

The Drift of #MyBodyMyChoice Discourse on Twitter

Cristina Menghini

Brown University

USA

cristina_menghini@brown.edu

Justin Uhr

Brown University

USA

justin_uhr@brown.edu

Shahrzad Haddadan

Brown University

USA

shahrzad_haddadan@brown.edu

Ashley Champagne

Brown University

USA

ashley_champagne@brown.edu

Björn Sandstede

Brown University

USA

bjorn_sandstede@brown.edu

Sohini Ramachandran

Brown University

USA

sohini_ramachandran@brown.edu

ABSTRACT

#MyBodyMyChoice is a well-known hashtag originally created to advocate for women's rights, often used in discourse about abortion and bodily autonomy. The Covid-19 outbreak prompted governments to take containment measures such as vaccination campaigns and mask mandates. Population groups opposed to such measures started to use the slogan "My Body My Choice" to claim their bodily autonomy. In this paper, we investigate whether the discourse around the hashtag #MyBodyMyChoice on Twitter changed its usage after the Covid-19 outbreak. We observe that the conversation around the hashtag changed in two ways. First, semantically, the hashtag #MyBodyMyChoice drifted towards conversations around Covid-19, especially in messages opposed to containment measures. Second, while before the pandemic users used to share content produced by experts and authorities, after Covid-19 the users' attention has shifted towards individuals.

KEYWORDS

twitter, covid, topic change, hashtags

ACM Reference Format:

Cristina Menghini, Justin Uhr, Shahrzad Haddadan, Ashley Champagne, Björn Sandstede, and Sohini Ramachandran. 2022. The Drift of #MyBodyMy-Choice Discourse on Twitter. In *14th ACM Web Science Conference 2022 (WebSci '22)*, June 26–29, 2022, Barcelona, Spain. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3501247.3531570>

1 INTRODUCTION

Many factors impact what and how users write on online social networks (OSNs). These factors include platform-level changes such as when Twitter changed the number of characters per tweet from 140 to 240 in 2017 [10], external events such as a presidential election [7, 28], social movements such as the Arab spring or the Black Lives Matter movement [15, 23], or crises such as earthquakes or terrorist attacks [3, 18]. In this paper, we focus on how users adopt hashtags for a different purpose than originally intended

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebSci '22, June 26–29, 2022, Barcelona, Spain

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9191-7/22/06...\$15.00

<https://doi.org/10.1145/3501247.3531570>



Figure 1: Word cloud of the hashtags that co-occur with #MyBodyMyChoice in the entire stream of tweets from 1-Jan-2018 to 20-Dec-2021.

or *hijack* hashtags. Understanding how users hijack hashtags is important for tracking and predicting social movements within public discourse.

The global health crisis resulting from the Covid-19 pandemic has impacted many aspects of daily life, ranging from health to economics, and has changed in certain ways how users behave on OSNs. Researchers have shown the role of OSNs in the distribution of false news [26], in part leading to our current *infodemic*. Many researchers have studied the effect of Covid-19 on Twitter, which is one of the most potent online platforms for distribution of information. Some research shows that during the Covid-19 pandemic, user behavior changed on Twitter and other platforms, such as Wikipedia [20, 22]. Gligorić et al. [11] shows that Twitter users’ attention shifted towards accounts associated to healthcare, science, government and politics as opposed to accounts related to religion or sports. As the impact of OSNs for spreading hate speech and false news have raised concern, it is vital that online social network administrators concentrate their efforts on excluding organized hate speech or the spread of false news. A drift in the usage of a hashtag can make it more difficult to identify false news or hate speech [5, 6].

On Twitter, discussions and topics are organized by usage of hashtags. When users adopt a hashtag that already exists for another purpose than originally intended, this is often called “hash-jacking” or referred to as hashtag drift or co-option. Researchers have studied hashjackings through a *linguistic perspective*, by analyzing the tweet itself, and through a focus on *user behaviours*, by researching how user groups interact with tweets [1, 5, 14, 25, 27] (see Section 3.3 for a full discussion).

In this work, we examine the hashtag #MyBodyMyChoice (Figure 1) to see how the content of the tweets associated with it evolved during the pandemic. The hashtag originated to support women’s rights around abortion and bodily autonomy, but during the pandemic a separate set of users utilized the hashtag to communicate how vaccine mandates related to Covid-19 restricted their bodily autonomy.

Our study shows that the hashtag #MyBodyMyChoice can no longer be uniquely associated to the original discourse centered around women’s rights. Indeed, after the Covid-19 outbreak there is an increasing number of monthly tweets, up to 75%, related to *covid & vaccines*. Although the discourse about #MyBodyMyChoice branches out into several directions, the characterization of users interactions on the corpora of tweets related to *women’s rights & abortion* and *covid & vaccines* suggests that the topic drift was not intentionally planned by a group of users. In Section 3.3, we provide a taxonomy of hashjackings including (1) semantic hashjacking; (2) hashjacking with polarization intents; (3) hashjacking to capitalize on the success of the original hashtag. We associate the case #MyBodyMyChoice to a semantic hashjacking, which involves a natural topic shift.

2 DATA

We collected the stream of #MyBodyMyChoice tweets from 1-Jan-2018 to 20-Dec-2021. For the purpose of the analysis, we break down the stream $S_{\#MyBodyMyChoice}$ into three equal-in-length periods to study the content of the hashtag before and after the Covid-19 outbreak. The three reference windows are: (1) Before pandemic (BP) from 1-Jan-2018 to the first officially reported case of Covid-19 outside China on 13-Jan-2020. (2) Initial pandemic (IP) from 13-Jan-2020 to the beginning of the vaccination campaign with the approval of Moderna vaccine by the FDA on 17-Dec-2020. (3) Covid-19 coexistence (CC) from 17-Dec-2020 to 20-Dec-2021.

Among the tweets in the stream, we retain the around 280K tweets written in English, corresponding to 86% of the entire retrieved corpus (excluding tweets from unidentified languages that are too short or contain too many hashtags to infer the language). It follows that throughout the analysis we assume that the content of the tweets mostly characterizes the ongoing debates in English-speaking countries.

2.1 Data Sub-topics

In this section, we explain how we identify the main sub-topics of the #MyBodyMyChoice hashtag, and how we assign a tweet to the respective sub-topic. Particularly, given the stream $S_{\#MyBodyMyChoice}$, we call a *sub-topic* the topic characterizing the content of a subset of the tweets in the stream.

Identify sub-topics. To characterize the presence of sub-topics we use hashtags. We refrain from the usage of topic-modeling algorithms for two main reasons: (1) In preliminary analysis, we found the hashtags to be representative of the ongoing discussions. Indeed, we observed that the text of tweets about either abortion or Covid-19 vaccination might contain the same set of words and differ only by hashtags: two examples are “I have the right to choose #MyBodyMyChoice #reproductiverights” versus “I have the right to choose #MyBodyMyChoice #novaccine”. (2) On

Sub-topic	Hashtags
<i>Women rights & abortion</i>	#prochoice, #womensrights, #abortion, #abortionishealthcare, #prolife, #abortionrights
<i>Politics & civil rights</i>	#blm, #blacklivesmatter, #lgbtq, #votethemout
<i>Freedom of choice</i>	#freedom, #medicalfreedom, #righttochoose, #freedomofchoice, #bodilyautonomy
<i>Covid</i>	#covid, #covid19, #nomorelockdowns, #freedomofchoice, #bodilyautonomy
<i>Vaccination</i>	#vaccine, #novaccinepassports, #vaccinesideeffects, #vaccinemandates, #iwillnotcomply

Table 1: For each sub-topic, we list some of the most relevant hashtags contained in their signature.

Twitter, users can navigate content using hashtags, so we will use them to quantify the ease of moving across sub-topics. Moreover, we point out that previous work focusing on change of topics also exploited hashtags [25].

Since 2018, more than 396K unique hashtags (excluding those consisting of a single character) have appeared in $S_{\#MyBodyMyChoice}$. To find the most relevant hashtags in the stream, we sample 100K tweets from the reference windows BP and the union of IP and CC¹. We take the union of the samples and extract a set R_i of hashtags such that at least 50% of the tweets in $S_{\#MyBodyMyChoice}$ contains at least one hashtag in R_i . We repeat the sampling 100 times and define the set of hashtags $R = \bigcup_{i=1}^{100} R_i$ composed of 141 hashtags covering 57.39% of the entire stream of tweets.

The list R can be grouped into the set S of five main sub-topics: *women rights & abortion*, *politics & civil rights*, *freedom of choice*, *covid*, and *vaccination*. We note that, for the periods after the pandemic started, we group the *covid* and *vaccination* sub-topics together, since we observed that (1) the two appear often together, and (2) there is no other vaccine debate open during the time in analysis. We soft-assign² the hashtags in the list R to each of these sub-topics. We refer to the set $\mathcal{T}_i$ of hashtags assigned to a sub-topic i as its *signature*. In Table 1, we list some of the hashtags contained in the signature of each sub-topic.

Assign tweets to sub-topics. The process of assigning tweets to sub-topics consists of two steps. In the first unsupervised step, given a tweet t , we denote by R_t the list of its hashtags that lie in R and assign t to the sub-topic i via

$$i = \operatorname{argmax}_{j \in S} \sum_{h \in R_t} \mathbb{P}(\mathcal{T}_j | h) \quad . \quad (1)$$

We apply this rule to all tweets that have at least one hashtag in R . Unfortunately, some tweets may not have any hashtags in R . To address this issue, we use the tweets labeled using Equation (1) as the ground truth to train a supervised classifier on the text of the remaining set of tweets to assign a tweet to a corresponding sub-topic. Our classifier consists of a Random Forest [2] of depth 35 for each time period, which is fed with the text of tweets represented as

¹Because the volume of the stream differs across reference windows, we sample with replacement to equally weight the hashtags appearing in different periods. The size of the sample is set to get a reasonable approximation of the real hashtags distribution.

²A hashtag can belong to multiple topics.

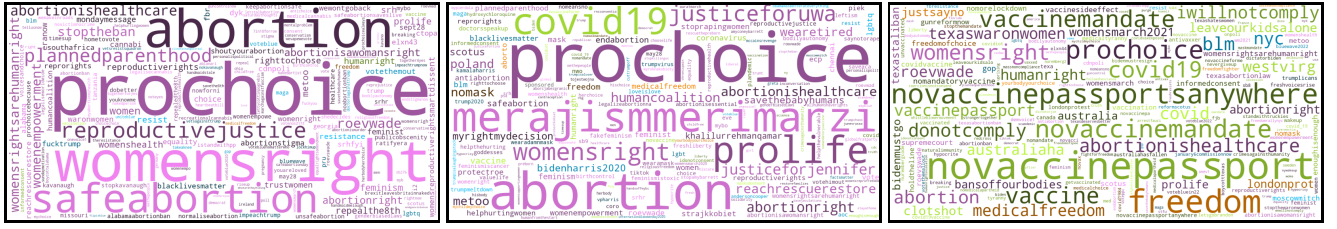


Figure 2: Hashtags in our reference windows. The plots show the most occurring hashtags appearing in tweets from BP, IP, and CC, from left to right. The larger and darker a word is, the more often it appears during the time window. Colors correspond to the respective topics: pink to *women rights & abortion*, green to *covid & vaccination*, light blue to *politics & civil rights*, and orange to *freedom of choice*.

a count vector³. The average test F1-score on the sub-topics is 82% for the BP period, 84% for the IP period, and 90% for the CC period. We decided not to use *covid-twitter-bert* [19] since a Random Forest classifier is an interpretable and simple model that already yields good results. Interpretability of our model is important since our ground truth is unsupervised, which provides extra flexibility if the labeled data we use require additional cleaning.

3 ANALYSIS

We want to understand whether and how the discourse around the hashtag #MyBodyMyChoice has changed after the Covid-19 outbreak by answering the following questions.

RQ1. How has the usage of #MyBodyMyChoice evolved?

RQ2. How do user bases and content interactions differ across sub-topics?

RQ3. Has #MyBodyMyChoice been hashjacked?

3.1 RQ1: #MyBodyMyChoice Evolution

Hashtags function as means of labeling, archiving and distributing information on OSNs. As such, we should not underestimate the importance of associating a topic to a hashtag. #MyBodyMyChoice was originally created to group content related to the *abortion* debate. Thus, an average consumer of information on Twitter would likely expect to find news and opinions about this topic when searching for content associated with the hashtag.

In our attempt to answer RQ1, we investigate if a consumer’s experience is likely to be consistent with this expectation, or whether the consumer is more likely to encounter tweets related to the online protests against Covid-19 vaccine mandates in which the slogan of “My Body My Choice” is widely used to claim their bodily autonomy.

#MyBodyMyChoice overview. For an overview of the discourse around #MyBodyMyChoice, we analyse the hashtags appearing in our three reference windows. In Figure 2, we observe that the hashtags used during the BP period relate more frequently to the discourse centered around women’s rights, and in particular abortion and body autonomy. During the IP period, women’s rights still appear central in the discussion, with a specific focus on asking for justice for women who were subjected to violence, e.g., #justiceforjennifer, #justiceforuwa, and #merajismmerimarzi;

at the same time, the hashtag #covid19 starts to gain relevance. During the CC period, corresponding to the first vaccine approval in the U.S., we see that the conversation about #MyBodyMyChoice is no longer centered around a topic associated to women’s rights: instead, and in contrast to the periods preceding the Covid-19 pandemic, multiple hashtags related to the vaccination campaign become more dominant.

Change of most relevant hashtags. We sample 100K tweets with replacement for each period BP, IP, and CC. Given one of these periods T , let H be the set of the 25 hashtags used most often within the period T and build the vector $\mathbf{v}_H^T \in \mathbb{R}^{25}$ whose entries are co-occurrences of #MyBodyMyChoice with each of the 25 hashtags in H . In the time period E following T , we count the co-occurrences of #MyBodyMyChoice with the same set H of hashtags, obtaining $\mathbf{v}_H^E \in \mathbb{R}^{25}$. To understand whether the frequency of co-occurrences of #MyBodyMyChoice with the hashtags is independent of the time window, we run a χ^2 test with a significance threshold of 5%. Applying this framework to the periods $T = BP$ and $E = IP$, we observe that the frequency of the hashtags is not independent of the time window. For the periods $T = IP$ and $E = CC$, we find similarly that the frequency of hashtags is not independent of the time window according to the χ^2 test.

Observation 1. The Covid-19 outbreak caused a decrease of occurrences of the 25 hashtags used most often before the pandemic. Similarly, the beginning of the vaccination campaign caused a change of occurrences of the hashtags that were most dominant during IC. For instance, the frequency of the hashtags #abortion, #prochoice, and #merajismmerimarzi was reduced by the 48.66%, 52.82%, and 98.35% respectively, while the hashtags #covid19, #vaccine, and #freedom increased in frequency by 97.73%, 422.66%, and 690.82% respectively.

Sub-topics monopolization. Following the process explained in Section 2.1, we separate the tweets into sub-topics. We refer to the fraction of tweets assigned to a sub-topic as the sub-topic’s monopolization factor. We note that, for the periods after the pandemic started, we grouped together *covid* and *vaccination* sub-topics, since we observed that (1) the two appear often together, and (2) there is no other vaccine debate open during the time in analysis. Figure 3 shows the monthly monopolization factor of each of the sub-topics. Consistently with our observations of the most dominant hashtags shown in Figure 2, the fraction of tweets about

³We evaluate each classifier on 5-folds (70/30) whose F-1 average is 83.75% (BP), 87.40% (IP), and 89.13% (CC).

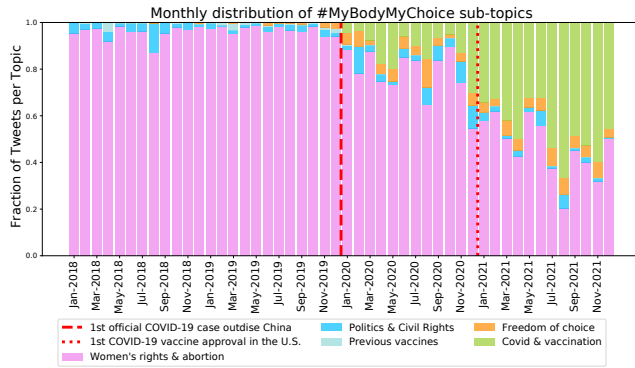


Figure 3: Sub-Topic Monopolization: for each sub-topic, we show the monthly fraction of tweets belonging to it.

women's right & abortion starts to decrease after the first official Covid-19 case was registered outside China, and the start of the vaccination campaign further reduces the original sub-topic of the hashtag #MyBodyMyChoice with the introduction of hashtags such as #novaccinepassport, #novaccinemandate, and #dontcomply. In fact, during the second half of 2021, the tweets about *women's right & abortion* represent less than 40% of the total number of tweets. Note, however, that the monopolization factor of *covid* is reduced after September 2021 when the abortion ban in Texas went into effect.

Observation 2. After the Covid-19 outbreak, the usage of the hashtag #MyBodyMyChoice changed on Twitter, with the conversation drifting towards Covid-19. In particular, the most relevant hashtags after the outbreak reflect opinions opposing the current vaccination campaign, typical of antivaxxers.

Sub-topics separation. So far, we have investigated how the original topic of *abortion* in the #MyBodyMyChoice stream was affected by the Covid-19 breakout. Next, we want to understand how connected or separated the four sub-topics are from each other. To this end, we built a hashtag network, where each node is a hashtag, and two hashtags are connected by an edge if they co-occur in at least 30 tweets. Each edge is weighted by the number of co-occurrences. Figure 4 shows a planar embedding of the resulting network in which nodes are colored by sub-topic. The figure indicates that each sub-topic corresponds to a separate set of hashtags. To quantify this phenomenon, we employ the Random-Walk Controversy score (RWC) [9]. The RWC score measures the degree of separation between two sets of nodes in a graph by comparing the probabilities of moving within and between the two sets. A negative score indicates that the two sets of nodes are not separated well so that the number of connections within and across the sets are comparable. When the RWC score is larger than zero, the two sets are well-separated, and there are likely very few connections between them relative to the number of connections within them. In our framework, the RWC score quantifies the overall separation of sub-topics by measuring the frequency with which hashtags associated with different sub-topics co-occur. On Twitter, hashtags



Figure 4: Network of hashtags in the #MyBodyMyChoice stream. We draw the largest connected components of the network. Pink nodes refer to the *women rights & abortion* sub-topic, green nodes to *covid and vaccination*, blue nodes to *politics & civil rights*, and the orange nodes to *freedom of choice*.

Sub-topic	Women's rights & abortion	Politics & civil rights	Freedom of choice	Covid & Vaccination
Women rights & abortion	-	0.49	0.32	0.46
Politics & civil rights	0.49	-	0.31	0.48
Freedom of choice	0.32	0.31	-	0.11
Covid & Vaccination	0.46	0.48	0.11	-

Table 2: For each pair of sub-topics, we report the RWC score. The closer the score is to one, the more separated the two sub-topics are. Note that the RWC score is symmetric.

are hyperlinks that allow users to navigate the associated content. Thus, in this context, RWC scores reflect how much users who navigate the content of a sub-topic are exposed to links related to content about other sub-topics. In particular, a lower RWC score implies that users who explore a sub-topic are less likely to reach content about the other sub-topic if they are navigating Twitter by following hashtags.

Table 2 reports the RWC scores for each pair of sub-topics. The *Women's rights & abortion* and the *covid & vaccination* sub-topics are the most separated, thus illustrating the topics shift of #MyBodyMyChoice. In contrast, *covid & vaccination* is close to *freedom of choice*, which, in turn, is close to *politics & civil rights*.

Observation 3. The *women's rights & abortion* and *covid & vaccination* sub-topics are well-separated. In particular, we expect that the hashtags characterizing the first sub-topic rarely co-occur with the hashtags representing the second sub-topic. Users who access content related to one of the two sub-topics therefore have a smaller probability of seeing hyperlinks to content related to the other sub-topic than to content of the same sub-topic.

3.2 RQ2: Compare Sub-Topics

The analysis in the previous section showed that Covid-19 has become a prominent sub-topic of the stream of tweets tied to #MyBodyMyChoice. In this section, our goal is to understand how

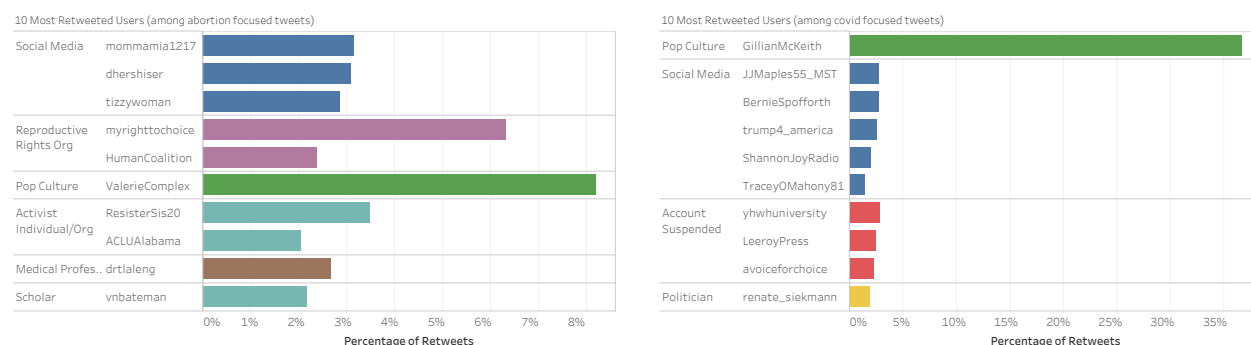


Figure 5: A list of the ten most retweeted users within the *women’s rights & abortion* and the *covid & vaccination* sub-topics. Percentages are computed with respect to the total number of retweets.

the sub-topics related to Covid-19 and abortion differ from each other in terms of their user bases and the popularity of content. We consider the tweets posted during the IP and CC periods, because the *covid & vaccination* sub-topic pop up after the beginning of the pandemic. The number of unique users that posted content about *women’s rights & abortion* is 1.82 times the number of users tweeting about *Covid-19 & vaccines*. Furthermore, 90.7% of users tweeted about only one of the two topics. Thus, we conclude that the user base of the two branches of #MyBodyMyChoice are separated.

Next, we investigate whether the number of posts, retweets, replies, and quotes, which constitute the different ways in which tweets are generated, differ between these two sub-topics and find that there are no significant differences: For the sub-topic *women’s rights & abortion*, 56.71% of tweets are retweets, 18.0% are replies, 7.34% are quotes, and the remaining are posts; the tweets that generate the retweets constitute 11.85% of the total tweets in this sub-topic. In the corpus of *covid & vaccination*, 55.94% of tweets are retweets, 20.92% are replies, 6.38% are quotes, and the remaining are posts; the tweets that generate the retweets make up 12.51% of all tweets in this sub-topic. Finally, we observe that the distributions of the number of tweets per user is comparable for both sub-topics: the median is one with few outliers in both populations. We speculate that the similar distribution across the sub-topics indicates that users tweeting about *covid & vaccination* were not part of an organized effort to intentionally shift the use of the hashtag #MyBodyMyChoice.

Observation 4. The user bases of *women’s rights & abortion* and *covid & vaccination* are different and display a similar characterization of content interaction (i.e., posts, retweets, replies, and quotes are distributed similarly). This indicates that there was no organized intervention to intentionally create a shift of the #MyBodyMyChoice discourse.

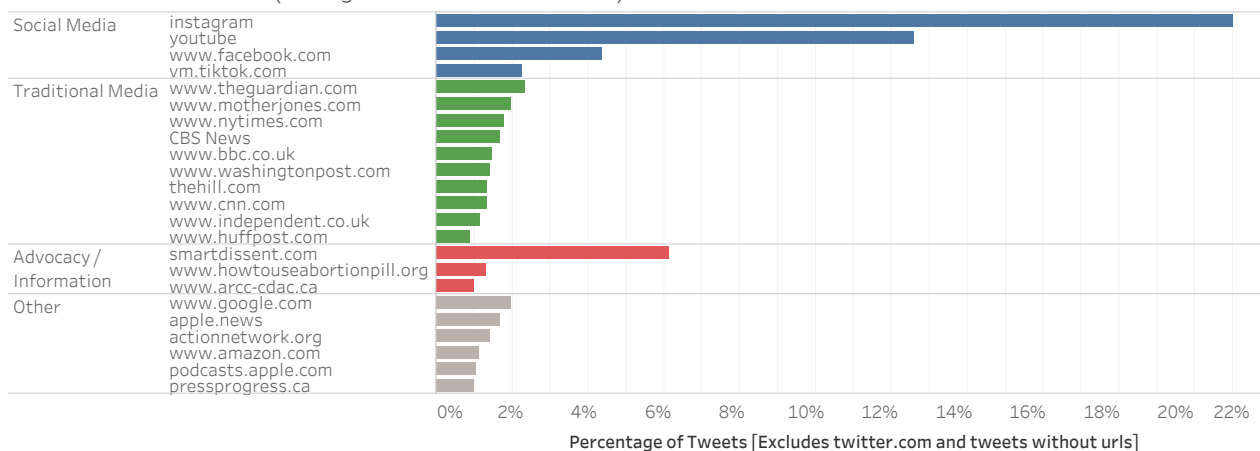
In Figure 5, we report the top-10 retweeted users appearing in tweets about *women’s rights & abortion* and *covid & vaccination*, respectively. The group of most frequently retweeted users focused on the topic of *women’s rights & abortion* are divided fairly evenly

between a few categories of users. These include activist and reproductive rights organizations, individuals from film, medical, and academic circles, and users known primarily for their social media presence. In contrast, the group of most retweeted users focused on the topic of *covid & vaccinations* includes more users from the social media category, fewer authority figures, and no organizations. In the latter sub-topic, there is also a striking gap between the most influential user’s share of retweets and everyone else’s: the user @GillianMcKeith has 37% of the retweets in this category, while the rest of the top ten have 1-3% each. Furthermore, three of the top ten accounts in the second sub-topic have been suspended.

Observation 5. The majority of the most retweeted users of *women’s rights & abortion* are organizations, activists, or professionals. In contrast, the set of accounts retweeted by the user base of *covid & vaccination* suggests that it gives more prominence to users that individually select, create, and share content without the support of established organizations.

In Figure 6, we report the URLs appearing in tweets from the *women’s rights & abortion* and *covid & vaccinations* sub-topics. The domains linked to in tweets from the *women’s rights & abortion* subcategory come primarily from three categories of sources; social media, traditional news outlets, and websites advocating for a position or providing information. The group of top domains linked to in tweets from the *covid & vaccinations* subcategory adds four additional social media sites to the four used by the abortion subcategory. These additional sites are known for having less restrictive censorship policies. Advocacy groups have a larger share of the links amongst the *covid and vaccination* subgroup, and they are a different set of organizations than those linked to by the women’s rights and abortion subcategory. Finally, traditional news outlets make up a much smaller portion of the top domains in the *covid and vaccinations* subcategory, and again the specific outlets are very different. In addition, one outlet known to disseminate ‘fake news’ is amongst the top twenty.

20 Most Linked Domains (among abortion focused tweets)



20 Most Linked Domains (among covid focused tweets)

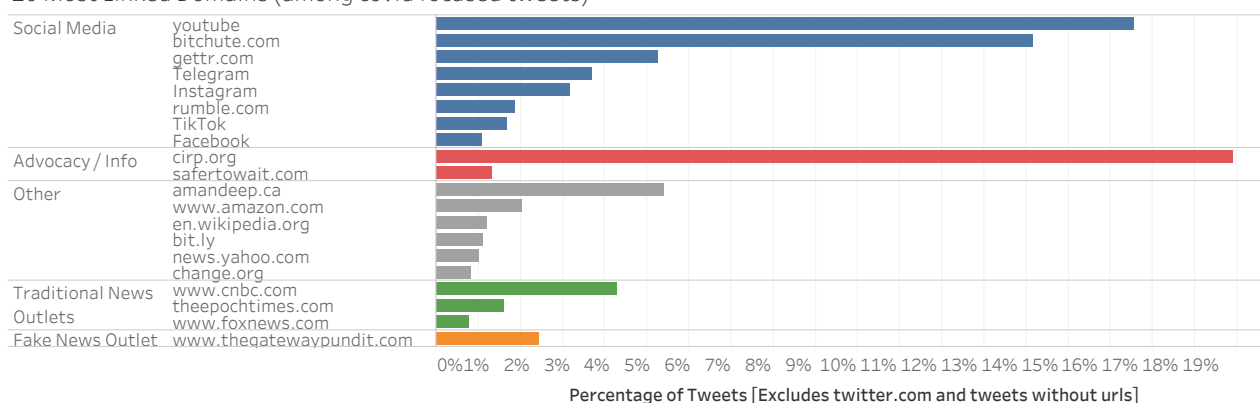


Figure 6: Listed are the domains of the most frequent URLs attached to tweets for the *women's rights & abortion* (top) and *covid & vaccination* (bottom) sub-topics. The percentages are computed over the total number of tweets that contain a non-Twitter URL. The percentage of tweets containing non-Twitter URLs is 33.1% for *women's rights & abortion* and 25.08% for *covid & vaccination*.

3.3 RQ3: Has #MyBodyMyChoice been hijacked?

There are a variety of different types of hashjacking, and this section situates the case of #MyBodyMyChoice in the context of other well-known hashjackings based on the analytical observations drawn so far. There are largely three types of hashjackings: those that semantically change the associated tweets, those that intentionally make a topic more polarized, and those that capitalize on the success of a given hashtag without taking part in the original conversation. These three categories are outlined below:

1) Hashjacking that semantically changes the associated tweets (whether intentional or not). When Alyssa Milano, in

2017, revitalized the #MeToo movement on Twitter by encouraging thousands of women to share their stories of sexual harassment, the movement garnered widespread global attention that did not attribute its founder, Tarana Burke, or use the language of the movement within its intended context of the Black community. As a result, #MeToo turned from a movement focused on Burke's community to women's experiences with sexual harassment broadly [24]. While the women responding to Milano's encouragement to post their own stories were not necessarily intentionally hashjacking Burke's words for the Black community, the result was that the semantic meaning of the hashtag and its associations changed.

2) Hashjacking that intentionally makes a topic more polarized. Within this category of hashjackings, users appropriate

a hashtag for the opposite purpose of the hashtag’s originally intended use. For example, McDonald’s launched a Twitter campaign using the hashtag #McDStories to encourage users to share their fond stories of Happy Meals. The hashtag took off as users posted critical comments about the food and how it impacted their health negatively. A similar example is the hashtag #Februdairy, which was launched by the dairy industry to support the dairy market after the creation of Veganuary, a challenge to join a global community to try vegan food for a month: users repurposed the hashtag to write about problems in the dairy industry.

3) Hashjacking to capitalize on the success of a given hashtag without taking part in the conversation. After #Metoo was popularized by Alyssa Milano, the hashtag was then also hashjacked in 2018 by the ‘120 decibels’ campaign by the Austrian Identitarian Movement in an attempt to latch onto to prominence of the #Metoo hashtag [16]. Similarly, the hashtag #Brazil2014 was hashjacked by ISIS in order to entice fans watching the World Cup in 2014 to click on links to the group’s propaganda videos [8]. In these hashjackings, the hashjackers did not intend to engage with the topic of the hashtag; rather they were solely using the hashtag to capitalize on its success.

Observation 6. #MyBodyMyChoice was a hashtag rooted in women’s rights and was then used by users during the pandemic to refer to something else entirely: anti-vaccine content. We situate #MyBodyMyChoice within the semantic hashjacking as (1) the discourse around the hashtag changed over time, and users who were hashjacking are distributed unique users rather than a given organization (Observations 1, 2, 4, & 5); (2) Users who hashjacked #MyBodyMyChoice used it to summarize their own positions on vaccines and masks rather than capitalize on the hashtag’s success. Furthermore, the hashtag was not hashjacked to make women’s reproductive rights more polarized but rather to express a different argument against COVID-19 vaccines (Observation 3).

4 RELATED WORK

We contextualize our work by providing a brief overview of previous studies about hashjacking. Some of these works analyse whether a hashtag has changed its usage (i.e., hashjacked), while some concentrate on developing tools to identify semantic shifts.

Past works emphasized the change of usage of a hashtag due to an event or by group of users. Rodriguez [21] showed that #FamilyTravelHacks, created to inform people about tips for safe travel with children, changed its usage after Executive Order 13769, a.k.a., the travel ban. As described in Section 3.3, Knüpfer et al. [16] demonstrated that the hashtag #metoo was hijacked by the ‘120 decibels’ campaign. Some other works analyzed how polarized users intentionally *hash-jack* a tweet to express their views contrasting the opinion of the original user base of the hashtag, e.g., in Germany’s presidential elections [5] or in U.S. politics [12].

To analyse these phenomena, several methodological works focus on designing computational tools to identify changes in the linguistic usage of a word [4, 13, 17] or of the meaning of a hashtag on Twitter [25]. Particularly, VanDam and Tan [25] use a matrix factorization-based method to build a day-by-day topic matrix and

use a standard change-detection hypothesis test to see if a change in the topic occurred. While they provide a strategy to say “if” the topic of a hashtag has changed or not, we focus on “how” the meaning of #MyBodyMyChoice has changed. For this reason, we adopted a qualitative approach instead of their method.

5 CONCLUSION

In this paper, we examined the hashtag #MyBodyMyChoice to see how the content of tweets associated with it evolved during the pandemic. The hashtag originated to support women’s rights around abortion and bodily autonomy, but we observed that during the pandemic a separate set of users utilized it to communicate their disagreement with vaccines.

We analysed the change of usage of #MyBodyMyChoice considering three factors: (1) the semantic of the discourse around the hashtag; (2) the users’ interactions with contents related to diverse sub-topics; and (3) the users’ attention toward specific users and external sources of information. We demonstrate the hypothesized semantic shift by employing supervised learning and statistical analysis. Furthermore, we quantify the shift by showing the evolution of dominant sub-topics. To understand whether the semantic shift was intentionally induced, we conducted a user behaviour analysis and found no evidence for such organized attempts. Finally our analysis of users’ attention to highly retweeted users and most commonly embedded URLs shows that attention has shifted from experts and authorities to individuals and that some unreliable sources appeared during the Covid-19 debate. We note that our analysis focused only on the stream of tweets using #MyBodyMyChoice and does therefore not necessarily generalize to the entire discourse of these topics. Moreover, although the methodology we used can be applied to other topics, it would still require human in the loop to first distinguish among sub-topics.

Disclaimer. It is important to note that the demographics of Twitter users is not representative of the overall population: 40% of Twitter users are in the 25-34 age group, another 20% are in the 35-49 age group, and the age groups 18-24 and above 50 each constitute around 17% of Twitter users (Statista); furthermore, 63% of Twitter users are male. Thus, as our study explores how the content associated with the #MyBodyMyChoice hashtag changed over the pandemic, it is important to keep in mind that our results may have been impacted by the demographics on Twitter.

ACKNOWLEDGMENTS

This work was a result of collaboration between Brown University’s Data Science Initiative, Brown University’s Computer Science Department and the Center for Digital Scholarship at Brown University. Haddadan was partially supported by the Data Science Initiative at Brown University and by the NSF via grant CCF-1740741. Ramachandran acknowledges support from US National Institutes of Health R01 GM118652 and R35 GM139628. Sandstede acknowledges support from the NSF grant CCF-1740741.

REFERENCES

- [1] Majid Alififi, Parisa Kaghazgaran, James Caverlee, and Fred Morstatter. 2018. Measuring the impact of ISIS social media strategy. *MIS2. Marina Del Ray, California* (2018).
- [2] Leo Breiman. 2001. Random Forests. In *Machine Learning*.

- [3] Christopher A Cassa, Rumi Chunara, Kenneth Mandl, and John S Brownstein. 2013. Twitter as a sentinel in emergency situations: lessons from the Boston marathon explosions. *PLoS currents* (2013).
- [4] Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd international conference on World Wide Web*.
- [5] Philipp Darius and Fabian Stephany. 2019. “Hashjacking” the Debate: Polarisation Strategies of Germany’s Political Far-Right on Twitter. In *International Conference on Social Informatics*.
- [6] Philipp Darius and Fabian Stephany. 2021. How the Far-Right Polarises Twitter: ‘Hashjacking’ as a Disinformation Strategy in Times of COVID-19. In *International Conference on Complex Networks and Their Applications*.
- [7] Gunn Enli. 2017. Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. *European journal of communication* (2017).
- [8] James P. Farwell. 2014. The Media Strategy of ISIS. *Survival* (2014).
- [9] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. Quantifying Controversy on Social Media. (2018).
- [10] Kristina Gligorić, Ashton Anderson, and Robert West. 2018. How constraints affect content: The case of Twitter’s switch from 140 to 280 characters. In *Proceedings of the International AAAI Conference on Web and Social Media*.
- [11] Kristina Gligorić, Manoel Horta Ribeiro, Martin Müller, Olesia Altunina, Maxime Peyrard, Marcel Salathé, Giovanni Colavizza, and Robert West. 2020. Experts and authorities receive disproportionate attention on Twitter during the COVID-19 crisis. *arXiv preprint arXiv:2008.08364* (2020).
- [12] Asmelash Tekla Hadgu, Kiran Garimella, and Ingmar Weber. 2013. Political Hashtag Hijacking in the U.S. (*WWW ’13 Companion*).
- [13] William L Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. *arXiv preprint arXiv:1605.09096* (2016).
- [14] Kohei Hayashi, Takanori Maehara, Masashi Toyoda, and Ken-ichi Kawarabayashi. 2015. Real-Time Top-R Topic Detection on Twitter with Topic Hijack Filtering (*KDD ’15*).
- [15] Jelani Ince, Fabio Rojas, and Clayton A Davis. 2017. The social media response to Black Lives Matter: How Twitter users interact with Black Lives Matter through hashtag use. *Ethnic and racial studies* (2017).
- [16] Curd Knüpfer, Matthias Hoffmann, and Vadim Voskresenskii. 2020. Hijacking MeToo: transnational dynamics and networked frame contestation on the far right in the case of the ‘120 decibels’ campaign. (2020).
- [17] Vivek Kulkarni, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2015. Statistically Significant Detection of Linguistic Change (*WWW ’15*).
- [18] Marcelo Mendoza, Barbara Poblete, and Carlos Castillo. 2010. Twitter under crisis: Can we trust what we RT?. In *Proceedings of the first workshop on social media analytics*.
- [19] Martin Müller, Marcel Salathé, and Per E Kummervold. 2020. Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter. *arXiv preprint arXiv:2005.07503* (2020).
- [20] Manoel Horta Ribeiro, Kristina Gligorić, Maxime Peyrard, Florian Lemmerich, Markus Strohmaier, and Robert West. 2021. Sudden attention shifts on wikipedia during the covid-19 crisis. In *Proc. Int. AAAI Conf. Web Soc. Media*.
- [21] Nathan J. Rodriguez. 2020. #FamilyTravelHacks: Humor and Political Commentary in Hashtag Hijacking.
- [22] Thorsten Rupprechter, Manoel Horta Ribeiro, Tiago Santos, Florian Lemmerich, Markus Strohmaier, Robert West, and Denis Helic. 2021. Volunteer contributions to Wikipedia increased during COVID-19 mobility restrictions. *Scientific reports* (2021).
- [23] Kate Starbird and Leysia Palen. 2012. (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work*. 7–16.
- [24] Times. 2018. I Founded ‘Me Too’ in 2006. The Morning It Went Viral Was a Nightmare. (2018). <https://time.com/6097392/tarana-burke-me-too-unbound-excerpt>
- [25] Courtland VanDam and Pang-Ning Tan. 2016. Detecting Hashtag Hijacking from Twitter (*WebSci ’16*).
- [26] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* (2018).
- [27] Ying Xiong, Moonhee Cho, and Brandon Boatwright. 2019. Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the #MeToo movement. *Public relations review* (2019).
- [28] Ussama Yaqub, Nitesh Sharma, Rachit Pabreja, Soon Ae Chun, Vijayalakshmi Atluri, and Jaideep Vaidya. 2020. Location-based sentiment analyses and visualization of Twitter election data. *Digital Government: Research and Practice* 1, 2 (2020), 1–19.