

Exploring Confirmation Bias Via Social Network Analysis

Denys Fenchenko

Why Visualization?

```
global_graph_pd.head()
```

	username	user_id	cluster_name	cluster_id	cluster_follow_count	group_id
0	Батал Гамгия	1599163293746302976	@ru_rbc	269770723	535342	0
1	Pletunoff	878179754804957184	@ru_rbc	269770723	535342	0
2	Mila	1599157008330096640	@ru_rbc	269770723	535342	0
3	Set Mortensen	1599155538507567104	@ru_rbc	269770723	535342	0
4	sweetdreamslizz	1599154015195086848	@ru_rbc	269770723	535342	0

Data Attributes:

- Tweet Text
- Tweet Positive Sentiment Score
- Tweet Negative Sentiment Score
- Tweet Entities

```
tweets_df.head()
```

	Cluster	Username	cluster_id	Tweet Text	Entities	Neg	Pos	cluster_follow_count
0		@ru_rbc	269770723	Урсула фон дер Ляйен призвала к...	[ЕС, Урсула фон дер Ляйен, Кита...	0.062	0.161	535342
1		@ru_rbc	269770723	Магомедов был председателем сна...	[Верховного совета, Дагестаном,...	0.000	0.126	535342
2		@ru_rbc	269770723	По предварительным данным, прич...	[]	0.311	0.000	535342
3		@ru_rbc	269770723	В Латвии «Дождю» (признан в Рос...	[России, Грузию, Дождю, Латвии]	0.135	0.000	535342
4		@ru_rbc	269770723	Альянс оставил в силе действующ...	[России, Альянс]	0.053	0.000	535342

Data Attributes:

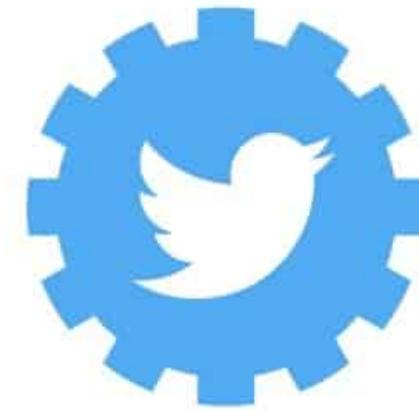
- Username
- User ID
- Cluster Name
- Cluster ID
- Cluster Follow Count
- Group ID



```
tweets_df.head()
```

	Cluster	Username	Cluster ID	Tweet Text
0		@1tvru_news	160881696	Первые -50 этой зимы. Психологи...
1		@1tvru_news	160881696	Из США приходят новости, что по...
2		@1tvru_news	160881696	Вулканический финал недели. Пра...
3		@1tvru_news	160881696	Новые планы по поставкам Западо...
4		@1tvru_news	160881696	Мирные жители Донецкой республи...

Data



Twitter API

```
def get_user_followers(user_name, user_id, num_pages=1, num_followers=10):
    followers = []
    followersPaginator = tweepy.Paginator(client.get_users_followers, id =
    user_id, max_results = num_followers, limit = num_pages).flatten()
    for follower in followersPaginator:
        followers.append(follower)
    return (user_name, user_id), followers
```

```
def get_user_tweets(user_id, num_tweets):
    tweets = client.get_users_tweets(id = user_id, max_results = num_tweets,
                                     exclude = ['retweets'], expansions='entities.mentions.username')
    return tweets
```

```
def get_user_follower_count(user_id):
    # fetching the user
    user = api.get_user(user_id = user_id)
    return user.followers_count
```



Techniques

Data Frame Visualization as a 3D Scatter plot

Tools: Plotly Express

Grammar of Graphics: Cartesian Coordinates,
Points Correspond to Followers, Colors

Correspond To Group Numbers, Twitter data source

Tooltip: username, cluster number, group number

Graph Visualization

Tools: NetworkX, Bokeh, Matplotlib, DiGraph

Grammar of Graphics: Points Correspond to Clusters,
Node Sizes Directly Proportional to the Number of Cluster Followers,
Arrows Represent Follow Relations Between the Users.



Word Cloud Visualization

Tools: WordCloud, NLTK

Grammar of Graphics: Word Scale is
Directly proportional to the Word Frequency
Word Color Improves Visibility

Sentiment-Entity Visualization as a 2D Scatter plot

Tools: Plotly, NLTK, Vader

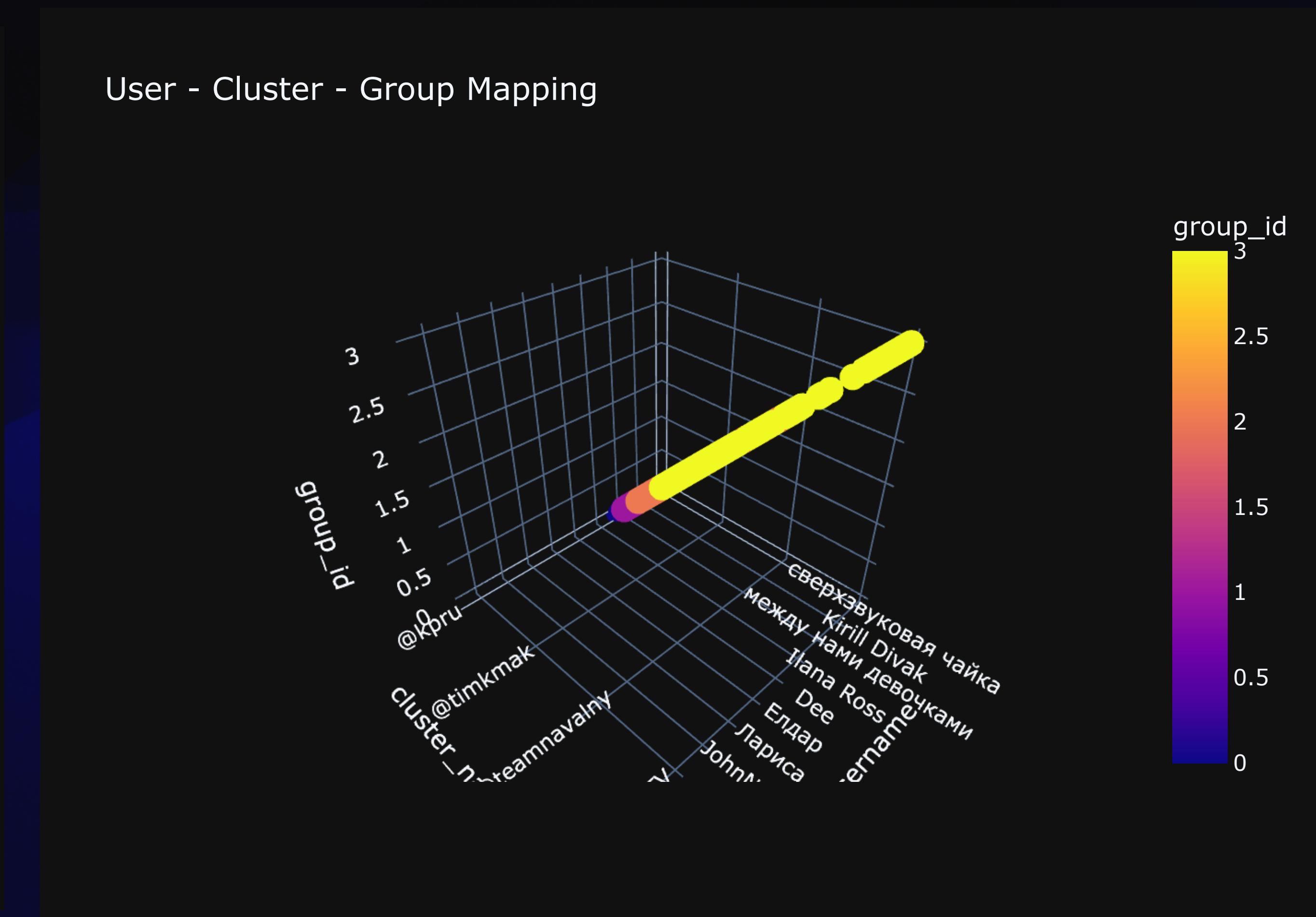
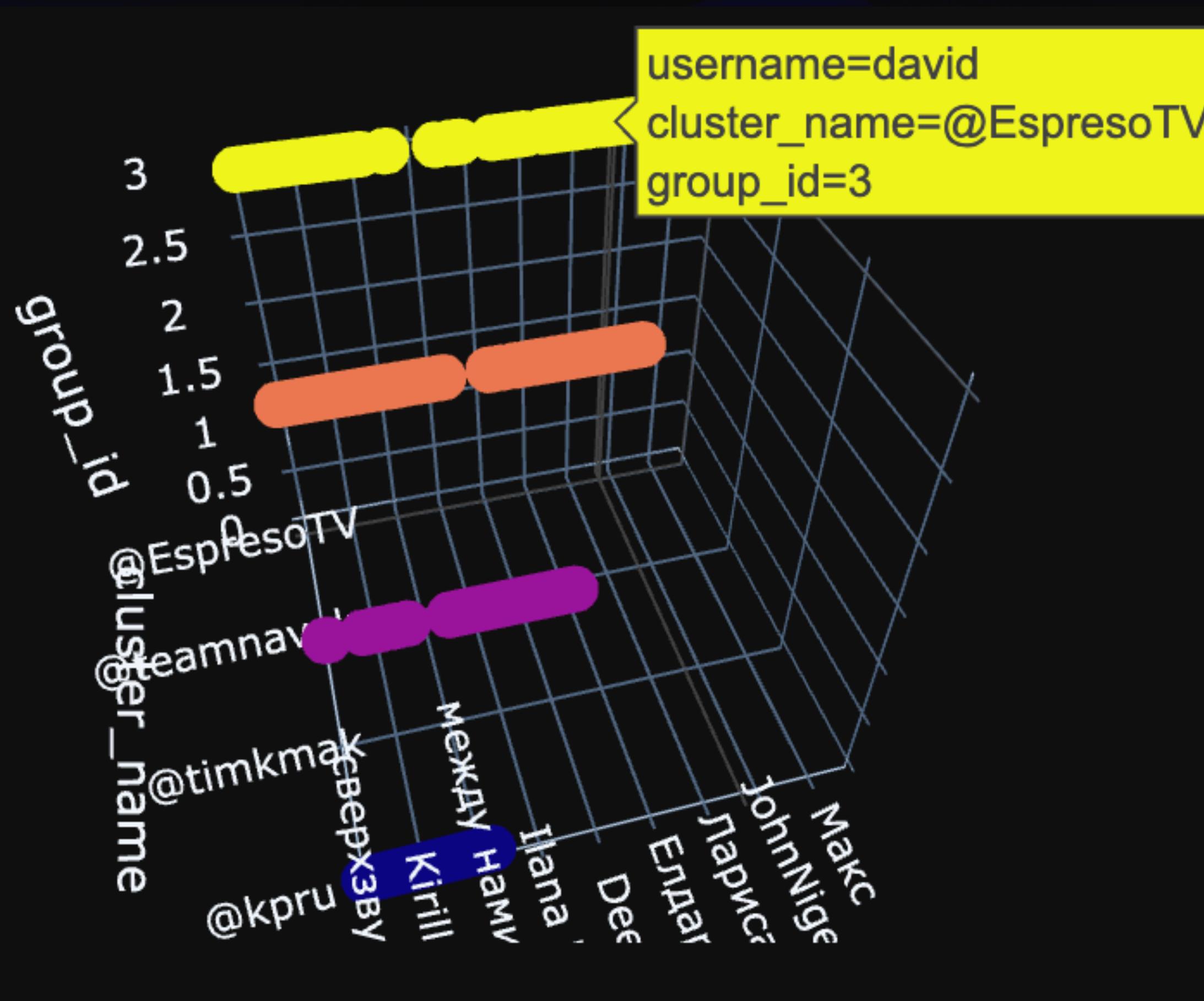
Grammar of Graphics: Points Correspond to
Individual Tweets, Point Size is Proportional to the
Number of followers, x & y Axis Correspond to Sentiment Scores,
Colors Correspond to the Cluster Names

Grammar of Graphics
Element on this Slide?

Results

User-Cluster-Group Visualization:

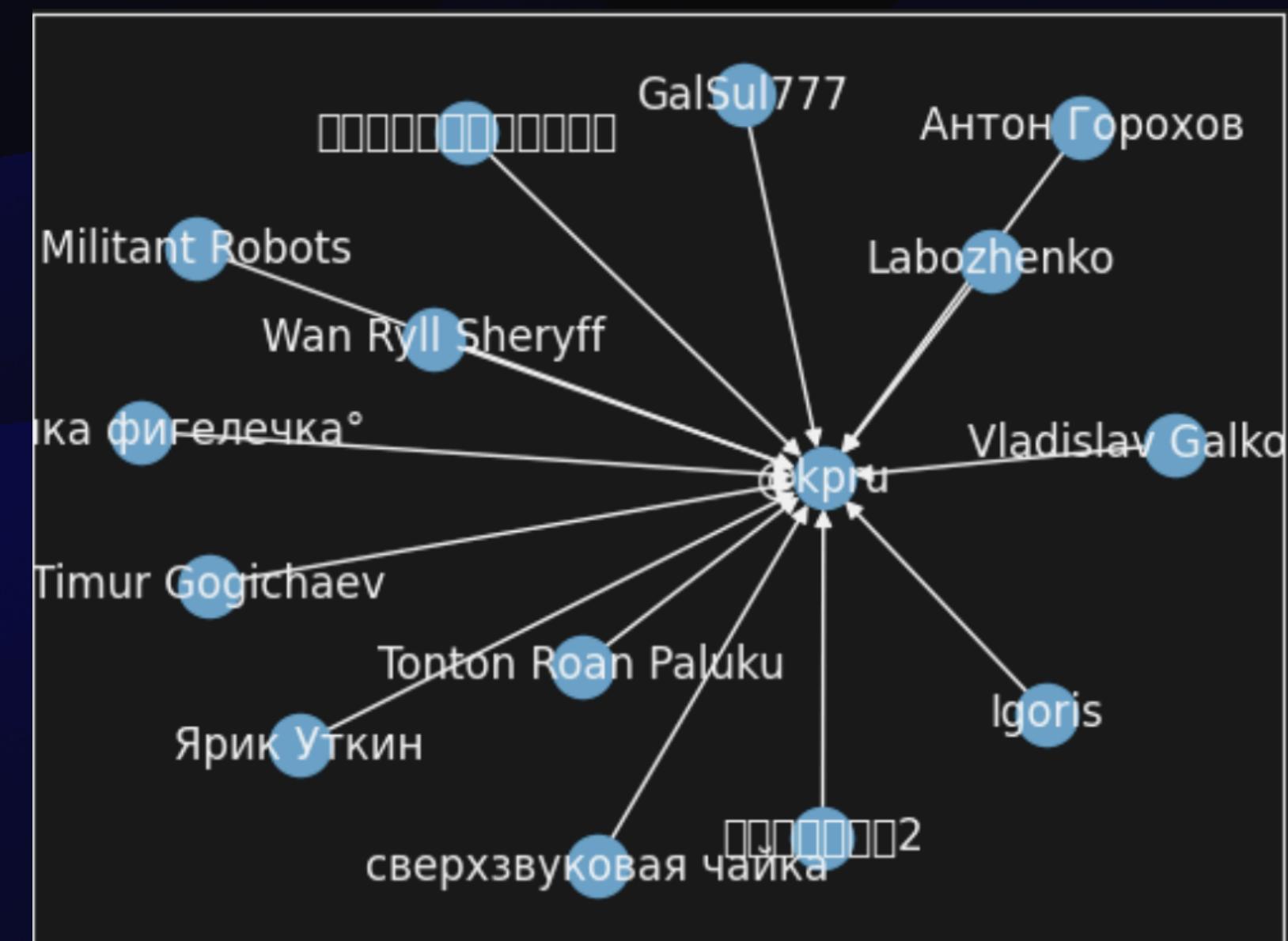
file:///Users/dennifenchenco/NYU-Fall-2022/info_vis/confirmation-bias-vis/visualizations/user-cluster-group-mapping.html



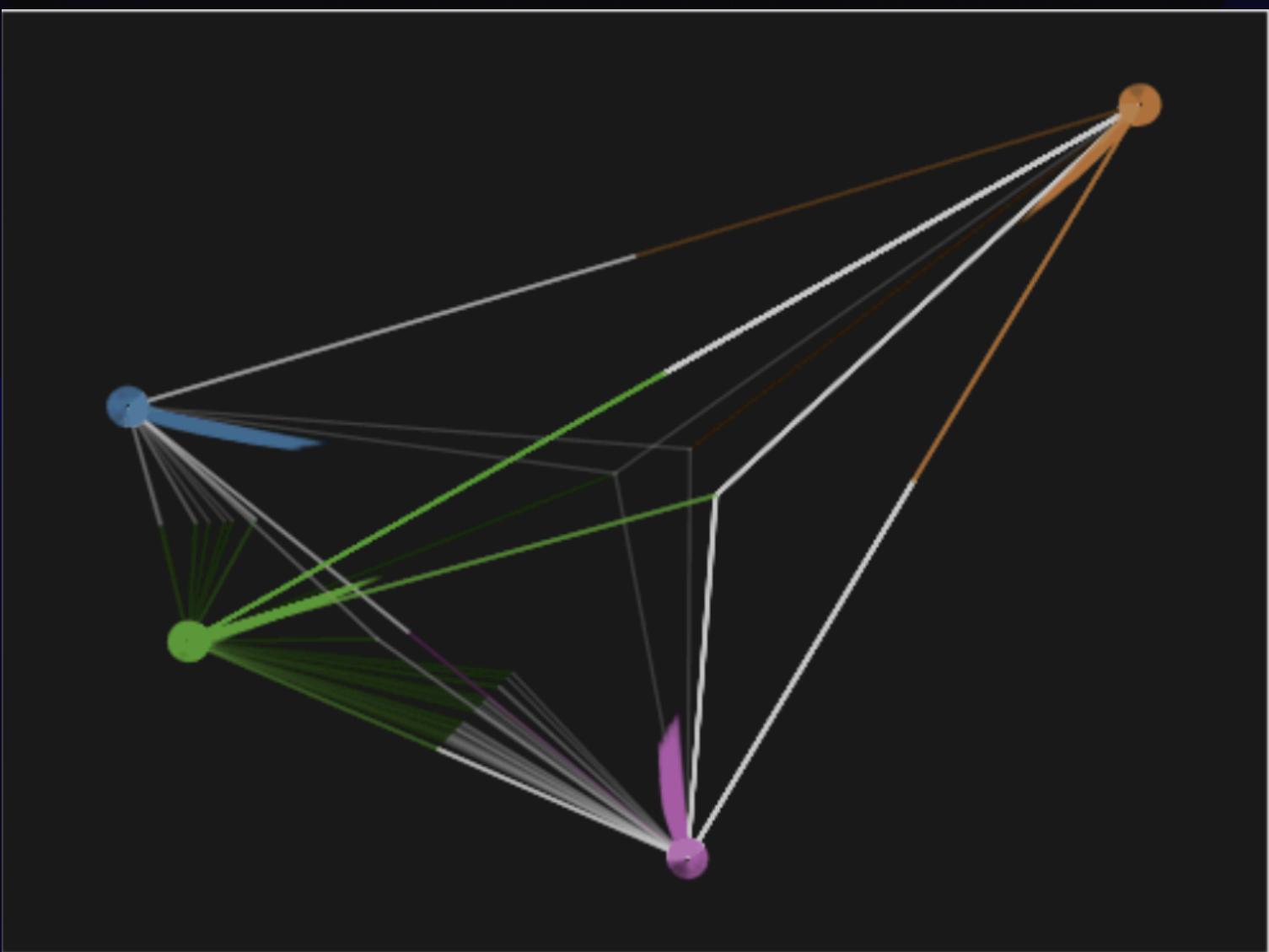
Results

Graph Visualization

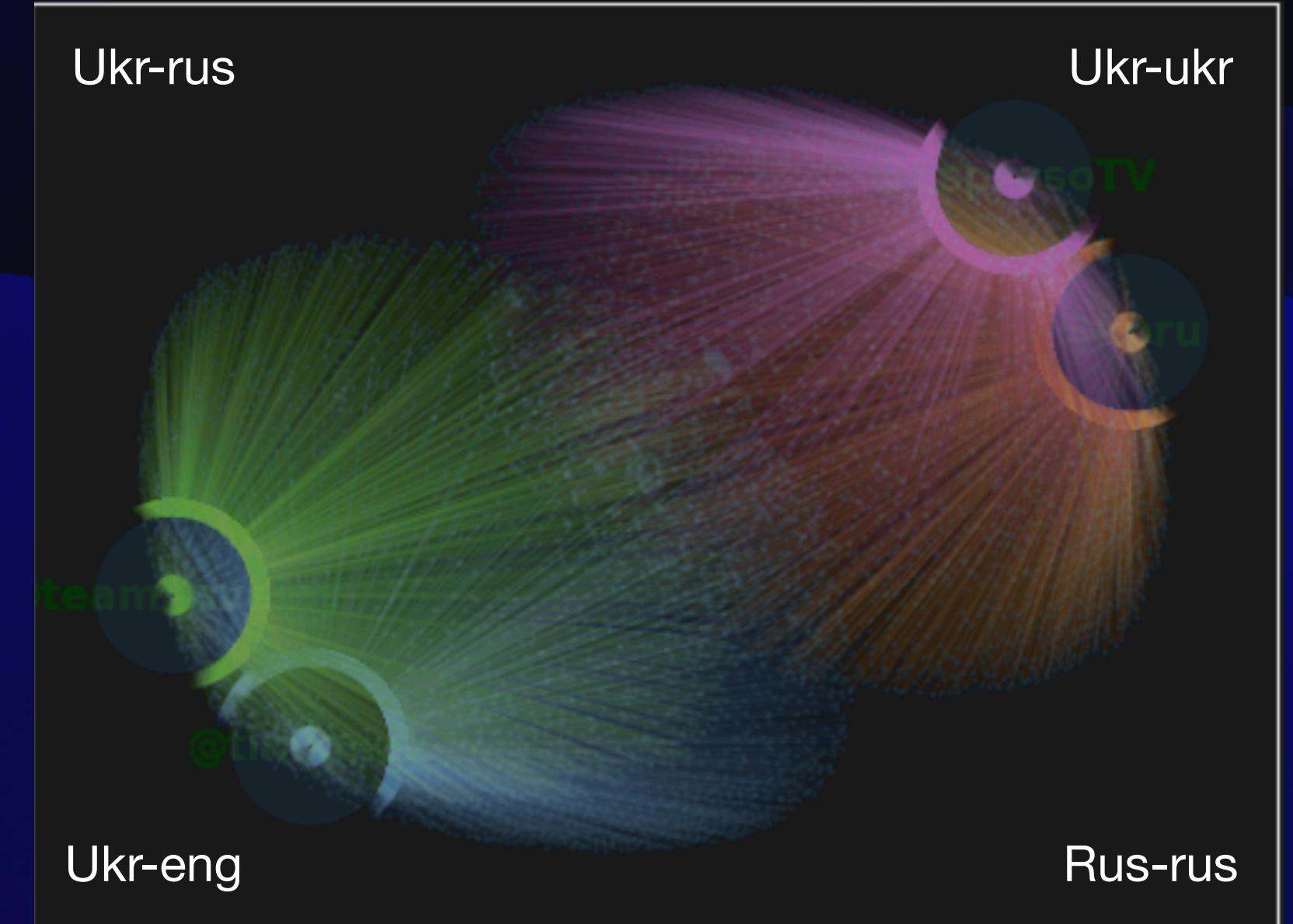
Simple Subgraph Visualization:



Graph Clusters Visualization:



Graph Clusters Visualization $k = 0.15$:



Visualization Attributes:

- Cluster center position
- Follow connection directions

Visualization Attributes:

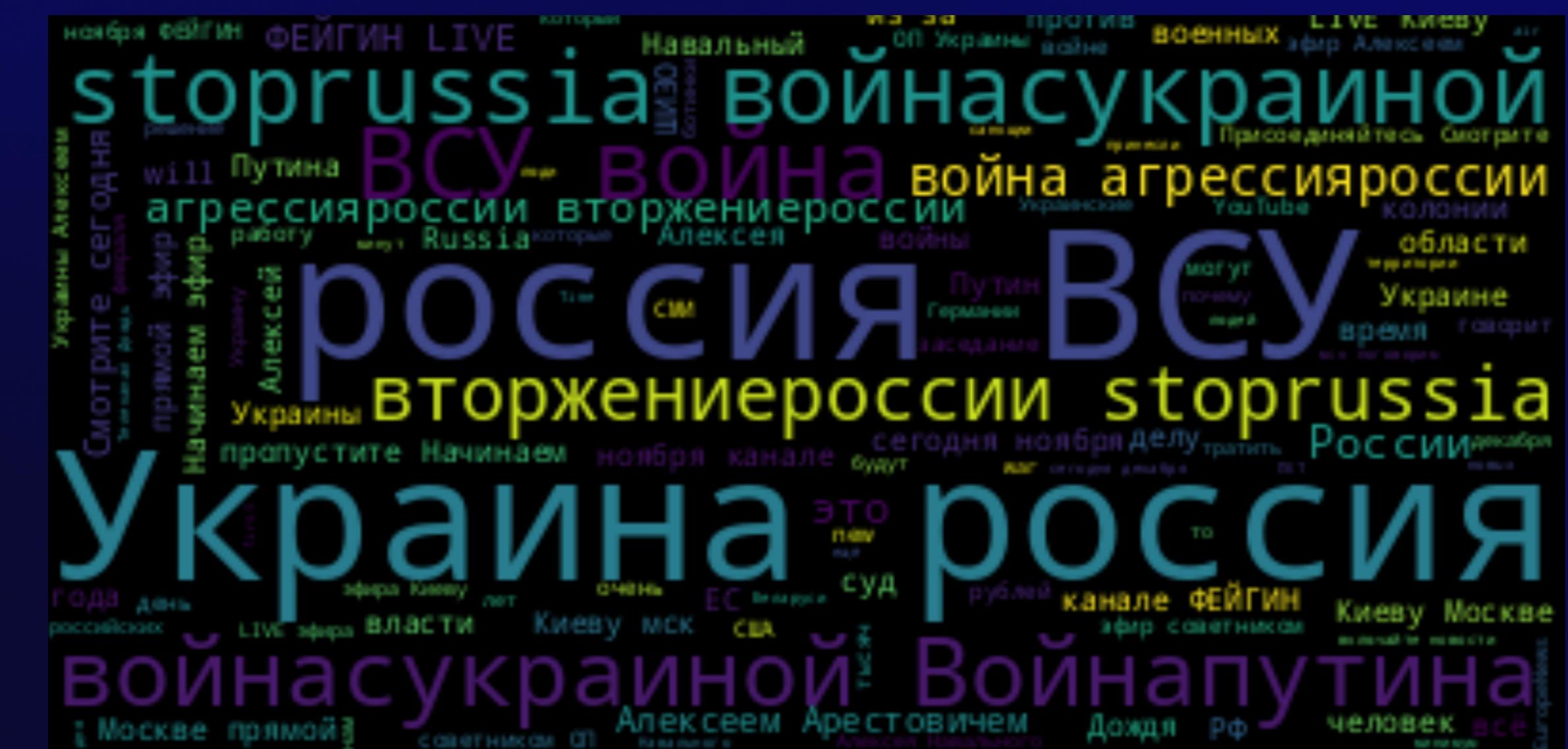
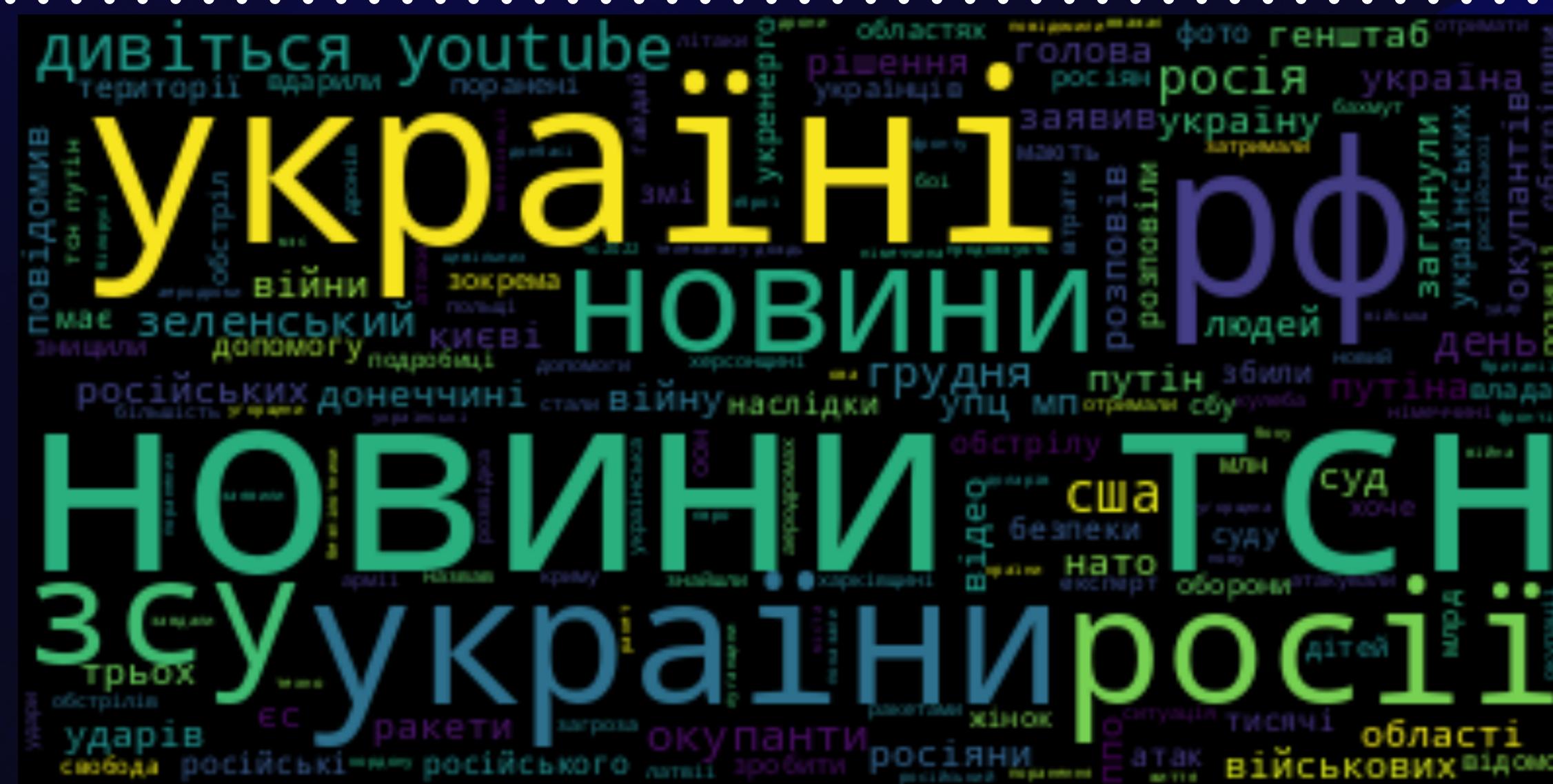
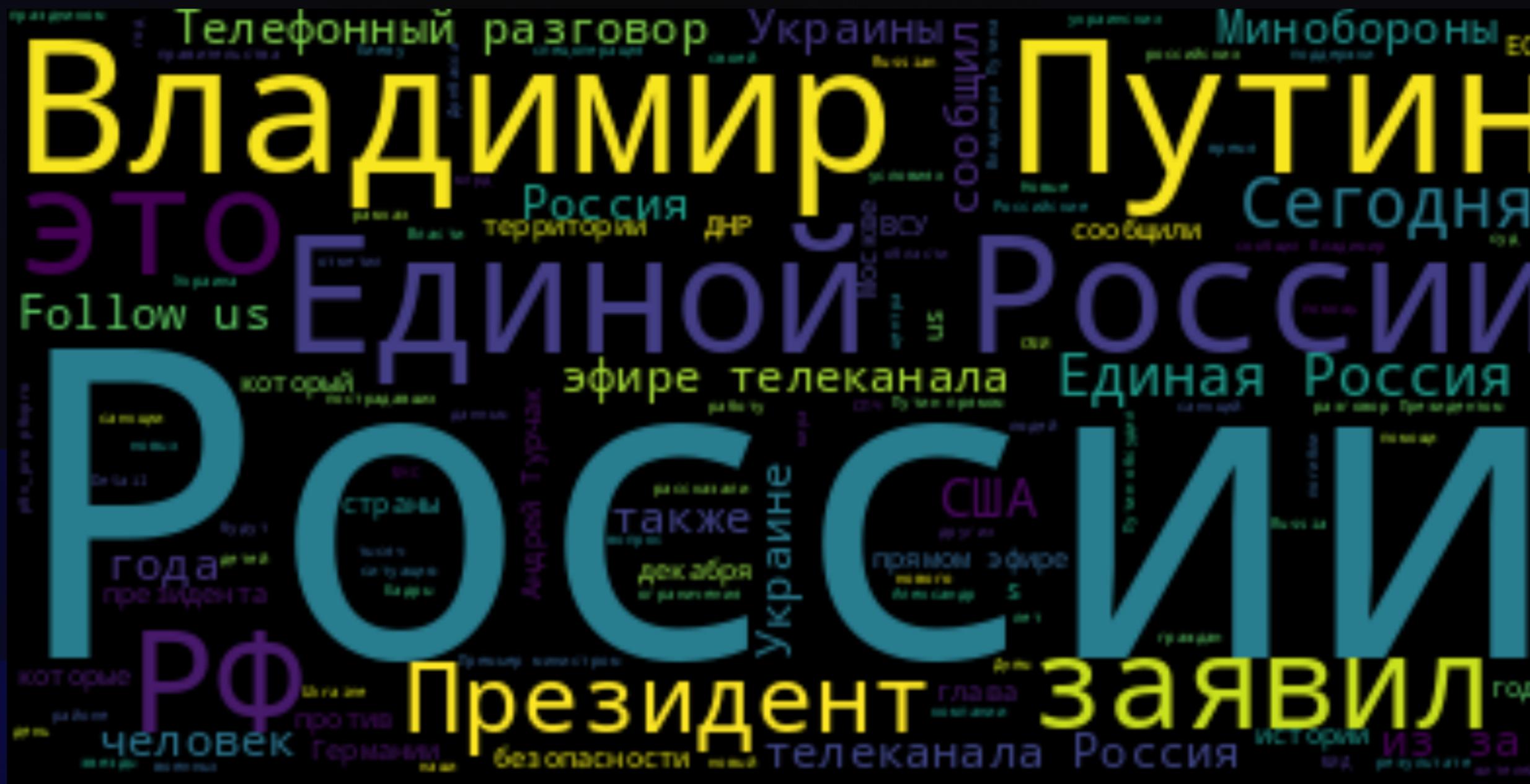
- Cluster center position
- Follow connection directions

Visualization Attributes:

- Cluster center position (NetworkX spring_layout)
- Cluster Sizes \leftrightarrow Number of Followers
- Color \leftrightarrow Community (greedy_modularity_communities)

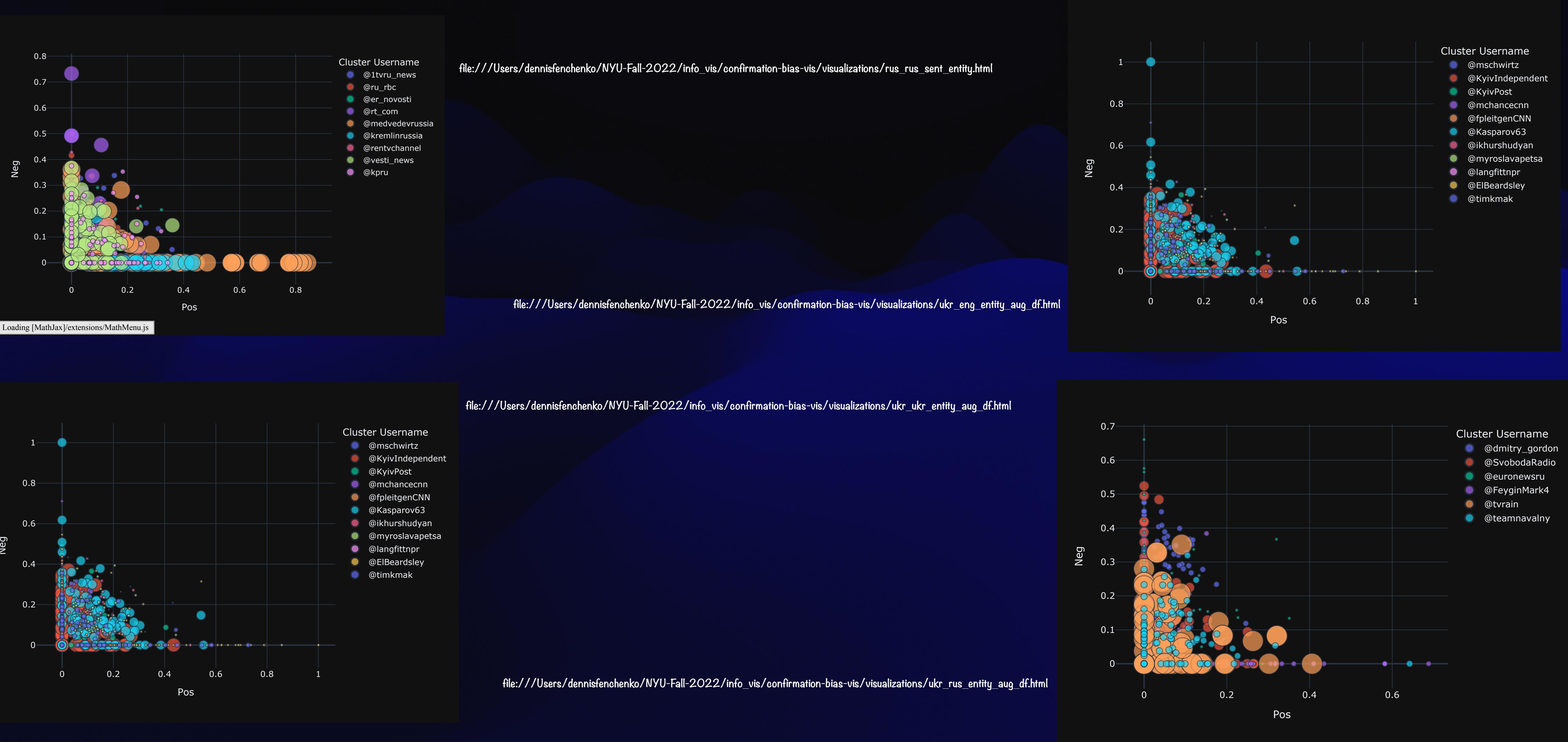
Results

Word Cloud Visualization



Results

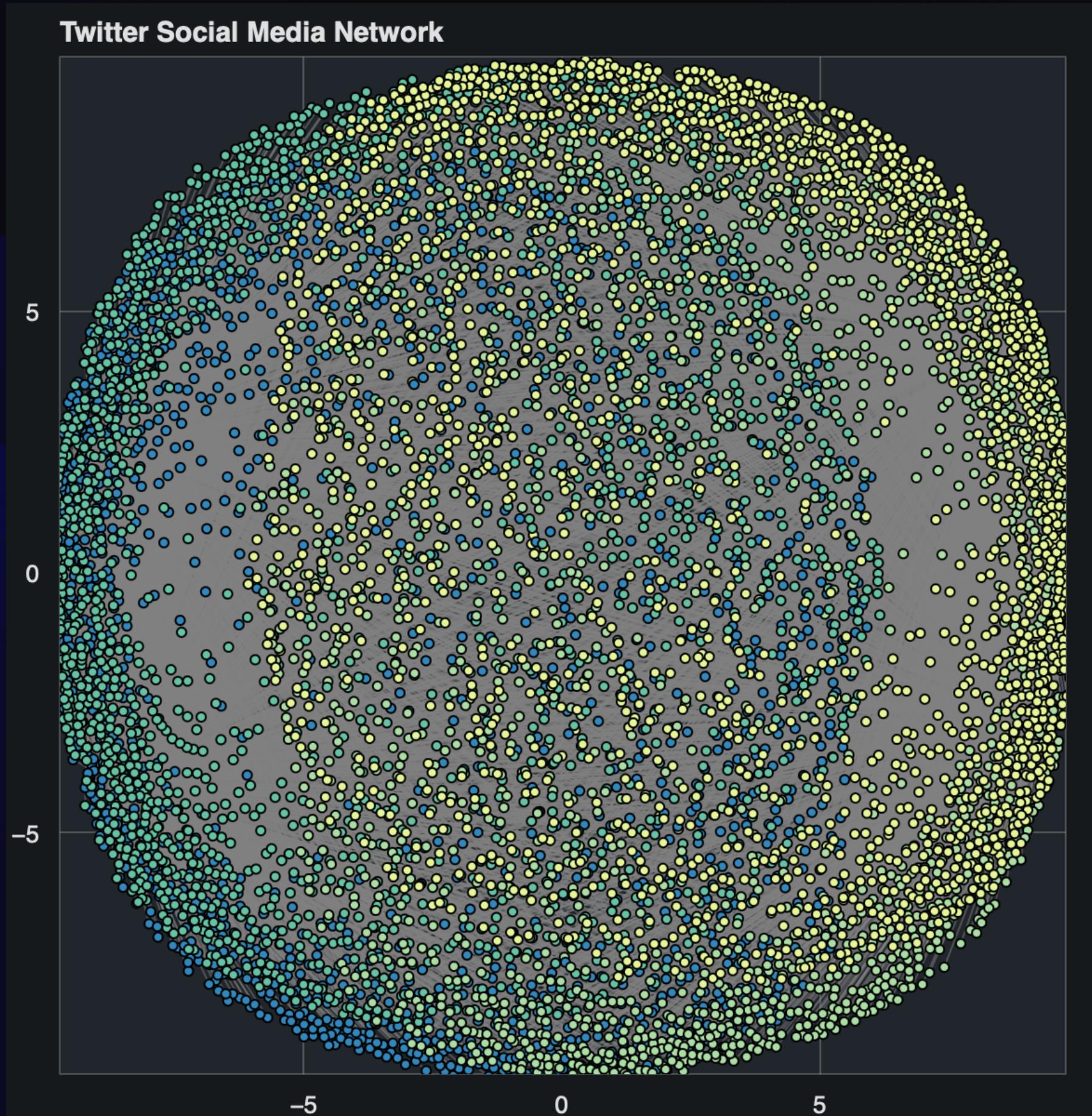
Sentiment-Entity Visualization



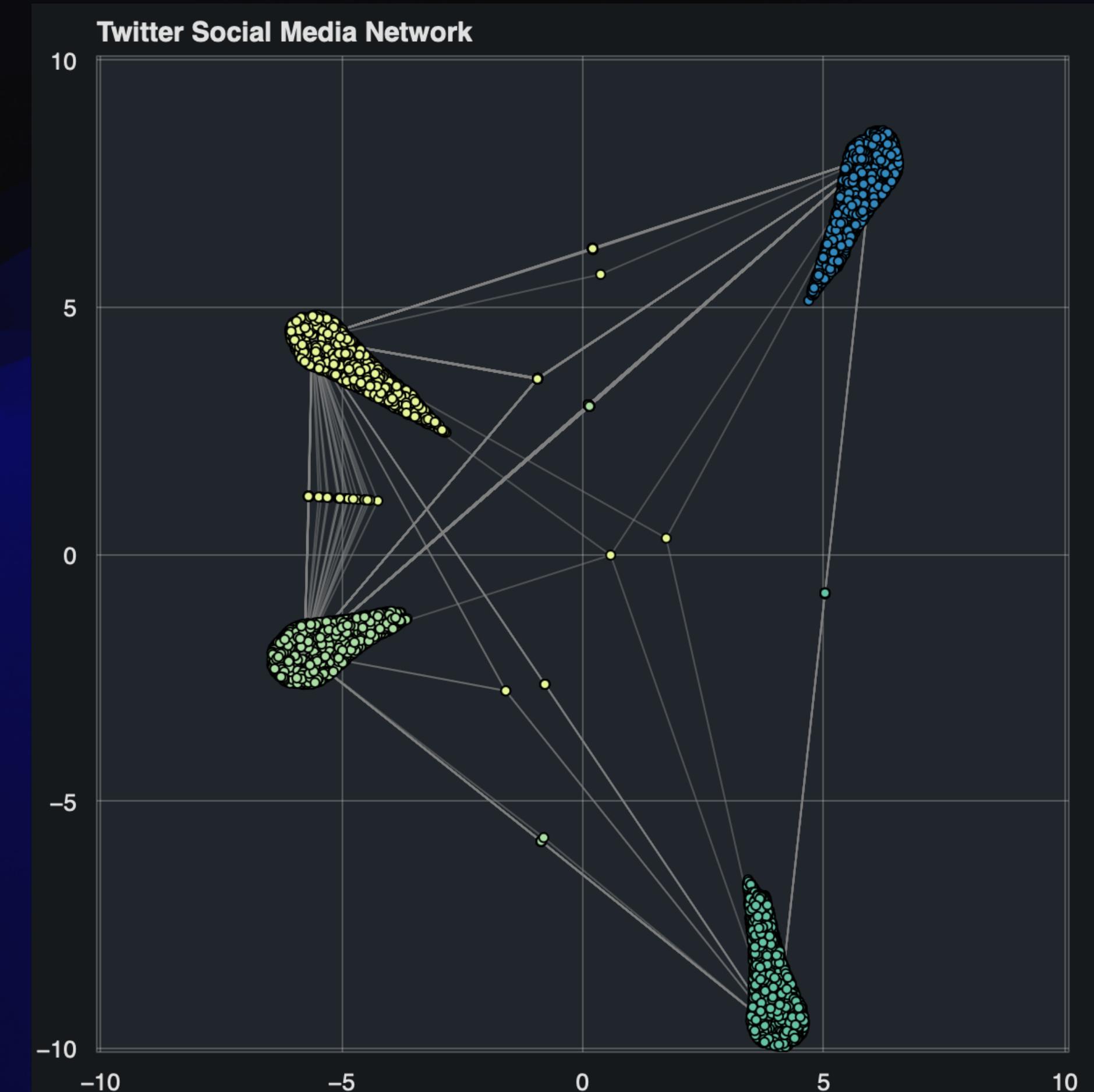
Results

Interactive Graph Visualization of the Overall Network

$K = 0.2$



$K = 0.02$



K : optimal distance
between nodes.

Interpretation

Graph Visualizations:

- People Cluster in groups and there is little communication in between those groups

Word Clouds:

- Both Russia and Ukraine affiliated sources discuss each other
- Russian Cluster Word Cloud mainly contains the terms Russia, Putin, President, etc. This might indicates a certain “personality cult” around the leader on Russian state media
- Ukrainian Cluster: mainly contains the terms Ukraine, Russia, Putin, Ukrainian Armed Forces, etc. This indicates that most tweets are centered around the topic of ongoing war with Russia.

Sentiment-Entity Visualization:

- Russia-affiliated sources show a strongly negative sentiment towards Ukraine and the West
- Ukraine-affiliated sources show a strongly negative sentiment towards Russia and certain decisions made by the western countries

Limitations & Future Work

Limitations & Challenges

Complexity:

- 1100 lines of Python code
- 19 dependency packages

Twitter Rate Limits:

- 15 000 followers / 15 min
- 900 tweets / 15 min
- 2 M tweets per month

Entity-Sentiment relationships hard
to interpret

Large Number of Points is hard to visualize

Graph Construction is Computational Heavy

Data Cleaning

- Tweet Extraction
- Tokenization
- Content Cleaning
- Stop Words Removal

Multiple Languages:

- English
- Ukrainian
- Russian

Differences:

Sentiment analysis
Entity Recognition
Stop Words removal

Future Work

- Larger data pull
- Experiment with different graph visualization layouts
- 3D graph visualization
- Map Sentiments and Entities to specific events
- Add time dimension