

Exploring the COVID-19 Vaccine-related Opinion Shifting of Users on Twitter

Ran Sun¹ and Lu An^{1,2}

1. School of Information Management, Wuhan University, Wuhan, s_ran131497@whu.edu.cn
2. Center for Studies of Information Resources, Wuhan University, Wuhan, anlu97@163.com



SCHOOL OF INFORMATION MANAGEMENT
Wuhan University

Submission ID: 330

Abstract

Understanding public opinion on social media is vital for developing evidence-based interventions that drive positive change in global COVID-19 vaccination efforts. This ongoing study thus addresses the public opinion shifting on Twitter by analyzing over 200k users engaged in online discussions about the COVID-19 vaccine. We then operationalize opinion shifting in three dimensions: word 1-grams shifting, topic shifting, and sentiment shifting. Day-scale usage frequency distributions for 1-grams are compared by rank turbulence divergence, and the Mann-Kendall Mutation analysis is applied to determine the discussion shift based on the divergence score per two days. Text analysis methods are employed for topic clustering and sentiment analysis on tweets. Based on analyzing users' topic shifting, we explore opinion shifting from their sentiment changes on a topic. Our findings indicate that more than half of the social media users experience vaccine-related topic shifting. Users who experience topic shifting and focus on a single topic have slightly different patterns of sentiment shifting when they hold the same sentiment on a single topic. Individuals who focus on a single topic tend to experience increased sentiment polarization toward a particular sentiment direction over time.

Research Question

This study explores the following two objectives:

- to measure the topics and sentiment shifting of social media users during COVID-19.
- to investigate the difference in the patterns of opinion shifting among users.

Methodology

- Data source:** We select Twitter users who have posted at least two vaccine-related tweets per month within half a year. The sample dataset consists of 222,369 users and their 9,274,487 original COVID-19 vaccine-related tweets between January 4 and July 4, 2021, for our empirical study.
- Word 1-grams:** Rank-turbulence divergence is a tunable instrument for comparing the Zipf distributions of any two ranked lists. By taking the frequency of 1-grams published on two consecutive days, we can quantitatively measure the shifting of public discussion on the COVID-19 vaccine from the level of the word.
- Topic modeling:** Topics for COVID-19 vaccine-related tweets are generated by the topic modeling algorithm, BERTopic.
- Sentiment analysis:** The TweetEval evaluation framework based on the pre-trained model BERTweet is used to perform sentiment analysis on tweets.

Results

Rank-turbulence Divergence Allotaxonograph for 1-grams

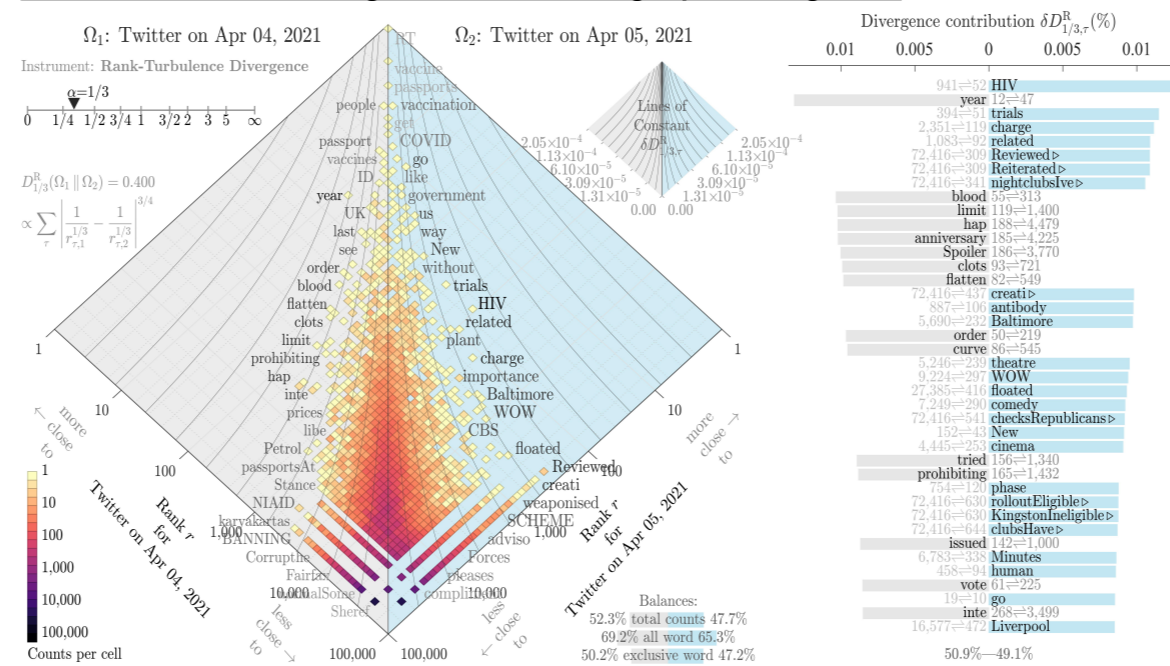


Fig. 1. Allotaxonograph comparing 1-grams usage ranks on two days of 2021/04/04 and 2021/04/05

A sample of 1-grams for the two days of April 4, 2021 and April 5, 2021 contrast diagram is shown in Figure 1. The 1-grams "HIV", "trials", and "antibody" are ranked higher on April 5, 2021. In the graph on the left side, words on either side of the central axis represent words with higher frequencies in the corresponding period. For example, "HIV" was ranked 941 on April 4, 2021 and 52 on April 5, 2021. Words near the central axis have little change in frequency ranking compared to those around the edges of the square. We can also observe the shifting in rankings of a single word in the graph on the right side. The ordered list on the right is generated based on the contributions of each word to the divergence of the two ranked lists. For example, the word "HIV" contributes the highest divergence across the two systems. For exclusive words that only appear in one system and do not appear in the other, their appearance in the other system is counted as zero time.

Analyzing COVID-19 vaccine-related Topic Shifting

- We count the number of topics for each user and find that the number of topics and the proportion of users conform to the long-tail effect. 34.97% of users only focus on a single vaccine-related topic within half a year, and the average vaccine-related tweets of these users is 5.12.
- "COVID-19 vaccines", "Vaccination and anti-vaccination", and "Racial disparities in vaccine distribution and hesitancy" are the top three common topics among these users.
- 65.02% of users tend to shift their focus over time.

Analyzing Different Opinion Shifting Patterns

Grouping the data by user and topic, we get a temporal sentiment series of each user under each topic.

Users who focus on a single topic with the same sentiment:

- More than half of them hold a neutral sentiment, followed by negative (38.1%) and positive (6.6%).
- The ratio of sentiment series showing an increasing trend under the same positive or negative sentiment (33.98% and 40.6%, respectively) is slightly higher than those showing a decreasing trend (31.4% and 36.8%, respectively).

Users who focus on a single topic with different sentiments:

- The most common sentiment shifting is a back-and-forth shifting between neutral and negative such as "neutral \leftrightarrow negative" and "negative \rightarrow neutral \rightarrow negative".
- The least common sentiment shifting is between negative and positive such as "negative \rightarrow positive \rightarrow negative". Among them, the shifting is more common beginning with positive and ending with negative.

Users who shift across multiple topics over time:

- Most of them hold the same sentiment on a topic (74.3%), followed by two sentiments (14.28%) and three sentiments (11.4%).
- Similar to users who focus on a single topic, those who shift across multiple topics over time and hold different sentiments for each topic exhibit a common pattern of back-and-forth shifting between neutral and negative.

Conclusion

- 65.02% of users change their topics over time, driven by external events or other factors.
- People who focus on a specific topic are more likely to increase their emotional polarization toward a particular sentimental direction over time.
- Users who engage with multiple topics maintain the same sentiment in subsequent tweets.
- The pattern of a negative state changing to neutral has the highest probability.
- When users change their sentiment on a topic, the probability of shifting toward negative sentiment is higher than positive sentiment.