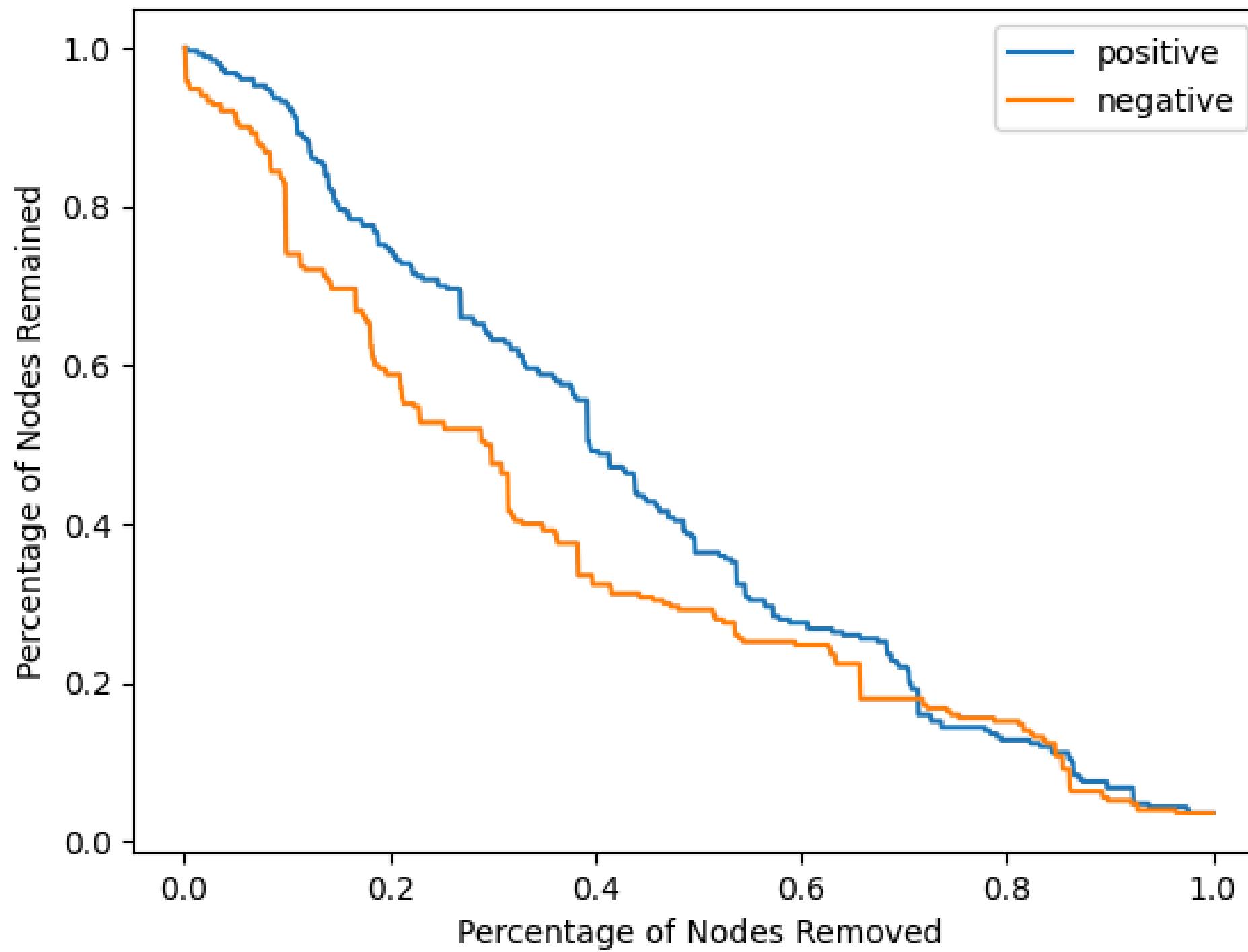
Just Stop Ambulances for Climate Change
Climate Change Denier? You're Not Alone
Just one of the many psychopathic climate change delusions

ECHO CHAMBER EFFECT



- Neighbors of a post = comments related to the post
- Single cluster centered in the origin, the network does not show echo chamber effect

Robustness to Positive and Negative words removal



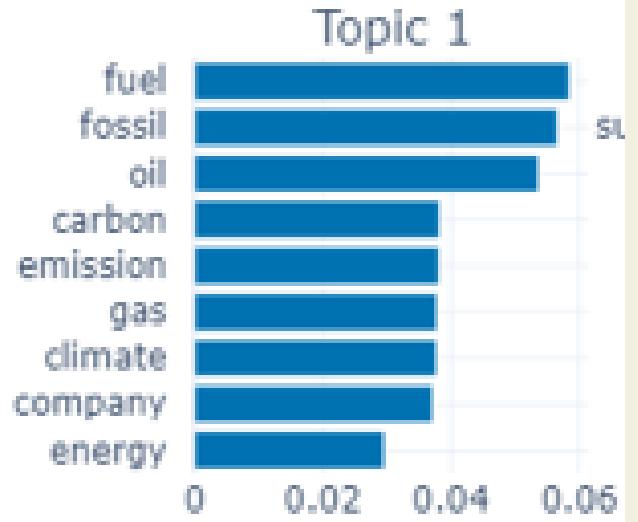
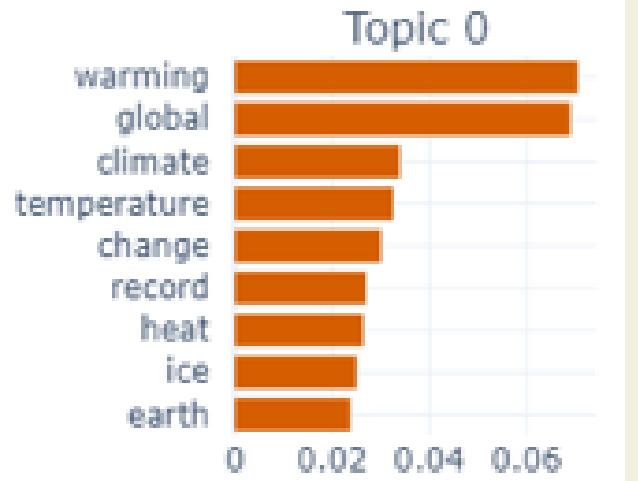
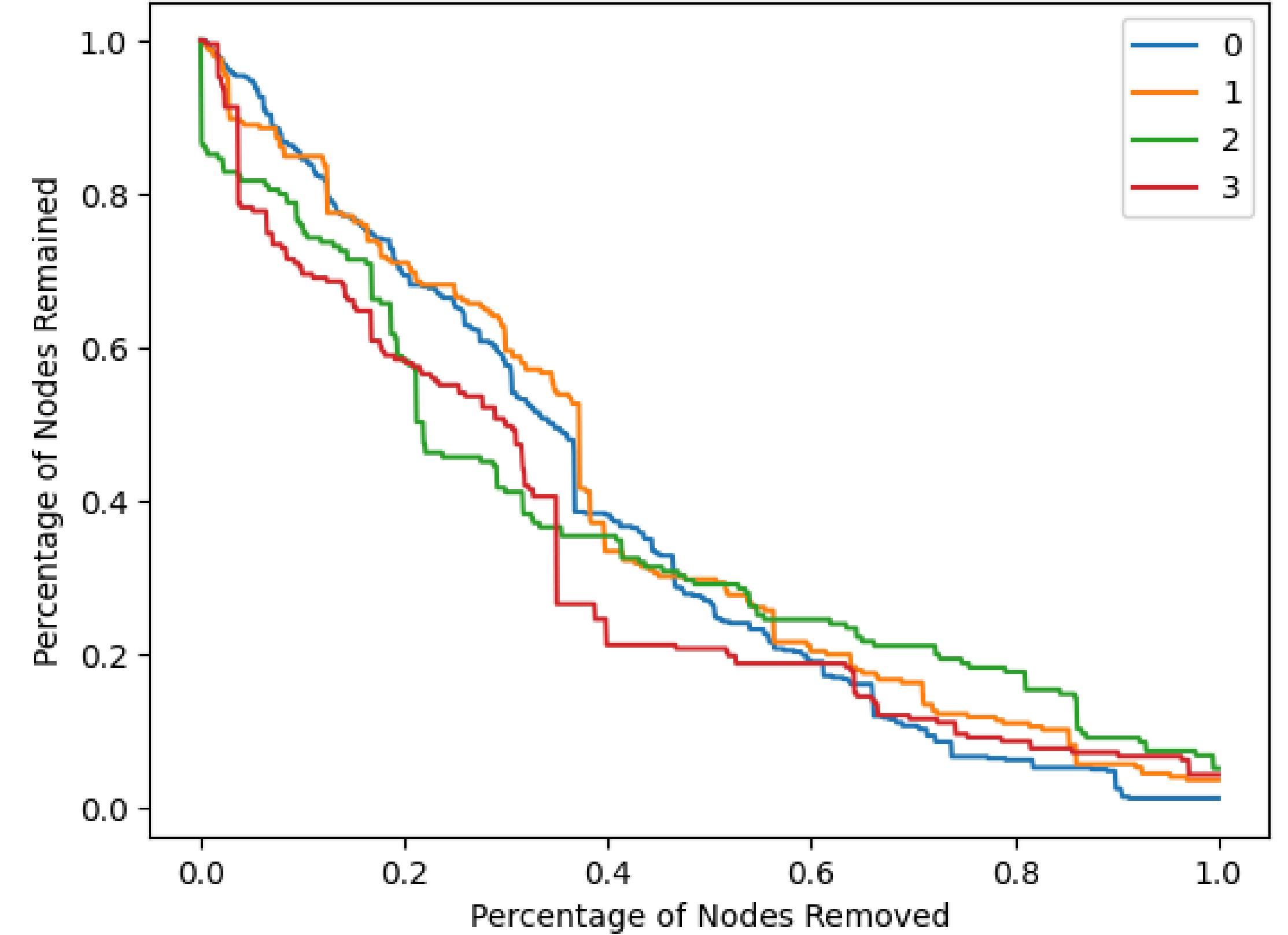
Result
NEGATIVE NETWORK

ROBUSTNESS

- Sentiment performed through the NLTK implementation of SentiWordNet
- Neutral words are removed
- Only strongest edges are kept

PER-COMMUNITY ROBUSTNESS

Robustness to Positive words removal for the most relevant communities



LIWC

- Analyzes text and examines the emotional contents of a document based on a set of linguistic and psychological markers.
- It counts the occurrences of words that fall into specific categories. E.g. “happy”, “joyful” are counted when measuring positive emotions

MARKERS

A	B	C	posemo	negemo	anx
id	subreddit	title	0	0	
j7uvvzp	worldnews	Some comments are	1.82	5.45	
j7ugj2i	worldnews	This kind of account:	0	0	
		Good.			
j7ubqhy	worldnews	People in power nee	5	0	
		Oh wow, someone is			
j7ue0yf	worldnews	edit: clearly I got my	9.09	4.55	
j7uh5a1	worldnews	This should be new c	1.41	5.63	
		It might sound good			
		It's standard languag			
j7uvw6r	worldnews	Even if they are succ	3.92	1.96	
		Sure, this is great, bu			
j7uhicp	worldnews	Forget a lawsuit - we	7.02	1.75	



Quantitative measure for each marker from 0 to 100

MARKERS

From the 93 markers provided
only 23 were considered
(regarding emotions, sense of
time and of community, tone)

Other 6 markers were defined in
order to account for 2 opposing
factors at the same time:

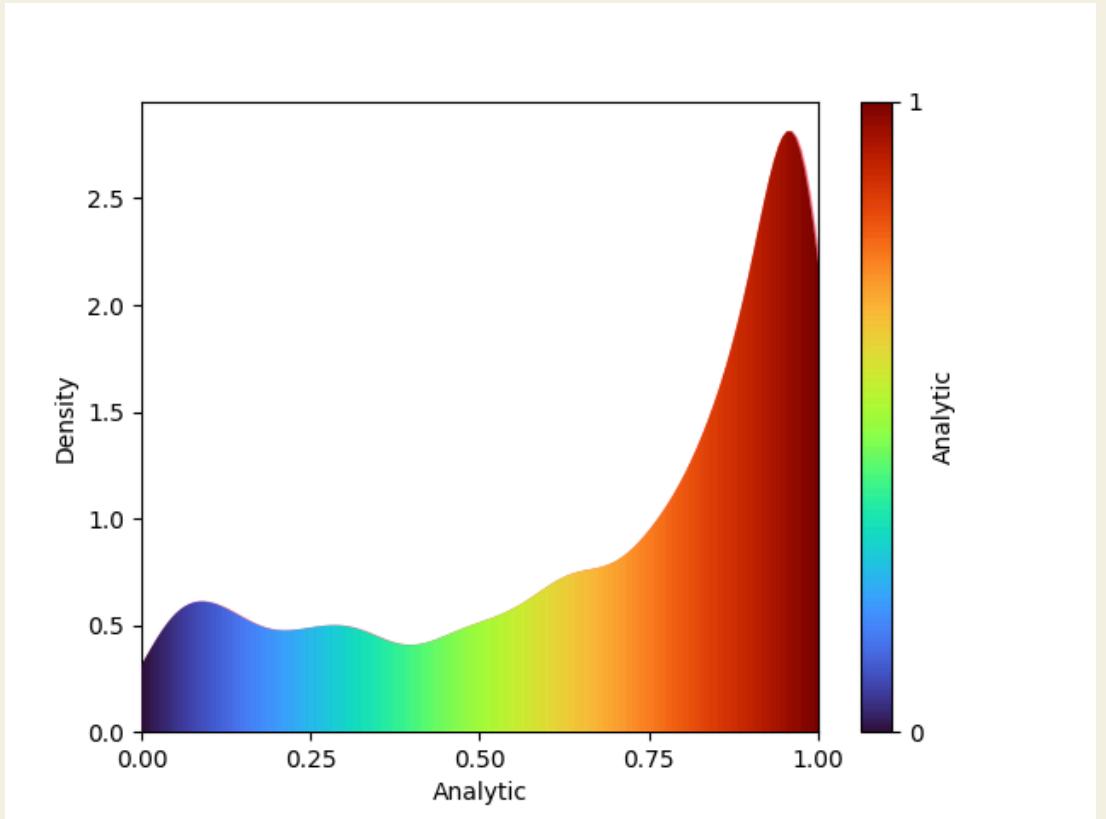
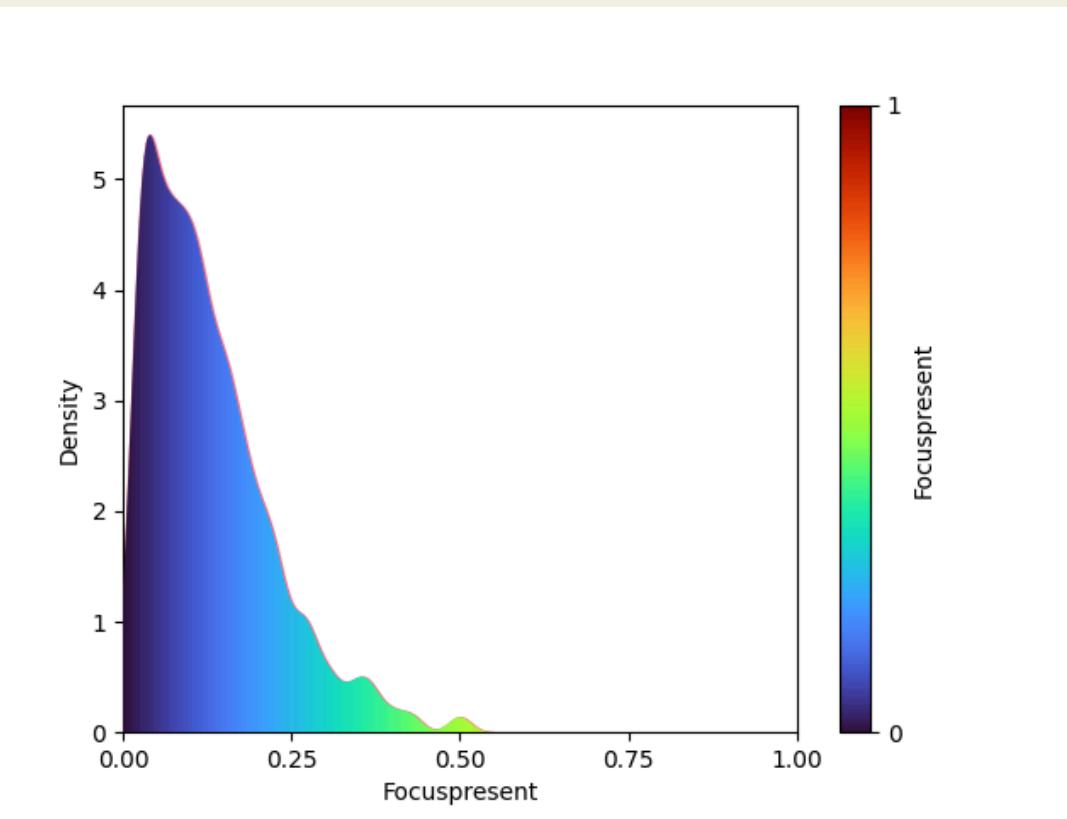
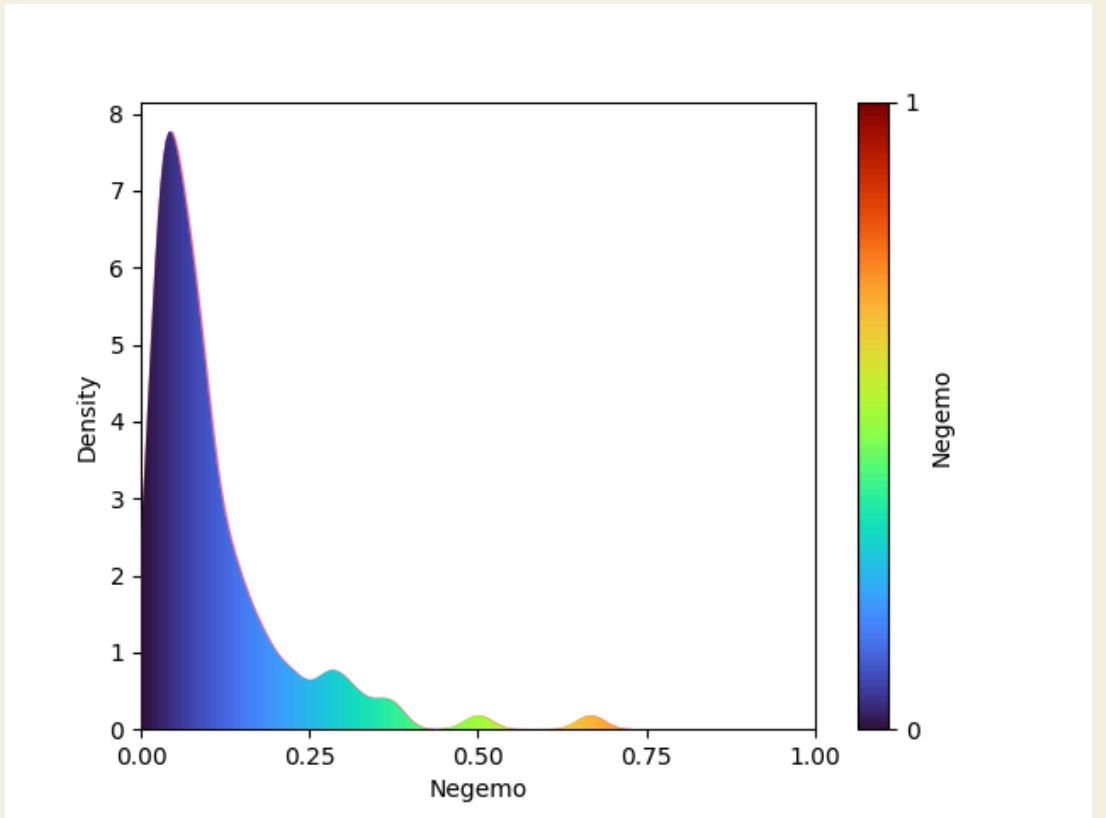
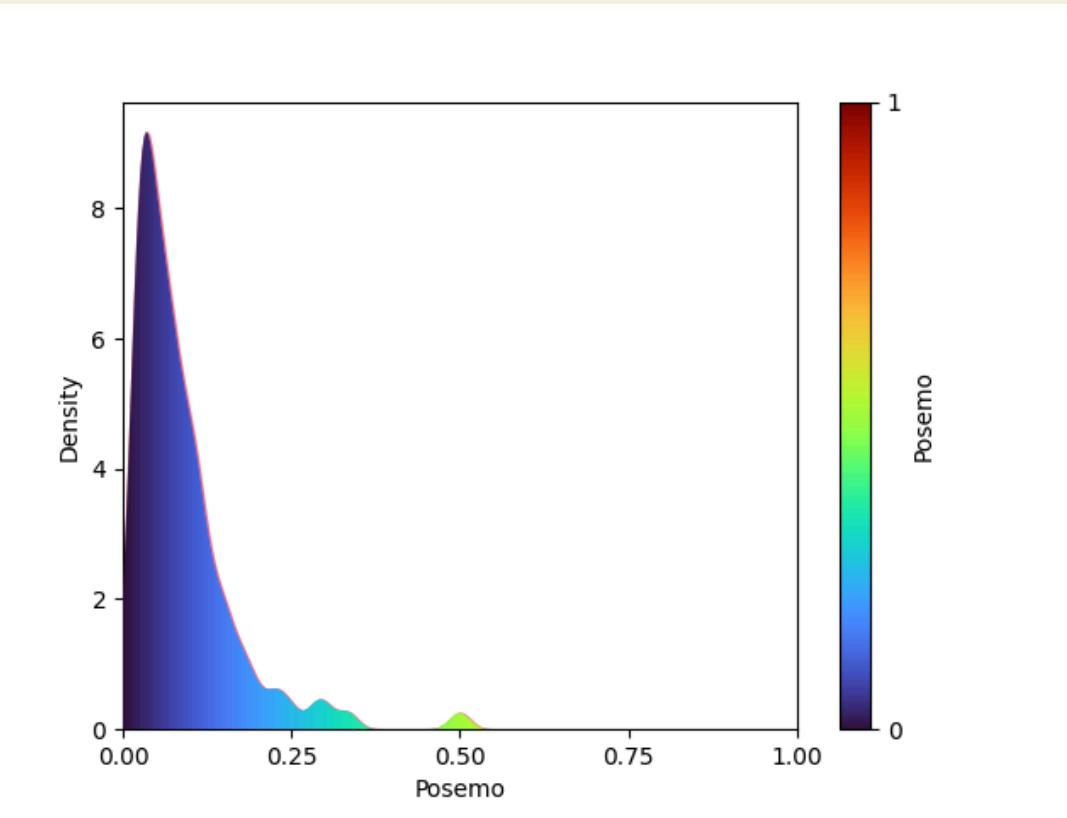
positive emotions ←
negative emotions ←

```
[ 'Analytic',  
  'Clout',  
  'Authentic',  
  'Tone',  
  'affiliation',  
  'posemo',  
  'negemo',  
  'anx',  
  'anger',  
  'sad',  
  'swear',  
  'family',  
  'friend',  
  'social',  
  'focuspresent',
```

```
df['emotions']=(df['posemo']-df['negemo'])/100  
df['female_v_male']=(df['male']-df['female'])/100  
df['leisure_v_work']=(df['work']-df['leisure'])/100  
df['risk_v_reward']=(-df['risk']+df['reward'])/100  
df['us_v_them']=(df['shehe']+df['they']-df['i']-df['affiliation'])/200  
df['past_v_present']=(df['focuspresent']-df['focuspast'])/100
```

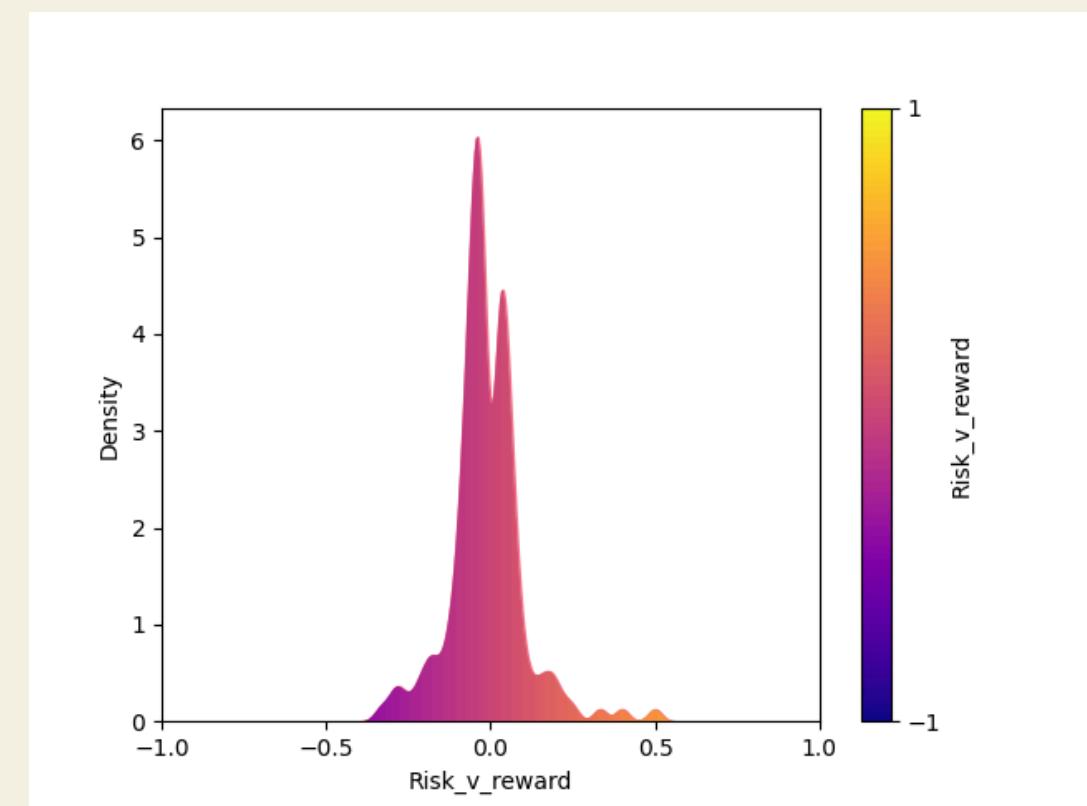
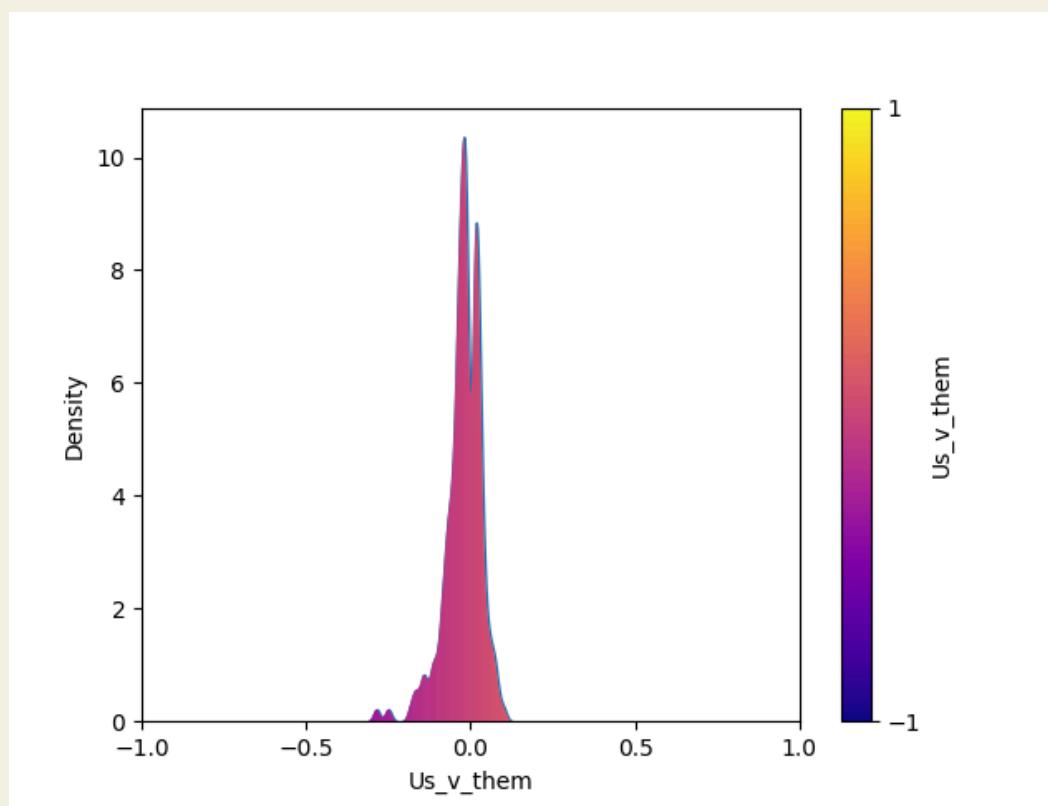
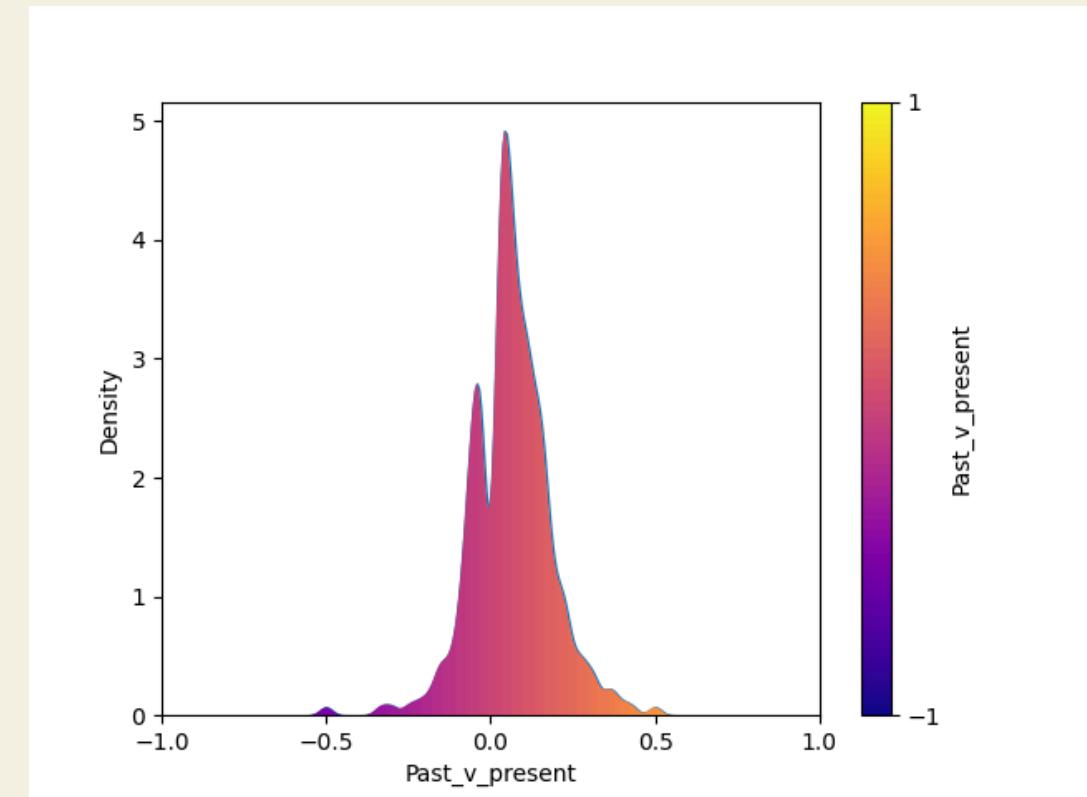
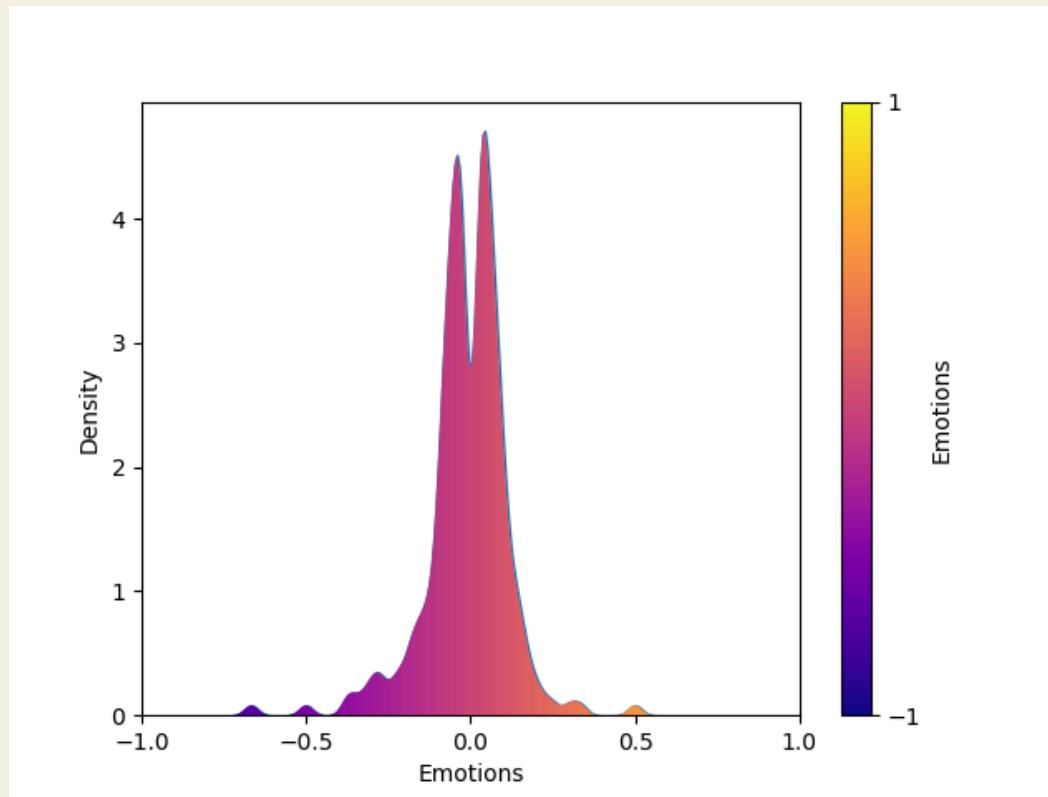
PLOTS

- Density plots for the markers (on the entire dataset):



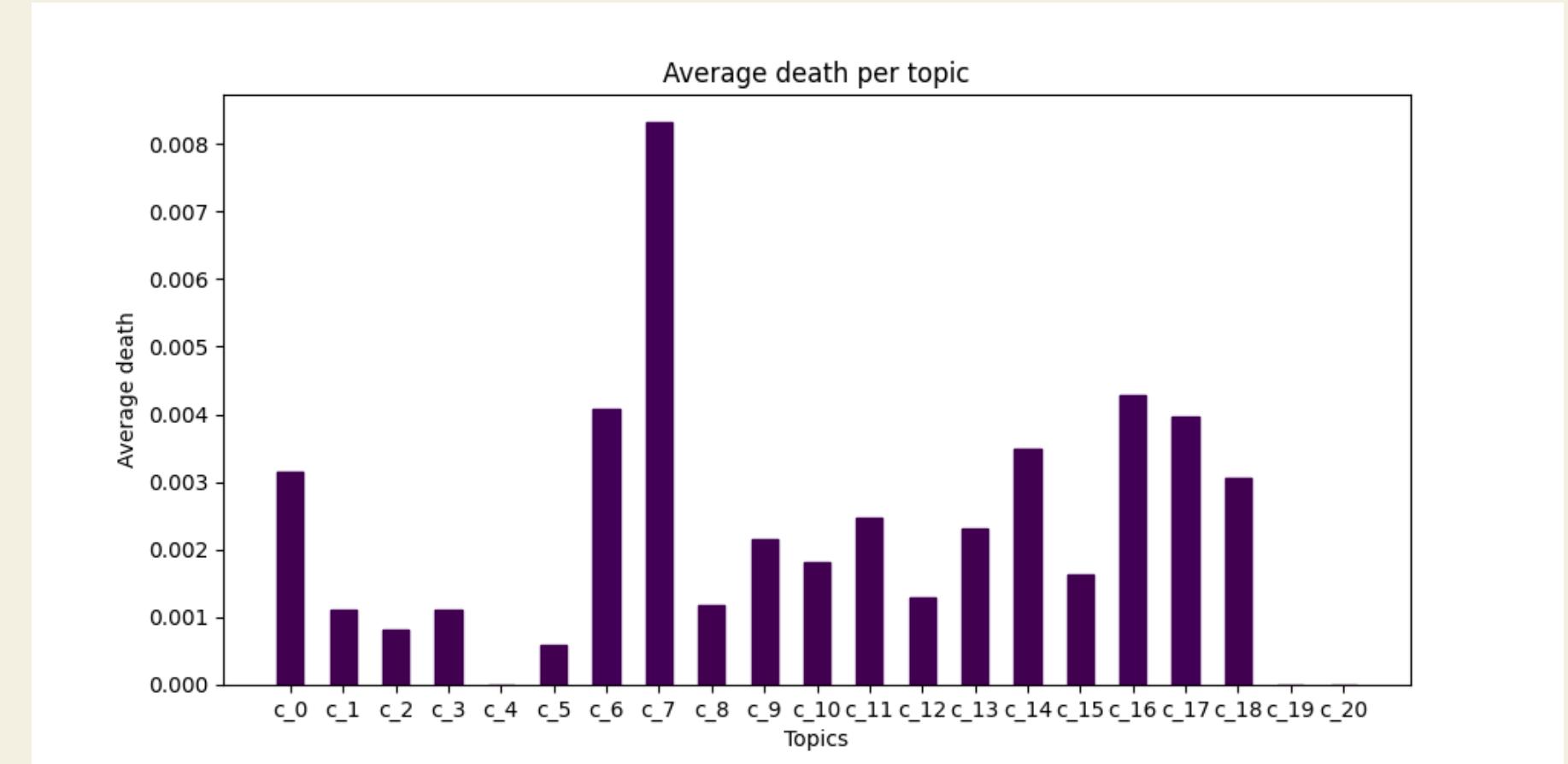
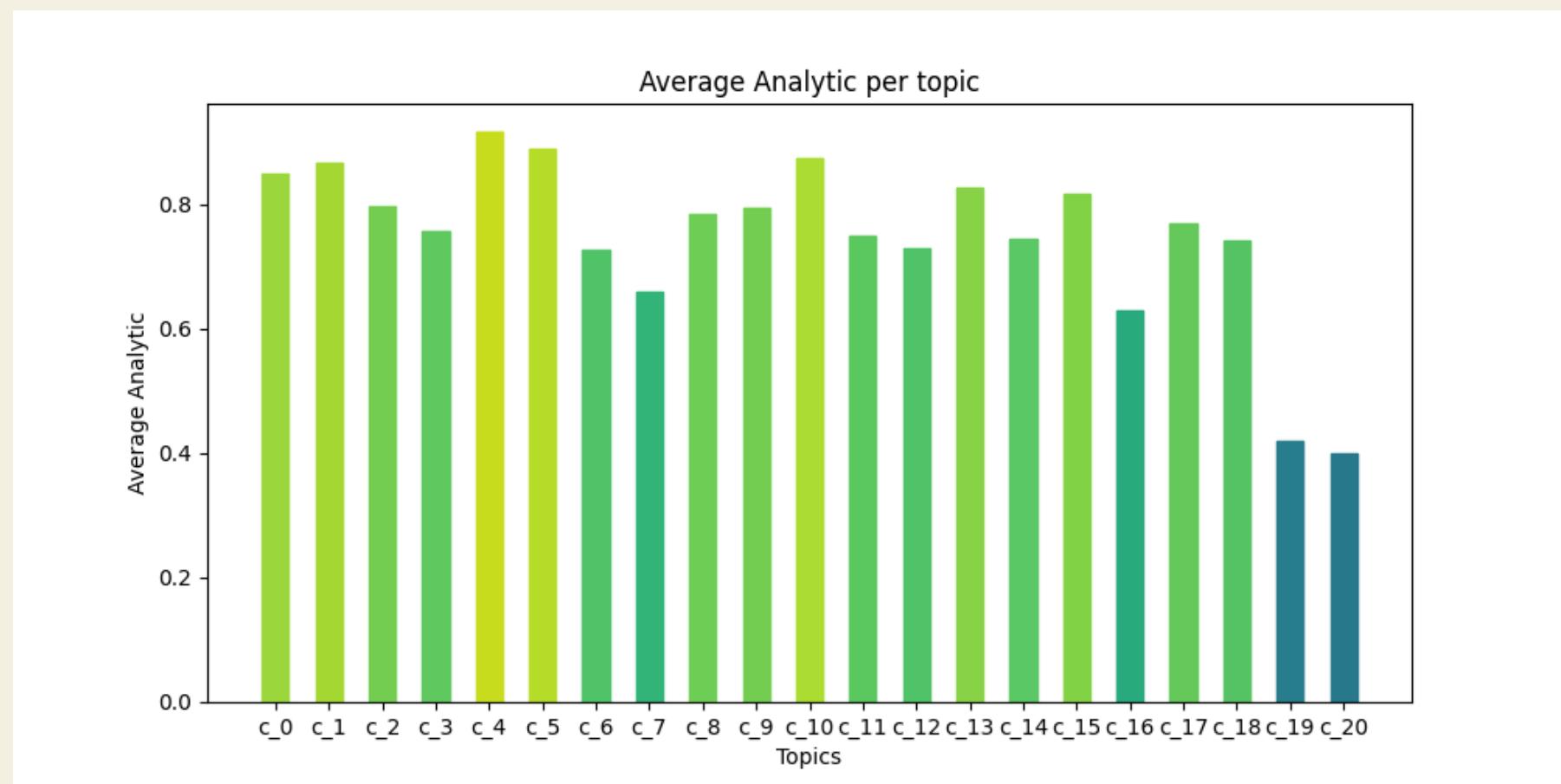
PLOTS

- Polarization plots for the markers (on the entire dataset):



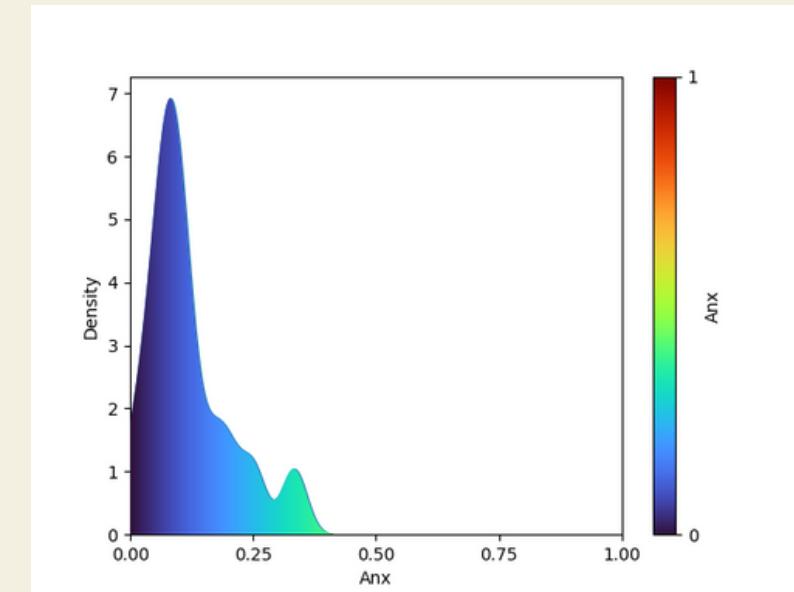
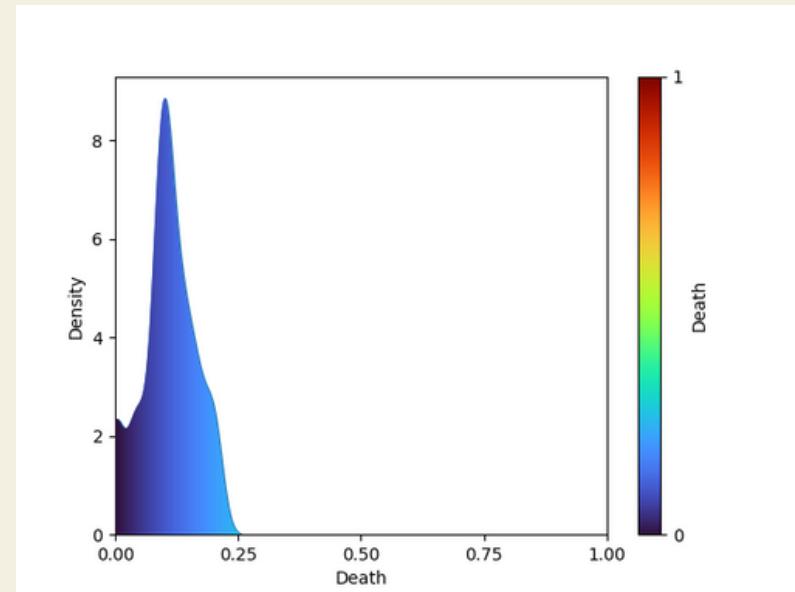
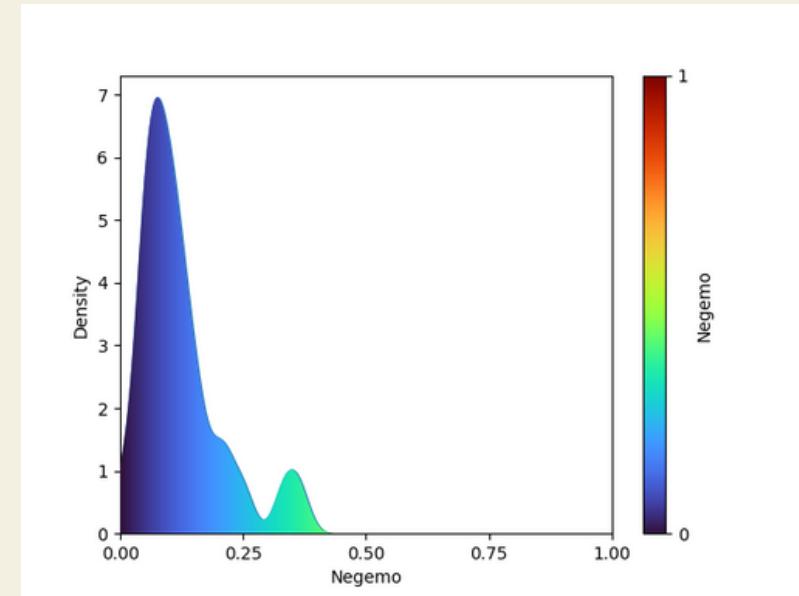
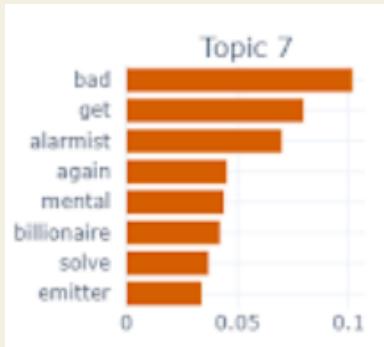
PLOTS

- Averages of markers (per topic)

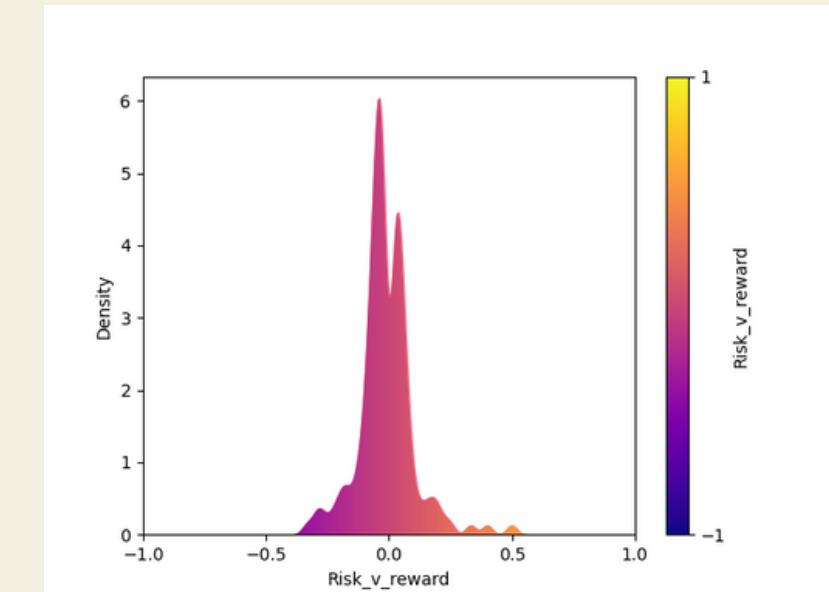


PLOTS

- Density and polarization plots for the markers (per topic):

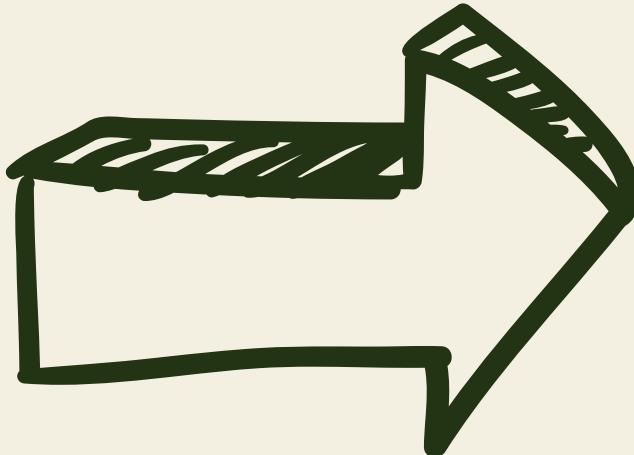


Statistics show that awareness around climate ...
Diseases Are Bred on All Animal Farms, Even th...
How are you beating climate anxiety?
Malthus was right - we are overpopulated. Unti...
Psychedelics and collapse awareness
“Ah Shit, Here We Go Again!” - A Casual Critiq...
Neoliberalism: It's Bad History, It's Bad Scie...
Just some thoughts I had this morning on the s...
The reality of the war in Ukraine.



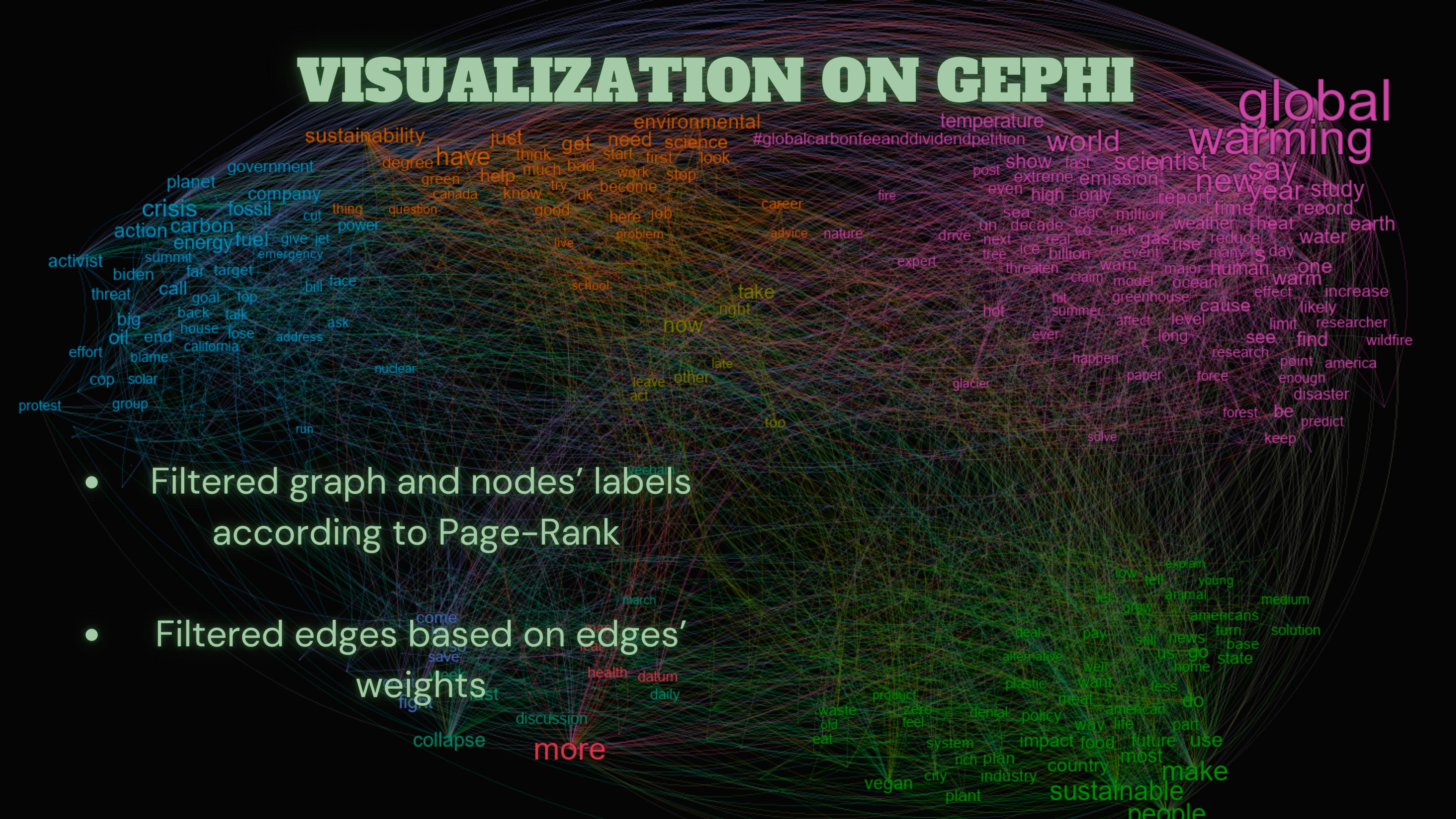
FROM PYTHON TO GEPHI

- Modified the occurrency matrix (just posts) by summing rows corresponding to the same word (e.g. "work NOUN" and "work VERB")
- Used the resulting P_{ww} as reference adjacency matrix instead of P_{dd} (better visualization)
- Degree of node = sum of entries of row in P_{ww}
- Edge weight = entry of P_{ww}



Ran the Louvain algorithm and found just 8 communities but relevant enough to the most important topics of P_{dd}

VISUALIZATION ON GEPHI



- Filtered graph and nodes' labels according to Page-Rank

- Filtered edges based on edges' weights



DISCUSSION ON THE RESULTS

HIGHLY CONNECTED NETWORK

- The network analyzed is highly connected, indicating a significant exchange of information and opinions on climate change within the online community.
- Despite many attempts to reduce connectivity by excluding common words like "climate" and "changes," the network remains highly connected.
- The high connectivity suggests a widespread awareness and engagement among social media users regarding the urgency of the climate change debate.
- It may reflect a growing public interest and concern about environmental issues and the imperative to address climate change.
- The homogeneity of the discussion within the network is also indicated by its high connectivity.

COMMUNITY DETECTED

- Focusing solely on posts proved more effective than including comments, as comments introduced noise and hindered analysis clarity.
 - The Louvain algorithm applied to the post network yielded the most effective results in identifying communities.
 - Topics identified through posts provided satisfactory results in terms of objective values, allowing for the perception of thought orientation towards climate change.
 - Various algorithms helped trace numerous communities with general themes, facilitating the classification of activist-oriented (such as alarmist) and denialist perspectives.
- Topic 3

Word	Probability
sustainable	0.07
plastic	0.06
waste	0.06
vegan	0.04
bag	0.04
product	0.03
use	0.03
alternative	0.03
economy	0.03
- ex. activist community
- Topic 10

Word	Probability
collapse	0.32
last	0.25
week	0.23
era	0.05
positive	0.03
inevitable	0.03
civilization	0.03
gender	0.03
landslide	0.03
- ex. alarmist community
- Topic 20

Word	Probability
feeling	0.58
nominee	0.40
alone	0.40
deni	0.38
speaker	0.25
denial	0.25
separating	0.22
classic	0.22
representative	0.20
- ex. denialist community

AGENCY

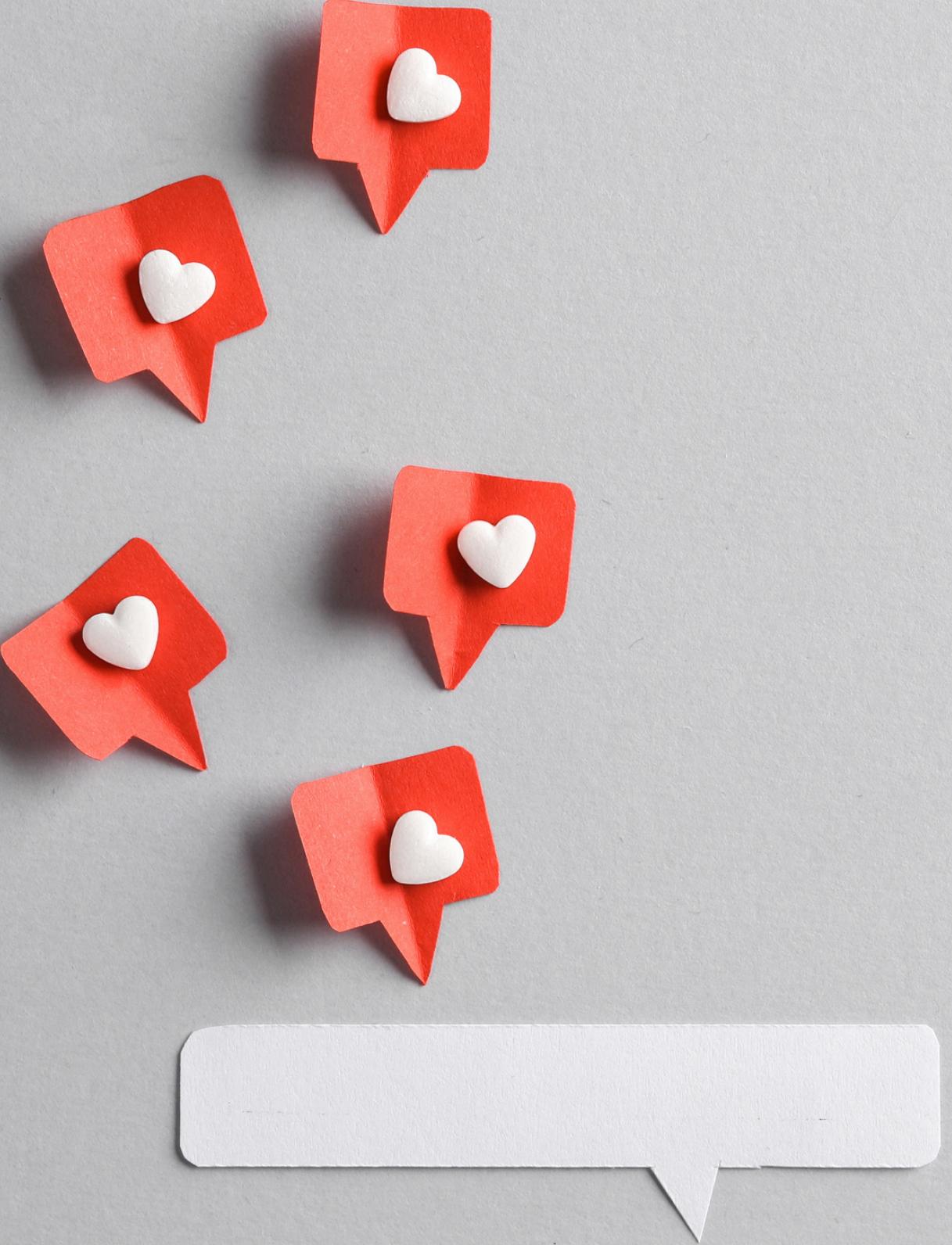
- Most documents in the dataset analyzed with BERTagent exhibit low levels of agency, indicating a tendency to avoid action-related words.
- While the lack of specific thresholds for defining significance complicates interpretation, noticeable differences are still observed in community orientations.
- Communities with higher agency values tend to be more activist-oriented, demonstrating a willingness to take action towards addressing climate change.
- Communities with null agency values may either lack action-related words, representing noise, or express a resignation towards the perceived irreversibility of the current climate situation (resulting in a more alarmist approach).
- Denialist positions are identified in communities with negative agency values, characterized by the absence of action-related words and a belief that climate change is not a real problem.

ROBUSTNESS

- The first graph presented in the robustness analysis includes all documents and evaluates the overall sentiment within the entire network rather than focusing on distinct approaches.
- The analysis reveals a prevalence of negative sentiments over positive ones, which may be more pronounced among alarmists emphasizing the severity of the climate change issue.
- Activists may also express negative sentiments, driven by a sense of urgency to promote concrete actions and solutions.
- Both factions, activists, and alarmists, are united in recognizing the existence of a problem.
- Although denialist factions represent a minority, their skepticism or denial of climate change science may contribute to the overall negative sentiment observed in the network from a different perspective.
- Despite interpretations regarding different factions, the distinction is not clear-cut and objective.

ROBUSTNESS PER COMMUNITY

- A graph was created to analyze the robustness of four communities, aiming to investigate whether alarmists have more negative sentiments compared to activists and to explore differences.
- The graph does not provide clear-cut distinctions, but some differences can still be noticed: lines representing more activist-oriented communities are below, while those characterized by more alarmist words are above.
- It can be concluded that there are different approaches in general, but it is uncertain whether there are more activist or alarmist posts due to the subjective nature of defining activism and alarmism.
- Interpretations are based partially on subjectivity, as there is no objective parameter to define activism and alarmism.



REDDIT

- The data analysis is influenced by the platform used, Reddit, which is known for its political alignment but also for being more open to dialogue compared to other platforms like Facebook or Twitter.
- It is considered normal for certain effects to be more moderated on Reddit due to its nature.
- The limited presence of denialist posts may be linked to the fact that Reddit fosters more dialogue and less echo chamber effects.
- Additionally, the focus on climate-related words in the research may have led to fewer denialist posts being captured, as they may focus on broader conspiratorial trends not specifically related to climate change.



ECHO CHAMBER EFFECT

- The echo chamber effect, where like-minded individuals only interact with each other, was not observed in the analysis.
- The absence of the echo chamber effect suggests that there is significant dialogue between individuals with different perspectives.
- Despite differences in post orientations, such as activism or alarmism, there is a presence of diverse opinions and dialogue under both types of posts.
- This result is attributed to the nature of the platform used, Reddit, which fosters a more open and diverse environment compared to other social media platforms.



LIWC RESULTS

- LIWC analysis provided insights into different community orientations in climate change discourse, with varying levels of anger observed, particularly in Community 17, suggesting potential activism or alarmism.
- Community 12 exhibited markers of anxiety and negativity, indicating a possible denialist stance, while Community 7 showed strong negative markers but also hints of activism, making categorization challenging.
- Community 4 appeared moderate and open to dialogue, while Community 9 emerged as clearly activist, promoting sustainable solutions with a peaceful tone.
- The presence of diverse sentiments and themes within communities highlights the complexity of climate change discourse on social media.
- Overall, the analysis underscores the nuanced nature of online climate change discussions and the challenges in categorizing communities based solely on linguistic markers.

LIWC RESULTS ON 2 OPPOSING FACTORS

- **Present Focus:** The analysis reveals a strong polarization towards focusing on the present in climate change discourse on social networks. This suggests heightened awareness and urgency regarding the current climate crisis, prompting immediate action rather than future planning. Individuals seem more responsive to events with direct impacts, emphasizing the need for immediate measures to address climate change.
- **Negative Emotions:** There is a notable polarization towards negative emotions in climate change discourse, reflecting the emotional impact of environmental threats. Fear, anxiety, and worry are prevalent, fueled by representations of environmental dangers in the discourse. This emotional response underscores a heightened awareness and concern for the devastating effects of climate change, amplified by media sensationalism in social networks.
- **Risk-Centric Perspective:** The discourse predominantly focuses on risks associated with climate change rather than potential rewards. This indicates a tendency to prioritize acknowledging the dangers posed by climate change, potentially overshadowing discussions about the benefits and opportunities of addressing it. However, there is a need to balance risk awareness with optimism about the opportunities for innovation and societal well-being stemming from climate action.

CONCLUSION



Throughout our project, we embarked on a journey through the intricate landscape of climate change discourse within social media networks. We focused on identifying activist perspectives and instances of excessive alarmism, delving into the depths of online communication on this crucial issue. Each algorithm offered us a different view, yet all revealed a wealth of communities and themes. The absence of echo chambers suggests fertile ground for open dialogue. Looking ahead, social media will be crucial in mobilizing the masses and promoting informed dialogue on climate change, shaping a sustainable future based on knowledge, collaboration, and collective action.