How anti-vaccination sentiment spreads on Twitter

Andrea Mock
21 Wellesley College Road
Wellesley, MA 02481

Abstract

The COVID-19 pandemic has heightened the discussions around vaccination and worries around the safety of vaccinations. Therefore we were curious to see how these developments have impacted information spread and discussions online surrounding this topic. This paper describes a study of 43,041 tweets from August 2017 to August 2020 on the topic of vaccination, specifically, tweets that had these hashtags: #vaccinesarepoison, #vaccineskill, #vaccineinjured, #vaccinescauseautism, #vaxxed. The study identifies the top accounts related to those tweets and examines their temporal progression. The study identifies three spikes of activity, which we hypothesize might be related to the coronavirus epidemic in 2020 and the release of the documentary "Vaxxed II" in 2019. In addition, we determined that there were three main topics that the corpus of tweets centered around: children and vaccination mandates, doctors and vaccine developments and vaccines in general, and the risks surrounding them.

1. Introduction

In the past months, there have been growing fears and misinformation surrounding the development of the COVID-19 vaccine development. The fears that vaccines can cause autism and other harmful effects have been around for a long time, but the recent news articles and politicization of the topic have made the discussion of the topic even more relevant. Especially because many individuals are confined to their homes, they are increasingly turning to the internet and social media platforms to interact with others on topics including COVID and vaccination. Therefore the study of social media messages related to attitudes toward vaccination has become an important area for study, specifically the study of anti-vaccination sentiment.

Numerous studies including (Blankenship et al. 2018) and (Mahajan et al. 2019) have laid their focus on studying the anti-vaccination community on Twitter. Twitter provides an easily accessible platform to study social networks as it allows researchers to access people's tweets and associated information, providing a wealth of data for further analysis. Furthermore, (Burki 2019) found that online platforms, including Facebook and Twitter, were key drivers for the anti-vaccination community and the preventable outbreak of diseases due to vaccine hesitancy. Even before the Coronavirus pandemic, there has been worry about social media platforms' role in augmenting anti-vaccination sentiment. UNICEF published a study conducted in 2013 (Unicef 2013) about the worrying effects of misinformation on social media platforms surrounding the vaccine debate. Their study looked at eastern European countries, among them Romania, Poland, and Bulgaria, and examined people's social media use and how widespread anti-vaccination beliefs were online. Their study underscores the importance of understanding how these communities organize and interact online, especially on Twitter. Thus, we also chose to focus on the anti-vaccination movement on Twitter. Our methodology involved looking at particular hashtags and frequency tracking of the most popular terms being used and identifying the most salient topics within these communities. Our paper focuses on answering the following three main research questions:

- RQ1: How has the anti-vaccination community's engagement on Twitter evolved over time in light of the COVID pandemic?
- RQ2: Are there members in the anti-vaccination community who play a prominent role within the community?
- RQ3: What are the most frequent terms and topics being used? What topics does the anti-vaccination community engage with on Twitter?

Our paper is structured in the following format: we begin by providing a brief overview of relevant literature, then continue with a description of our methodology and techniques, then discuss our results and finally conclude by presenting our conclusions and discussing future work.

2. Related Work

Vaccines are an important tool in the fight against diseases. Therefore with the COVID-pandemic, the search for a vaccine has garnered a lot of attention. Simultaneously, people have started to question the safety of the swift development process of a vaccine against COVID. The questioning of vaccine efficacy is not a new phenomenon. In the 1990s, a now-retracted study propagated the idea that vaccine usage

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

is linked to autism. This idea has stayed with many people, nevertheless. In the age of technology, these sentiments can be studied in an online setting. Twitter has been one of the places of interest for these types of analyses. Below is a brief review of the literature surrounding the study of pro- and anti-vaccination sentiment over time, the authorities, and relevant communities in the vaccine discussion, and the use of particular hashtags.

2.1 Longitudinal study of pro- and anti-vaccination sentiments over time

A popular approach taken by many researchers consists of tweets over a longer period of time. (Blankenship et al. 2018) and (Gunaratne, Coomes, and Haghbayan 2019) collected tweets for the hashtag #vaccine over multiple years. (Blankenship et al. 2018) focused on tweets from 2010 to 2016 and (Gunaratne, Coomes, and Haghbayan 2019) focused on tweets from 2010 to 2019. In a subsequent step, the collected tweets were then identified as pro- or anti-vaccine. Over time they witnessed an increase in pro-vaccine tweets and a decrease in anti-vaccine tweets starting in 2014. Similarly, (Mitra, Counts, and Pennebaker 2016) examined tweets collected over a span of four years from January 2012 to June 2015, collecting a total of 15,240 tweets generated by 144,817 unique users. They used five different search terms, including 'vaccination+autism,' 'vaccine+autism,' 'mmr+vaccination,' 'measles+autism,' and 'mmr+vaccine.' The researchers then classified the tweets as holding anti-vaccination or pro-vaccine standpoints. Their analysis involved looking at the users who recently joined the anti-vaccination conversation on Twitter and examined these users' timelines to see how their views have evolved.

2.2 Influencers and communities in the vaccine discourse

Twitter is structured such that users can follow others, comment, and retweet tweets. This creates room for different communities to emerge and interact with one another. (Sanawi, Samani, and Taibi 2017) examined the key influencers within the vaccine community on Twitter. The researchers focused on anti-vaccination communities and identified six different influencer groups. These groups included celebrity doctors, governments and government institutions, medical journals and bloggers, and the media. Using social network analysis and tools, such as NodeXL and Gephi, they visualized these networks and identified users with the highest number of mentions or retweets.

(Blankenship et al. 2018) and (Gunaratne, Coomes, and Haghbayan 2019) focused on classifying the collected tweets as anti- or pro-vaccine for the hashtag #vaccine to measure the size of the communities and their standpoints with regards to vaccines. (Blankenship et al. 2018) found that around 32% of the tweets they had collected were pro-vaccine, 43% neutral, and 24% were anti-vaccine. From those who voiced anti-vaccine sentiments, around 47% of the anti-vaccine tweets talked about vaccines' dangers, and 26% focused on government mistrust and distrust of pharmaceutical companies. Similarly, (Gunaratne, Coomes, and Haghbayan 2019) identified 154 pro-vaccine and 125 anti-vaccine hashtags within the 1.6 million tweets they had collected. 86% of the tweets were pro-vaccine, while around 12% only held an anti-vaccine stance.

(van Schalkwyk, Dudek, and Costas 2020) put their focus on analyzing the anti-vaccine supporters on Twitter. In their analysis, they focused on Twitter users who cited scientific articles for making claims about anti-vaccination. They were interested in seeing how Twitter users interacted with academic journal articles and thus focused on 113 open-access journal articles, and utilized a website called Altmetric that tracks the number of accounts and tweets, including these articles. Their analysis included identifying the account with the highest centrality score, meaning those who mention the most articles in their tweets and have the most mentions in common with other Twitter users. They found that the most active Twitter users are in the middle, where they span both the anti-vaccination and pro-vaccination communities. Also, they found the density of anti-vaccination clusters to be very high. Simultaneously, for mixed groups, defined groups that are center around one or two Twitter accounts, include people pro- and against vaccines. Anti-vaccine supporters were more likely to cluster around multiple hashtags, while the pro-vaccine tweets largely belonged to one group.

A different approach to measure the role of particular users and tweets was described by (DeMasi, Mason, and Ma 2016). To gauge engagement for any particular hashtag, they defined engagement as having received a retweet or favoring of a tweet. Thus, a particular tweet's engagement is then used as the probability of engagement for the hashtag. They also looked at the diversity of adoption and the number of users using that particular hashtag. This approach allowed them to conclude that hashtags surrounding events are widespread, although they do not have as high an engagement factor as content-oriented to a specific group.

Aside from measuring engagement and identifying the most frequent hashtags and the frequency of tweets that involved anti- and pro-vaccine tweets, sentiment analysis was also a popular method used by researchers to analyze Twitter data, as seen in (Mahajan et al. 2019) and Brooks (Brooks 2014). Using the NRC Word-Emotion Association Lexicon (EmoLex), they were able to identify the most common words used in anti- and pro-vaccination tweets and identify the emotions most present in particular tweets. They found that anger, disgust, fear, sadness, and negative sentiment were more present in anti-vaccination tweets.

2.3 Strategic use of hashtags

Within Twitter, hashtags are utilized for different purposes. (DeMasi, Mason, and Ma 2016) were able to identify four categories of hashtags used in different settings: events, stable, periodic, and stochastic. They found that certain hashtags were only used during particular time periods, such as #friday, which had its peak weekly on Fridays. On the other hand, there were stable hashtags used over the years but did not have a high engagement.

(Gunaratne, Coomes, and Haghbayan 2019), through their analysis, were able to identify the prominent hashtags used by the anti-vaccine as well as the pro-vaccine community. The most frequent hashtags with anti-vaccine sentiment were #cdcwhistleblower, #vaxxed, #hearthiswell, #novax, and #cdcfraud. On the other hand, pro-vaccine tweets were largely centered around the hashtag #vaccineswork instead of multiple hashtags.

3. Data and Methods

3.1 Data Collection

Since we were interested in the temporal changes of anti-vaccination sentiment on Twitter, we collected tweets from a three year time period from August 2017 to August 2020. To gather the tweets, we first used the Python library snc-scraper that allowed us to scrape a total of 43013 tweets from the following five hashtags: #vaccinesarepoison, #vaccineskill, #vaccineinjured, #vaccinescauseautism, #vaxxed. These hashtags all showed anti-vaccination sentiment and were partly banned from Instagram in 2019 (Yurieff 2019) to curb the spread of misinformation about vaccines. Since these hashtags had received considerable attention on Instagram, they were a good starting point to learn more about Twitter's anti-vaccination community.

Snc-scaper has its limitations, as it only collects Twitter ids and allows for the collection of Tweets in batches of roughly 3000 tweets at a time. Therefore we collected tweets for each hashtag separately and by year. Since snc-scrapeR only provided the Tweet ids, we utilized the Twitter API to gather the actual tweets and additional information on the users who had written these tweets. This information included when the user had joined Twitter when a tweet was written, the hashtag it included, the number of retweets, followers, mentions, and people a particular user follows. We went through this route of using snc-scrapeR and the Twitter API as the original library GetOldTweets3 that allowed an easy extraction of tweets without any additional credential had been discontinued during our research phase, which limited the amount of subsequent analysis possible due to the limited time.

Our data was stored in separate files that allowed us to first look at each data set individually. Comparing all the datasets, we saw that the hashtag #vaxxed had received the most attention for our collection period with a total of 34,994 tweets while #vaccinescauseautism had 1,423 tweets, #vaccineskill 4,542 tweets, #vaccinesarepoison 802 tweets, and #vaccineinjured had 1,780 tweets. For these tweets, the average tweet length was 186 characters.

3.2 Data Cleaning

Before identifying the most popular hashtags and model topics, the raw data had to be prepared and cleaned first. Twitter allows users to write 280 character messages, which are often not written in clear language so that a data cleaning step is required before interpreting them. To do this cleaning, the NLTK library was used. First, we used it to filter out stop words, which are pre-defined in NLTK. Next, we removed links, hashtags, and mentions from the tweet's text. We converted all text to lowercase in each tweet to enable us to calculate the actual frequency of words in the tweets and enable for the identification of the most salient topics.

3.3 Topic Identification

After that preparation, we used the Latent Dirichlet Allocation method (Mahajan et al. 2019), a topic modeling algorithm to assess the most popular hashtags and identify topics. Topic modeling is an unsupervised learning algorithm that allows identifying topics within texts. The Python library PyLDA, which implements this algorithm, was used to perform this step. However, we also tried different LDA models, using the sklearn and the gensim libraries and their predefined LDA algorithms. To find the most salient topics, we tried a different number of topics and, with 10 topics, were able to identify words that occurred the most within each topic. Using a visualization of the topics and their inter-topic distance, we thus were able to identify topics that did not have much overlap but, at the same time, presented main topics that were being discussed in our corpora of tweets.

4. Results

Using the collected data, we explored the following three questions:

- RQ1: How has the anti-vaccination community's engagement on Twitter evolved over time, given the COVID pandemic?
- RQ2: Are there members in the anti-vaccination community who play a prominent role within the community? What are the most frequent terms and topics being used?
- RQ3: What topics does the anti-vaccination community engage with on Twitter?

4.1 RQ1 Engagement over time

In general, we have seen the number of anti-vaccination tweets fluctuate over time. Certain hashtags have gained popularity, but also lost popularity over time. From 2017 to 2020, there are three visible spikes: May 2020, another around March 2019, and another around September 2017, as seen in Figure 1.

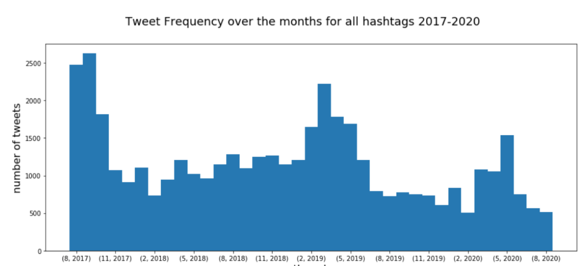

Figure 1: Number of tweets that appeared over the time period from August 2017 to August 2020 that included at least one of our five hashtags.

To better understand what hashtags are related to these spikes in anti-vaccination tweets, we also examined the tweets over time for each hashtag individually. In the hashtag #vaxxed, (shown in figure 2) we see that it follows roughly the same pattern as our overall hashtag data. The large number of tweets can partly explain this in our dataset that used this hashtag. The two big spikes seen in the #vaxxed data are around September 2017 and March 2019. A possible explanation for these spikes is the release of the movie "Vaxxed II: The People's Truth," which was released in March of 2019. The second movie in an anti-vaccination documentary purports to document vaccine-related injuries around the United States. However, the spike in 2017 cannot be explained by the release of the first documentary "Vaxxed", which was released in 2016, a year earlier, when we saw the spike in our data. However, since a large amount of #vaxxed tweets could indicate people still actively engaging with the movie and the anti-vaccination community after its release and translation in other languages, but to ascertain this hypothesis, further research would be required.

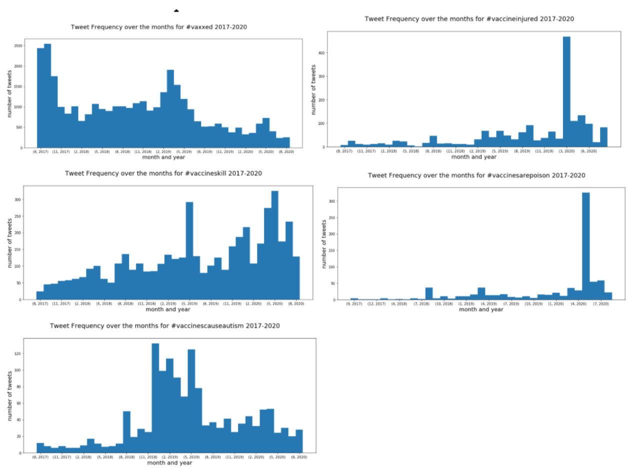

Figure 2: The tweet frequency for each hashtag from the start of the collection period in August 2017 to August 2020. (Note: the axes differ for each subplot)

The hashtag #vaccineinjured shows a very different timeline, with a spike in March of 2020. In general, more users are engaging with the hashtag since March 2020. This might indicate that the COVID pandemic led to people engaging more around the search for a COVID vaccine and the related risk that people fear from a vaccine. To examine this hypothesis, we filtered the tweets from March 2020 to June 2020 and identified the most frequent words being utilized. Among the top 15 terms, the coronavirus appeared 394 times.

Similarly, the hashtag #vaccinesarepoison also did not have very much engagement before the COVID pandemic. Beginning in May 2020, however, we can see a big spike in the number of tweets that use this particular hashtag. Again, similar to the hashtag #vaccineinjured, the heightened discourse around finding a vaccine against the coronavirus most likely fueled the number of users engaging in the anti-vaccination community. As we look at the terms that appear in the individual tweets, we also see the name of Donald Trump and Bill Gates, two public figures who often get mentioned when it comes to conspiracies, especially surrounding the Coronavirus. The hashtag #vaccineskill shows more engagement than the other hashtag, and as Figure 2 details have gained traction over time. By contrast, the hashtag #vaxxed has become less popular over time. Finally, the hashtag #vaccinescauseautism shows to have had the most traction during the middle of 2019. Searching for news surrounding autism in 2019, we find that a big study with 500,000 participants was published in early 2019, again refuting the claim that vaccines cause autism. This study collected data over 20 years and thus held considerable weight in emphasizing that vaccines are not harmful. At the same time, it could have resulted in more individuals engaging with the topic and becoming aware of vaccines and autism. To prove this in a statistically rigorous way, future work might perform a correlation study of trending news topics and these Twitter tags.

4.2 RQ2 Anti-vaccination community

Our dataset contained 43,041 tweets, and we can see that the top Twitter account has contributed to around 16.5% of the total tweets in our dataset. Similarly, the next frequent Twitter users also contributed to roughly 6% of the tweets. In addition, we see that a select number of users who were part of the top tweeters for a particular hashtag also engaged with multiple hashtags simultaneously and were also among the most frequent users of different hashtags. For example, a couple of users were very active across four of the five hashtags we examined. Interestingly, the most active user overall only engaged with the hashtag #vaxxed and did not appear to use the other hashtags as often. Examining these accounts in more detail showed that the median number of followers, the top thirty users in our dataset had was 1,834, and the number of accounts they followed had a median of 1,516. The earliest of these accounts were created in March of 2008, and the latest was created in May 2020, with a mean creation date in November 2013. The creation dates of these accounts indicate that they have been active for quite some time.

4.3 RQ3 Topics in the anti-vaccination discourse

To examine which terms were the most popular across the different hashtags, we identified the top terms used within each hashtag. To avoid including commonly used English words, we removed stop words and other symbols that would have altered our frequency calculation, as mentioned in section 3.2.

Comparing the different words that show up in the tweets using different hashtags, we see terms including negatively connotated words such as "risk" and "injury." At the same time, popular words include health-related words such as "doctors" and "health." Furthermore, we also see references to figures such as President Trump and Bill Gates to be prevalent. As can be seen from the graphics, terms such as Bill and Gates occur almost in equal amounts suggesting that they should be considered as one. A bigram analysis for each hashtag provided further clues of the actual sentiments expressed within each hashtag and tweet. Creating bigrams creates an even clearer picture of what terms are used together. For example, Bill Gates and Obama's names get

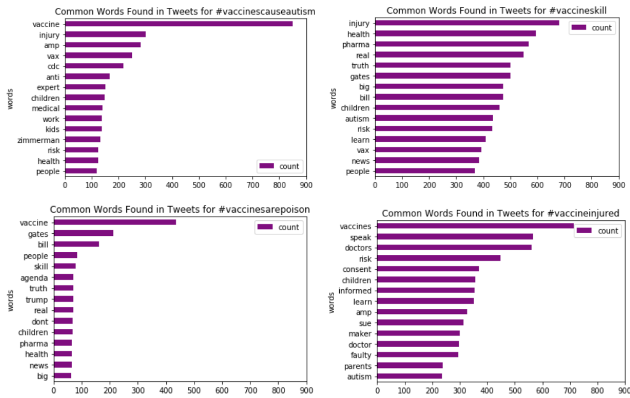

Figure 3: The number of tweets that contain the fifteen more frequently used words in the corpus of documents (tweets) for the hashtags #vaccinescauseautism, #vaccineskill, #vaccinesarepoison and #vaccineinjured

used with the word "evil" a lot in the bigrams for #vaccinesarepoison. These bigrams were created by looking at what words in our text occurred at the most together. Using the networkx library, the terms occurring as bigrams were visualized as graphs, where edges between two terms represent two words forming a bigram.

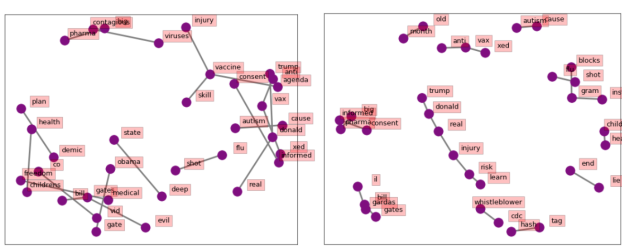

Figure 4: Bigram visualization for #vaccinesarepoison and #vaccineskill.

In #vaccineskill, we see an even clearer picture where Donald Trump is associated with many tweets and the term injury in terms of vaccination injuries. Also, terms that are often used in conspiracy theories are part of the graph such as Bill Gates, big pharma, and #cdcwhistleblower, a hashtag that was popularized through the movie "Vaxxed" which claimed that the CDC is hiding the negative effects of vaccination from the public.

From our examination, we found a group of selected accounts that are very active in the anti-vaccination community, while a big part of those who used the hashtags we examined only used them one or two times in total over the last three years. We found that there were a total of 184,208 hashtags used, of which 14,695 were unique hashtags. The most popular hashtags included #vaxxed and #vaccineswork.

Looking at the co-occurrence of hashtags, we found several hashtags related to pregnancy that appeared in our dataset, including #baby, #ObyGyn, and #pregnant. We

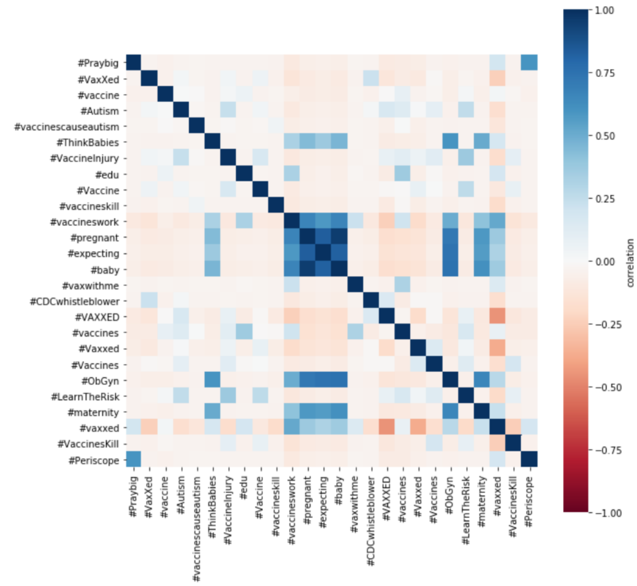

Figure 5: Correlation matrix of the most popular hashtags in our dataset. We can see that there are multiple maternity-related hashtags, which are also positively related to the hashtag #vaccineswork.

found that these pregnancy-related hashtags were positively correlated with the hashtag #vaccineswork as seen in Figure 5. In addition to examining the top hashtags, we also looked at the accounts mentioned in the collected tweets. The second most popular user that was mentioned was Donald Trump (@RealDonaldTrump). Among the fifteen most frequently mentioned user handles was also @POTUS and @YouTube. The rest of the top mentions were accounts belonging to individual users.

To better grasp what topics are popular in our collected tweets, we try to identify overarching topics using the Python library pyLDAvis that uses Latent Dirichlet Allocation, a type of topic modeling. In general, we can see that the anti-vaccination tweets are centered around three topics: institutions and doctors and their role in vaccine development. Another popular topic is different vaccines and the risks associated with them. A third big topic discusses the children, families, vaccines' effects, and the possible mandates surrounding vaccines.

Figure 6 offers examples of the top three topics in our datasets and corresponding examples. These examples illustrate the general finding that these tweets tend to emphasize the risk and negative side-effects of vaccines.

5. Discussion

5.1 Discussion of Results

In general, we can say that the COVID pandemic has played a role in an uptick in the number of tweets expressing anti-vaccination sentiment. It will be interesting to follow these hashtags further to see how their popularity evolves as the pandemic progresses. Overall we have seen that our data

Topic	Popular Terms	Example of tweets in topic
Children, vaccination risk and mandates	“parent”, “child”, “vaccine”, “mom”, “mandatory”, “exempt”	'So many of my parent friend's children are being injured and that's only what we see initally. Long term effects lead to other disease later in life. #vaccinescauseautism'
Doctors and vaccine development	“research”, “doctor”, “time”, “truth”, “immune”	'If all doctors were honest about the amount of education they receive in medical school about vaccination, parents would understand the absolute necessity to do their own research #humility #humble #vaxxed Are Doctors Experts On Vaccines? https://t.co/z9XNQcfofb via @YouTube'
Vaccines and the risks surrounding them	“risk”, “vax”, “cancer”, “mercury”, aluminum”, “death”	'#STUDY, 2016: #Chickenox Reduces the Risk of GLIOMA\n\nhttps://t.co/79XSUw9vSG\n\nThe beauty of NATURAL immunity.\nWith the #vaccination, we traded non-deadly childhood disease for an increased risk of brain #cancer.\n\n#LearnTheRisk #InformedConsent #VaccineInjured #WakeUp https://t.co/C5ZcQ1sKcA'

Figure 6: Illustration of the three main topics appearing in the corpus of tweets and corresponding terms and examples of tweets.

was centered around three main topics, including vaccination and families, doctors, and those creating vaccines and vaccines and disease in general. We also saw that the users that were most active within the anti-vaccination sphere had been active on Twitter for multiple years. In general, the tweets we collected followed temporal patterns that could be explained by the co-occurrence of events such as the COVID pandemic. In the future, it would be interesting to continue looking at future developments in this sphere.

5.2 Design suggestions

We have seen that platforms, including Twitter, can help heighten the spread of misinformation regarding vaccines and lead to more vaccine hesitancy in the general population. This can have severe effects, as it can lead to diseases that were previously thought to be eradicated to reemerge. Therefore, following the lead of other platforms, Twitter can start banning particular hashtags and limiting their spread. For example, Instagram has limited the appearance of Instagram posts that use specific hashtags in their search results, making posts that include disinformation harder to locate. Twitter could also take similar steps by creating posts that include hashtags such as #vaccineskill not appear in search results or not make them appear as the first results when searching for particular queries on Twitter. In addition, adding warning labels with links to sites that have accurate information on vaccination could help combat misinformation by providing users with additional sources that include real facts about vaccinations and their potential downsides. As of now, Twitter only has added such warning labels to the tweets of prominent figures such as Donald Trump. Using both labels, blocking, and suppressing tweets spreading misinformation could help Twitter combat the anti-vaccination community and the falsehoods they disseminate. At the same time, they have to be careful that censoring tweets based on specific criteria does not limit an individual's right to freedom of speech.

5.3 Limitations

One limitation that our research faces is that we have only identified five hashtags, which does not map the complete anti-vaccination community. These hashtags were also chosen by looking at which hashtags had gained particular traction on Instagram and, as a result, had been banned from the platform. However, we cannot equate Instagram with Twitter, and it may well be that the most popular hashtags may be significantly different on Twitter. The total number of tweets we collected is also small compared to other Twitter datasets making it hard to reach broader conclusions. By examining the frequency of tweets in our dataset, we were able to see that some hashtags had multiple days without tweets utilizing that particular hashtags. On the one hand, this might indicate that the vaccine community is not particularly active on Twitter. On the other hand, it may also be a sign that the choice of hashtags for our analysis could have been better. Therefore looking at more users and tweets would offer a better picture of the whole community of anti-vaccination supporters on Twitter. In addition, to ascertain hypotheses on the influence of trending news, e.g., the release of certain movies, a correlation study between generally trending news topics and the hashtags examined here would be worthwhile.

6. Conclusion and Future Work

Our research focused on examining tweets from five different hashtags over the time period of August 2017 to August 2020. We were able to visualize the number of tweets over time and saw no significant growth of anti-vaccination sentiment in the last few months of the COVID pandemic. However, the limited number of hashtags and corresponding data we collected only provides a small sample of the tweets within the anti-vaccination community. It may well be true that this does not fully picture the whole anti-vaccination community on Twitter. Thus there is more exploration and research surrounding anti-vaccination on Twitter that needs to be done to allow for broader conclusions. One next step that could be taken is to look at the users in our dataset and examine how far back their anti-vaccination sentiment goes on Twitter. In addition, examining the tweets from all users and looking for other anti-vaccination hashtags could be a next step to expand the number of hashtags and users we are examining. Moreover, correlating the tweet topics with news topics in the news media might provide additional information useful in ascertaining possible explanations. Despite their limitations, our collected data and subsequent analysis provide interesting insights into the anti-vaccination community on Twitter and the topics they interact with.

References

[Blankenship et al. 2018] Blankenship, E. B.; Goff, M. E.; Yin, J.; Tse, Z. T. H.; Fu, K.-W.; Liang, H.; Saroha, N.; and Fung, I. C.-H. 2018. Sentiment, contents, and retweets: A study of two vaccine-related twitter datasets. *The Permanente journal* 22. Publisher: Kaiser Permanente.

[Brooks 2014] Brooks, B. 2014. *Using Twitter data to identify geographic clustering of anti-vaccination sentiments*. PhD Thesis.

[Burki 2019] Burki, T. 2019. Vaccine misinformation and social media. *The Lancet Digital Health* 1(6):e258–e259. Publisher: Elsevier.

[DeMasi, Mason, and Ma 2016] DeMasi, O.; Mason, D.; and Ma, J. 2016. Understanding communities via hashtag engagement: A clustering based approach. In *Tenth International AAAI Conference on Web and Social Media*.

[Gunaratne, Coomes, and Haghbayan 2019] Gunaratne, K.; Coomes, E. A.; and Haghbayan, H. 2019. Temporal trends in anti-vaccine discourse on twitter. *Vaccine* 37(35):4867–4871. Publisher: Elsevier.

[Mahajan et al. 2019] Mahajan, R.; Romine, W.; Miller, M.; and Banerjee, T. 2019. Analyzing Public Outlook towards Vaccination using Twitter. In *2019 IEEE International Conference on Big Data (Big Data)*, 2763–2772. IEEE.

[Mitra, Counts, and Pennebaker 2016] Mitra, T.; Counts, S.; and Pennebaker, J. W. 2016. Understanding anti-vaccination attitudes in social media. In *Tenth International AAAI Conference on Web and Social Media*.

[Sanawi, Samani, and Taibi 2017] Sanawi, J. B.; Samani, M. C.; and Taibi, M. 2017. # vaccination: Identifying influencers in the vaccination discussion on twitter through social network visualisation. *International Journal of Business and Society* 18(S4):718–726. Publisher: Universiti Malaysia Sarawak.

[Unicef 2013] Unicef. 2013. Tracking Anti Vaccination Sentiment in Eastern European Social Media Networks. *New York: UNICEF*.

[van Schalkwyk, Dudek, and Costas 2020] van Schalkwyk, F.; Dudek, J.; and Costas, R. 2020. Communities of shared interests and cognitive bridges: the case of the anti-vaccination movement on Twitter. *Scientometrics* 1–18. Publisher: Springer.

[Yurieff 2019] Yurieff, K. 2019. Instagram blocks #VaccinesKill hashtag after CNN Business report - CNN.