

VoterFraud2020: a Multi-modal Dataset of Election Fraud Claims on Twitter

Anton Abilov^{1,2}, Yiqing Hua^{1,2}, Hana Matatov³, Ofra Amir³, Mor Naaman^{1,2}

¹ Cornell Tech

² Cornell University

³ Technion

aa2776@cornell.edu, yiqing@cs.cornell.edu, hanama888@gmail.com, oamir@technion.ac.il, mor@jacobs.cornell.edu

Abstract

The wide spread of unfounded election fraud claims surrounding the U.S. 2020 election had resulted in undermining of trust in the election, culminating in violence inside the U.S. capitol. Under these circumstances, it is critical to understand the discussions surrounding these claims on Twitter, a major platform where the claims were disseminated. To this end, we collected and released the *VoterFraud2020* dataset, a multi-modal dataset with 7.6M tweets and 25.6M retweets from 2.6M users related to voter fraud claims. To make this data immediately useful for a diverse set of research projects, we further enhance the data with cluster labels computed from the retweet graph, each user’s suspension status, and the perceptual hashes of tweeted images. The dataset also includes aggregate data for all external links and YouTube videos that appear in the tweets. Preliminary analyses of the data show that Twitter’s user suspension actions mostly affected a specific community of voter fraud claim promoters, and exposes the most common URLs, images and YouTube videos shared in the data.

1 Introduction

Free and fair elections are the foundation of every democracy. The 2020 presidential election in the United States was probably one of the most consequential and contentious such events. Two-thirds of the voting-eligible population voted, resulting in the highest turnout in the past 120 years (Schaul, Rabinowitz, and Mellnik 2020). The Democratic Party candidate Joe Biden was elected as the president.

Unfortunately, efforts to delegitimize the election process and its results were carried out before, throughout and after the election (Election Integrity Partnership 2021). Claims of voter fraud, most of them unfounded (Frenkel 2020), were spread through public statements by politicians, by the media, and on social media platforms. As a result, 34% of Americans say that they do not trust the election results as of December, 2020 (NPR 2020). While prior research had shown that allegations of widespread voter fraud in the recent 20 years is not supported by credible evidence (Goel et al. 2020), claims of voter fraud have shown to have a significant negative impact on confidence in electoral integrity (Berlinski et al. 2021). Indeed, on January 6th, 2021,

believing that the election was ‘stolen’, mobs breached the U.S. capitol while the Congress voted to certify Biden as the winner of the election. In theory and in practice, unsubstantiated voter fraud allegations have great ramifications for the integrity of the election and the stability of democracies in the U.S. and beyond.

Social media platforms like Facebook, Twitter, YouTube and Reddit play a significant role in political events (Vittak et al. 2011; Allcott and Gentzkow 2017), and the 2020 election was no exception (Ferrara et al. 2020). In particular, Twitter has been the focus of public and media attention as a prominent public square where ideas are adopted and claims—false or true—are spread (Vosoughi, Roy, and Aral 2018; Grinberg et al. 2019). It is thus important to understand the participants, discussions, narratives, and allegations around voter fraud claims on this specific platform.

In this work, we release *VoterFraud2020*, a multi-modal Twitter dataset of 7.6M tweets and 25.6M retweets that are related to voter fraud claims. Using a manually curated set of keywords (e.g., “voter fraud” and “#stopthesteal”) that was further expanded using a data-driven approach, we streamed Twitter activities between October 23rd and December 16th, 2020. We performed various validations on the limits of our stream, given Twitter’s API constraints (Morstatter et al. 2013), and estimate that we were able to retrieve around 60% of the data containing our crawled keywords.

We further enhanced the *VoterFraud2020* dataset in order to make it accessible for a broader set of researchers and future research: (1) We cluster users into communities according to the network of their retweet patterns and release the community labels for each user; (2) Given Twitter’s widespread post-election suspension action, we crawl and include each user’s suspension status as of January 10th, 2021; (3) We compute and share the perceptual hashes of 168K images that appeared in the data; (4) We aggregate and share meta-data about 138K external links that appeared in the tweets, including 12K unique YouTube videos. Our dataset also allows researchers to calculate the amount of Twitter interactions with the collected tweets, users, and media items, including number of retweets and quotes from various clusters, or from suspended users.

A preliminary analysis finds a significant cluster of users who were promoting the election fraud related claims, with nearly 7.8% of them suspended in January. The suspensions

focused on a specific community *within* the cluster. A simple analysis of the distribution of images, based on visual similarity, exposes that the most broadly shared (by number of tweets) and the most retweeted images are different. Although recent research has shown that voter fraud claims are pushed mainly by mass media (Benkler et al. 2020), we also find that external links referenced by promoters of the claims point mostly to low-quality news websites, streaming services, and YouTube videos. Some of the most widespread videos claiming ‘evidence’ of voter fraud were published by surprisingly small channels. Most strikingly, all of the top ten channels and videos spreading voter fraud claims were still available on YouTube as of January 11th, 2021.

We believe that the release of *VoterFraud2020*, the largest public dataset of Twitter discussions around the voter fraud claims, with the enhanced labels and data, will help the broad research community better understand this important topic at a critical time.

2 Data Collection

Our data collection process involved streaming Twitter data using a data-driven manually curated set of keywords and hashtags. We report on the span and volume of the collected data, as well as on analyses estimating its coverage.

2.1 Streaming Twitter Data

We used a data-driven approach to generate a list of keywords and hashtags related to election fraud claims in an iterative manner. We started with a single phrase and two derived keywords: `voter fraud` and `#voterfraud`. We first used a convenience sample of 11M political tweets consisting of the tweets of 2,262 U.S. political candidates and the replies to those tweets, collected between July 21st and Oct 22nd, 2020 using the Twitter Streaming API (Twitter 2019). We then identified hashtags that co-occur with our meta-seed keywords, `voter fraud` and `voterfraud`. We selected all hashtags that appeared in at least 10 tweets and co-occurred with either of the meta-seed keywords at least 50% of the time. From the resulting set, we manually filtered out those that were not directly relevant to voter fraud. To this end, two members of the research team reviewed the hashtags, including, if needed, searching for them on Twitter to see whether they produce relevant results. Only the hashtags that were agreed on by both evaluators were added, resulting in an initial set of hashtags that was added to the two original keywords.

To find more related hashtags, we computed the Jaccard coefficient between each of our seed hashtags and all other hashtags that appeared in the new stream. We added to our set all hashtags that had a Jaccard coefficient greater than 0.001 with any of the seed hashtags. Three members of the team again reviewed this list by 1) excluding hashtags that were not related to voter fraud, 2) adding keywords corresponding to the hashtags (e.g. `#discardedballots` corresponds to `discarded ballots`), and 3) adding relevant hashtags or keywords that the researchers observed while searching for hashtags from the generated list. Both the seed list and the final list of keywords and hashtags we used for streaming are included in the Appendix (Table 4).

We collected data using the Twitter streaming API (Twitter 2019). The *VoterFraud2020* dataset includes tweets from 17:00, October 23rd, 2020 to 13:00 December 16th, 2020. We expanded the keywords list on Oct. 31st with additional keywords, and added `#stopthesteal` as it started trending on November 3rd. While streaming, we stored each tweet’s metadata (e.g., user ID, text, timestamp). We also downloaded all image media items included in the tweets. In total, we collected 3,781,524 original tweets, 25,566,698 retweets, and 3,821,579 *quote tweets* (i.e. tweets that include a reference to another tweet). Note that quote tweets are included in the Twitter stream when either the new tweet or the referenced (quoted) tweet include one of the keywords or hashtags on the list. In total, we collected tweets from 2,559,018 users who posted, shared or quoted one or more tweets with these keywords.

2.2 Coverage Analysis

Since the Twitter streaming API provides only a sample of the tweets, especially for large-volume keywords (Morstatter et al. 2013), we performed multiple tests to evaluate and estimate the coverage of the *VoterFraud2020* dataset. This analysis suggests that the dataset covers over 60% of the content shared on Twitter using the keywords we tracked.

Retweet and Quote Coverage. We evaluated the coverage of retweets and quote tweets by comparing the counts of these objects in the stream to Twitter’s metadata. When a new retweet for an original tweet appears in the stream, the API returns the tweet’s metadata including the current retweet count and quote count of the original tweet. In other words, if an original tweet t_i is retweeted, it will appear in the stream as a retweet rt_j , and the metadata for rt_j will include the total number of retweets of t_i so far. From this metadata, it is easy to define the *retweet coverage* as the ratio of the total number of retweets (rt objects) streamed and stored in our dataset, over the sum of all retweet counts of the original t tweets, returned by the API in the latest rt retweet of each original tweet. The quote coverage is defined analogously. According to this analysis, the *VoterFraud2020* dataset captured 63.2% of the retweets and 62.6% of the quote tweets. These findings compare favorably with previous work that shows a single API client captures only 37.6% of the retweets through the Streaming API (Morstatter et al. 2013).

Comparison with #Election2020. To further evaluate the coverage on the voter fraud tweets, we compared our dataset with a previously published Twitter dataset of the U.S. 2020 election (Chen, Deb, and Ferrara 2020). The creators of the *#Election2020* dataset used the streaming API to track 168 keywords that are broadly related to the election and 57 accounts that are tied to candidates running for president.

As in *VoterFraud2020*, the keyword ‘voter fraud’ was also used to collect data for *#Election2020*. We used this overlap to estimate our coverage. First, we can directly compare the relative volume and overlap between the ‘voter fraud’ tweets in both datasets. We expect our *VoterFraud2020* to have a

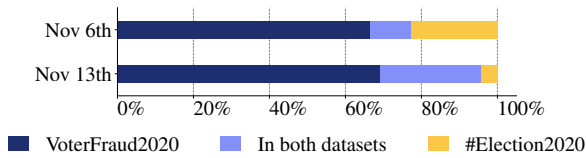


Figure 1: Coverage comparison between our dataset and #Election2020 for tweets containing ‘voter fraud’.

higher volume of such tweets because of its more focused set of keywords. Second, if we assume sampling for both streams is independent and random, we could estimate the coverage of *VoterFraud2020* by looking at the proportion of #Election2020 tweets that also appear in our data.

To this end, we extracted all tweets and retweets that contain this keyword from both datasets posted on two days following the November 3rd election data: November 6th and November 13th. The analysis, performed on December 17th, was limited to two days as we had to obtain the content of the tweets of the #Election2020 dataset by “hydrating” them (i.e. using the tweet IDs to get the full tweet text using the Twitter API). We were unable to hydrate the full data, presumably due to inactive accounts and deleted tweets. The hydration yielded 92.4% of the #Election2020 data from November 6th (a total of 1.4M tweets/3.5M retweets), and 91.1% of the data from November 13th (1.3M tweets/3M retweets).

In total, our *VoterFraud2020* data includes 45,322 ‘voter fraud’ related tweets on November 6th, 2.3 times as much as recorded in #Election2020. The ratio is even higher on November 13th, when we obtained 47,313 tweets, 3.1 times as much as in #Election2020. Figure 1 breaks down the coverage by dates (separated by rows), in the two datasets (by different colors). From left to right, the bars show the percentages of tweets that are available only in our dataset (dark blue), that are available in both datasets (light blue), and that are available only in #Election2020 (yellow). On any given day, the *VoterFraud2020* dataset contains substantially more tweets related to voter fraud, especially when the estimated total volume is lower. On November 13th (second row), *VoterFraud2020* contained 95.7% of the combined data (left two bars) while #Election2020 only collected 30.7% (right two bars) of the tweets. These numbers also indicate that *VoterFraud2020*’s sample includes 32.1% of the related samples in #Election2020 on November 6th and 85.9% on November 13th. We acknowledge that these two numbers are not consistent, presumably because of November 6th’s much higher volume of activity. If these samples are indeed independent, though, it means that our lower bound of coverage is November 6th’s 32.1%.

Based on these coverage analyses, we conclude that *VoterFraud2020* is, at the time of submission, the largest known public Twitter dataset of voter fraud claims and discussions.

3 Data Enhancement

To ensure the reusability of our data, we took the following steps to enhance the raw streaming data. We performed a

community analysis of users according to the retweet graph, and release the community labels. Given Twitter’s large-scale suspension of accounts and the public interest in those actions, we also queried Twitter for each user’s status on 10th of January, and share the user status as active/not-found/suspended. Furthermore, to enable research on spread of images and visual misinformation, we encode all images shared in the tweets with perceptual hash that allows for easy comparison and retrieval of similar content in the data. Finally, we release the set of URLs that appeared in the dataset, as well as the YouTube metadata for each YouTube video URLs.

Retweet Graph Communities. Retweet networks have been frequently analyzed in previous work in order to understand political conversations on Twitter (Arif, Stewart, and Starbird 2018; Cherepnalkoski and Mozetič 2016). Using community detection algorithms, researchers are able to study key players, sharing patterns, and content on different sides of a discussion surrounding a heated political topic.

To compute these communities, we first constructed a retweet graph of the *VoterFraud2020* dataset, where nodes represent users and directed edges correspond to one user (the target of the edge) retweeting another (the source). Edges are weighted according to the number of times the corresponding source user has been retweeted by the target user. The resulting network consists of 1,887,736 nodes and 16,718,884 edges.

To detect communities within the graph, we used the Infomap community detection algorithm (Bohlin et al. 2014), which captures the flow of information in directed networks. Using the default parameters, the algorithm produces thousands of communities. By excluding all communities that contain fewer than 1% of the nodes we are left with 90% of all nodes¹ which are clustered into five communities (see Table 1).

In Figure 2a, we visualize the retweet network using the Force Atlas 2 layout in Gephi (Bastian, Heymann, and Jacomy 2009), using a random sample of 44,474 nodes and 456,372 edges. The nodes are colored according to their computed community as described in Table 1. Edges are colored by their source node. The visualization indicates that the nodes are split between two distinct clusters: community 0 (blue) on the left and communities 1, 2, 3 and 4 on the right. By examining the top users in each community, we conclude that community 0 mostly consists of accounts that tend to refute and *detract* from the voter fraud claims, while the communities on the right consist of accounts that *promote* the voter fraud claims. For brevity, in the following analyses, we refer to the cluster on the left as the *detractor* cluster, and the cluster with community 1,2,3,4 on the right as the *promoter* cluster.

Community 2 is more deeply embedded in the *promoter* cluster compared to Community 1, as we observe tweets from Community 1 being retweeted by Community 0 on the left, but not from Community 2. Most of the tweets

¹Since the graph only includes retweeting and retweeted users, this number corresponds to 73.8% of all users in our dataset.

Community	Users	% of users suspended	Tweets	Images tweeted
0	860,976	1%	1,199,587	30,506
1	437,783	4.6%	644,219	14,423
2	342,184	14.1%	3,982,990	94,115
3	33,857	1.5%	27,699	867
4	23,414	1.6%	24,191	753

Table 1: Community statistics.

from in our data are written in English (not surprisingly, as we were tracking English language keywords), except for users in Community 3 who mainly post tweets in Japanese and users in Community 4 who write mostly in Spanish. We include the list of top five Twitter accounts in each community by the number of community retweets in the Appendix.

Note that due to the partisan nature of the U.S. politics, most of the *promoter* users are likely aligned with right-leaning politics, and *detractor* users align with left-leaning politics. By inspecting the top 10 retweeted domains in each cluster (Table 2), and correlating the sources with political alignment evaluations by AllSides² and Media Bias/Fact Check³, we confirm that 90% of the top news sources shared by the *promoter* cluster are right-leaning, and 90% of the top news sources shared by the *detractor* cluster are left-leaning.

To identify users that are prominent within each of these two clusters, we calculate the closeness centrality of the user nodes in each cluster. In a retweet network this metric can be interpreted as a user’s ability to spread information to other users in the network (Okamoto, Chen, and Li 2008). We compute the top-k closeness centrality to find the 10,000 most central nodes within the *detractor* and *promoter* clusters (Bisenius et al. 2018).

The dataset includes users’ community labels and their closeness centrality score in their cluster (*detractor* and *promoter* clusters). We also include additional community-based metrics for the tweet: *retweet count by community X* and *quote count by community X*. For a tweet t_i , the *retweet count by community X* is the total number of retweets rt_i it received from each user u_X in community X . The metric is also computed, analogously, for quote tweets.

Labeling Suspended and Deleted Users When the electoral college was set to confirm the election results on January 6th, 2021, the allegations of voter fraud took a dramatic turn, which culminated in the storming of the US Capitol. Subsequently, Twitter took a harder stance on moderating content on their platform and suspended at least 70,000 accounts that were engaged in propagating conspiracy theories and sharing QAnon-content (Twitter 2021). This ban has substantial implications for researchers seeking to understand the spread of voter fraud allegations on Twitter, since the Twitter API does not allow the “hydration” of Tweets from suspended users. In order to understand the distribution of suspensions in our dataset we queried the updated

²allsides.com

³mediabiasfactcheck.com

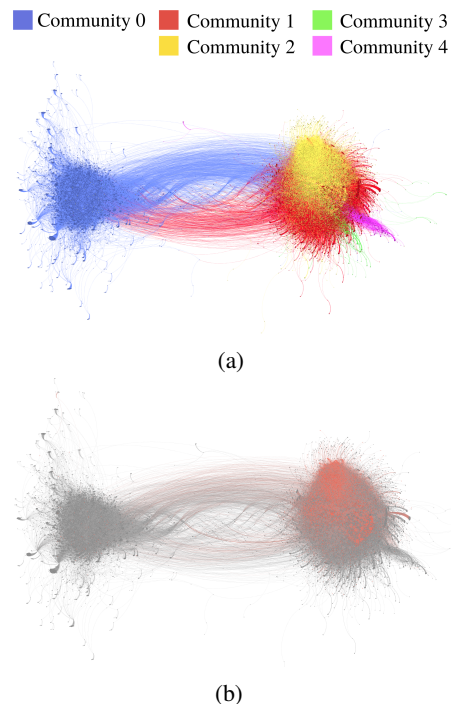


Figure 2: (a) Retweet graph colored by communities. (b) Suspension status (orange: suspended users).

user status of all users in our dataset on January 10th, a few days following the ban. The Twitter API returns a user status that indicates if the user is active, suspended or not found (presumably deleted). In total, 3.9% of the accounts (99,884 accounts) in our data were suspended.

In Figure 2b, we color the nodes in the randomly sampled retweet graph according to each user’s suspension status (suspended users in orange). The visualization shows how Twitter’s suspension efforts primarily targeted users within the *promoter* cluster. In our data, we find that 88.3% of the suspended users that were part of the top five communities were part of this cluster. Moreover, the figure shows that the suspensions greatly overlap with Community 2 in Figure 2a. The Data Analysis section below provides further details about this overlap and the suspended users.

We enhance the *VoterFraud2020* dataset by labeling tweets and users that were suspended. This metadata will enable research on the suspensions, and ease hydration of the data from Twitter by allowing hydraters to skip content that is no longer available. Relatedly, we also include two additional metrics for each tweet: *retweet count by suspended users* and *quote count by suspended users*.

Due to its immense public interest, we have retained the *full* data we retrieved from the 99,884 suspended users including 1,240,405 tweets and 6,246,245 retweets. This detailed data is not part of *VoterFraud2020*. However, we will distribute an anonymized version of this data to published academic researchers upon request.

Images. Because of their persuasive power and ease of spread, there is a growing interest in analyzing how visual misinformation spreads both within a platform or across platforms (Zannettou et al. 2018; Highfield and Leaver 2016; Paris and Donovan 2019; Moreira et al. 2018; Zannettou et al. 2020). However, visual information such as images or videos is difficult for many researchers to study due to computational and storage costs. Here, we make the information about image content shared in *VoterFraud2020* easier to use by researchers by sharing *perceptual hash values* for these images (Petrov 2017; Zauner, Steinebach, and Hermann 2011). Common perceptual hashes are binary strings designed such that the Hamming distance (Zauner, Steinebach, and Hermann 2011) between two hashes is close if and only if the two corresponding images are *perceptually* similar. In other words, an image that is only slightly transformed, for example, by re-sizing, cropping, or rotation, will have a similar hash value to the original image.

With these numeric hash values, researchers can easily find duplicates and near-duplicate images in tweets, without working directly with cumbersome image content. To this end, we download all image media items that were posted in the tweets in the streaming data, and encode them with three different types of perceptual hashes. As the definition of perceptual similarity is often subjective and the underlying algorithms are often different, various hash functions have different performance characteristics dealing with various types of image transformations. Therefore, we encode the images in our dataset with *three* perceptual hash functions: the Perceptive Hash (pHash), the Average Hash (aHash), and the Wavelet Hash (wHash) (Petrov 2017; Zauner, Steinebach, and Hermann 2011).

In total, our streamed tweets included 201,259 image URLs, 167,696 of them were successfully retrieved during streaming. We provide some more details about the distribution of these images in Section 5.

External Links. Misinformation campaigns are known to use broad cross-platform information, often via links to other sites (Wilson and Starbird 2020; Golovchenko et al. 2020). Hence, we extracted and publish the set of external (non-Twitter) URLs that were referenced in the tweets. For ease of use, we resolved URLs that point to a redirected location (e.g. bit.ly URLs) to their final destination URL. Our streamed tweets included references to a total of 138,411 unique URLs, appearing in 609,513 tweets.

Since a large portion (over 12%) of all URLs in the data were YouTube links, we further enhanced the data with YouTube-specific metadata. A key motivation for this specific focus was the known role YouTube plays generally in spreading misinformation (Hussein, Juneja, and Mitra 2020; Papadamou et al. 2020) and specifically its role in the 2020 election and voter fraud claims (Kaplan 2020; Wakabayashi 2020). For each YouTube video that was shared in the collected tweets, we used YouTube’s Data API (YouTube 2021), to retrieve the video’s title, description, as well as the id and title of the channel that posted it. We retrieved the YouTube metadata on Jan 1st, 2021. On that data, out of the

13,611 unique video ids that we have queried, 1,608 were no longer available resulting in 12,003 YouTube URLs with full additional metadata.

4 Data Sharing and Format

Our *VoterFraud2020* dataset is available for download under FAIR principles (Wilkinson et al. 2016) in CSV format⁴. The data includes “item data” tables for tweets, retweets, and users keyed by Twitter assigned IDs and augmented with additional metadata as described below. The data also includes the list of images that appear in the dataset, indexed by randomly generated unique IDs. Finally, the data includes aggregated tables for URLs and for YouTube videos including the information described in Section 3. The dataset tables and the fields they contain are summarized on Github⁵.

The *VoterFraud2020* dataset conforms with FAIR principles. The dataset is *Findable* as it is publicly available on Figshare, with a digital object identifier (DOI): 10.6084/m9.figshare.13571084. It is also *Accessible* since it can be accessed by anyone in the world through the link. The dataset is in csv format, hence it is *Interoperable*. We release the full dataset with descriptions detailed in this paper, as well as an online tool to explore the dataset at <https://voterfraud2020.io>, making the dataset *Reusable* to the research community.

The tables for Tweets and Retweets contain the full set of items that were collected, including from suspended users. These tables do not include raw tweet data beyond the ID, according to Twitter’s ToS. However, to support use of the data without being required to download (“hydrate”) the full set of tweets, we augment the Tweets table with several key properties. For each tweet we provide the number of total retweets as computed by Twitter (`retweet/quote_count_metadata`), as well as the number of retweets and quotes we streamed for this tweet from users in each of the five main communities (`retweet/quote_count_community_X`, X ranging from 0 to 4). Note that the latter do not add up to the Twitter metadata retweet count due to the coverage issues listed in Section 2.2. The Tweet table properties also include the `user_community` (0–4) for the user who posted the tweet, computed using methods listed in Section 3. Some of the Twitter accounts were not clustered into one of the five main communities. In this case, the `user_community` label is `null`. With this augmentation, researchers using this dataset could very quickly, for example, select and then hydrate a subset of the most retweeted tweets from non-suspended users in Community 2. As the tweet itself and the ID of the user who tweeted it is not available until hydration, Twitter’s users’ privacy is preserved.

The Users table is similarly augmented with aggregate information about the importance of the user in the dataset, including the community that they belong to, their centrality in the two meta-clusters, *detractor* and *promoter* (`closeness centrality detractor cluster`

⁴<https://figshare.com/account/projects/96518/articles/13571084>

⁵<https://github.com/sTechLab/VoterFraud2020>

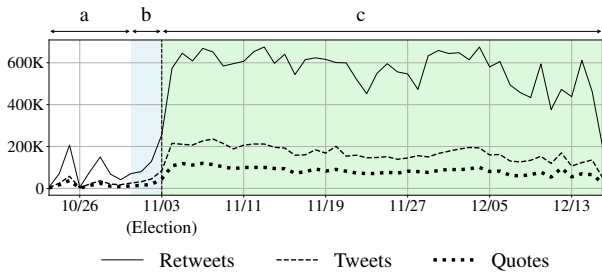


Figure 3: Temporal overview of the dataset showing number of streamed tweets, quotes and retweets per day. The shaded regions mark the expansions of the keyword set.

and `closeness centrality promoter cluster`), and the amount of attention (retweets and quotes) they received from other users in the different communities. We also note whether, according to the data we collected, the user had been suspended. With this data, interested researchers can quickly focus their attention and research on the main actors in each community.

The Images table includes a list of all the image media items retrieved in the stream, their unique media ID, and the ID of the tweet in which the image was shared. We augment this table with the image hash using three types of perceptual hash functions – aHash, pHash and wHash, as detailed in Section 3. This augmentation, together with the link to the Tweet ID, will allow researchers to quickly identify and hydrate popular images using the tweet metadata. Researchers also have the option of identifying matches in the data for images that come from any other source, by computing and comparing the perceptual hash values for new query images.

The two aggregate tables—the URLs table and the YouTube Videos table— provide information about the popularity of each of these objects in the dataset: aggregate retweet and quote counts for the URL or video link, both using the Twitter metadata and the count of objects in our stream from the various communities. In addition, these tables are augmented with metadata about the item (URL or YouTube video) as noted in Section 3.

5 Data Analysis

We performed a preliminary analysis of our dataset and its different modalities – tweets and users, images, external links – to demonstrate its potential interest and provide some initial guiding insights about the data.

Tweets and Users. Figure 3 shows the amount of retweets (solid), original tweets (dashed) and quote tweets (dotted) in the *VoterFraud2020* dataset over the time (X-axis) of the data collection. Three shaded regions, from left to right, mark the expansion of our set of keywords on October 31st (light blue, region b) and November 3rd (light green, region c). The Y-axis specifies the daily count. In general, except for the large increase after the election date (November 3rd, dotted vertical line), the volume of the stream remains roughly the same. On average, there are

170,938 tweets, 576,136 retweets, and 85,488 quote tweets per day after the election.

Our manual inspection shows that top tweets retweeted by the *detractor* cluster often condemn the alleged voter fraud claims, while top tweets on the *promoter* cluster indeed make voter fraud claims. Not surprisingly, among the top ten most retweeted tweets in the *promoter* cluster, nine were tweeted by President Trump. We refer readers to our project website for more details about popular tweets.

While the *promoter* cluster seems rather homogeneous (Figure 2a), users in Community 2 (yellow) stand out in both their level of activity and the rate in which they were suspended. Community 2 was highly active in our dataset. For example, Community 2 comprises 18.1% of the users, but contributed 68% of the *VoterFraud2020* tweets, and 74% of the retweets. Moreover, 14% of Community 2’s users were suspended by Twitter by the time we collected the account status data as described above, a much higher rate than the other communities, as shown in Table 1. In total, Community 2 was responsible for 46.1% of all suspensions amongst the users we associated with the top five communities. The suspension effect, and its focus on Community 2, can also be observed in Figure 2b, when compared visually against Figure 2a.

Promoters of the QAnon conspiracy theory were heavily involved in spreading unsubstantiated voter fraud claims (Tollefson 2021). We conduct preliminary analysis to evaluate the QAnon presence in the *VoterFraud2020* dataset and in the suspensions, based on early reports that suggested Twitter’s suspensions were targeting QAnon users (Conger 2021). We curated a set of QAnon-related hashtags⁶, and identified 50,385 users associated with QAnon by counting the users that have either tweeted or retweeted one of the hashtags, or mentioned the hashtag in their profile description. Of these users we find that 52.4% users have been suspended as of January 10th, providing indication that Twitter’s suspensions focused on the QAnon community. We find that for these QAnon users for whom we had network data, 82.7% were part of community 2, where suspension rates were highest. Further, the rate of QAnon hashtags in Community 2 was 5 to 99 times higher than other communities in the retweet graph.

A full analysis of the suspended accounts and their network communities, and the potential impact of the suspension is out of scope for this dataset paper, but can be easily performed using the data we share in *VoterFraud2020*. For example, the data shows that 35% of the *promoter* cluster users that were retweeted more than 1,000 times (1,596 in total) were suspended.

To conclude, our preliminary analysis shows that alleged election fraud claims mostly circulate in the *promoter* cluster, and in particular in the most active Community 2 within the cluster. The recent moderation efforts from Twitter seem

⁶The set of QAnon hashtags used in our analysis: #awake, #cabal, #calmbeforethestorm, #cbts, #enjoytheshow, #greatawakening, #neonrevolt, #outoftheshadows, #patriqts, #pizzagate, #q, #qanon, #qmovie, #qproofs, #savethechildren, #stqrm, #thegreatawakening, #theshow, #thestorm, #wga, #wwg, #wwg1wga

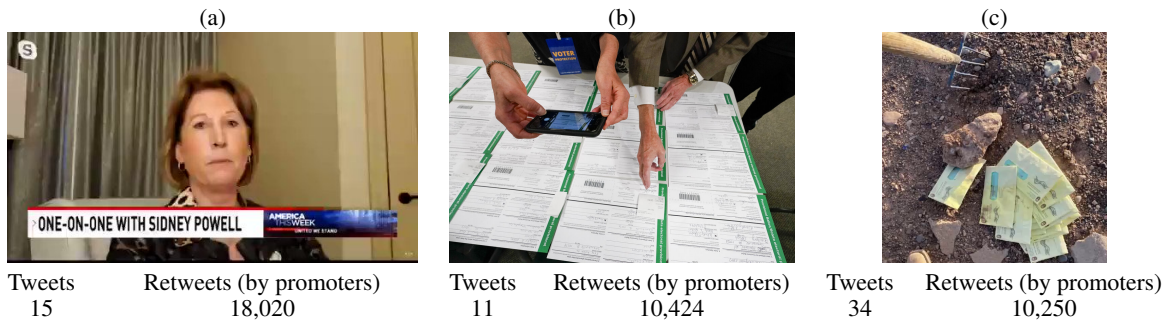


Figure 4: Top three most retweeted images in the *promoter* cluster: (a)–(c), with the number of unique tweets they appeared in and the number of retweets by users in the *promoter* cluster. Image (c) was cropped to fit the figure.

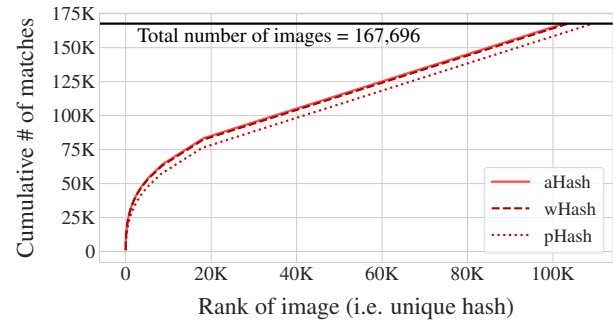
to have affected this active community, and did *not* broadly target all accounts involved in promoting such claims. There is some indication that Community 2 has significant overlap with followers of the QAnon conspiracy theory.

Images. We conducted a preliminary examination of matching and repeated images in *VoterFraud2020* to analyze the distribution of images related to voter fraud claims. Our data, using the perceptual hash functions described in Section 3, allows tracking of duplicate and near-duplicate images that were posted in multiple tweets. In this analysis, we experimented with three perceptual hash functions and refer to two images as matching if they have an *identical* perceptual hash value.

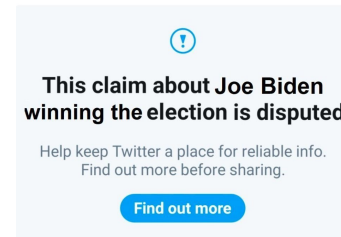
For example, there are 109,312 (out of 167,696) images with the same pHash value. Of these, 17,831 were shared in more than one tweet, an average of 4.27 times. In other words, 34% of the image instances in *VoterFraud2020* tweets appear in more than one tweet. Figure 5b presents the image that appeared in the most unique tweets: the image (based on the perceptual hash value) appeared in over 1,000 tweets, according to all three hash functions.

Figure 5a shows the cumulative distribution of the number of unique perceptual hashes in *VoterFraud2020* (Y-axis), with hash values sorted based on the number of unique tweets in which they appear, from the highest to the lowest (X-axis). For example, according to pHash, the 1,000 images shared in the largest number of unique tweets appeared together in 25,019 different tweets (not including retweets). Although in general the results are similar when using different hash functions, pHash is the most “conservative” of the three in terms of assigning matches. Overall, our results are similar when using different hash functions.

We further investigate the popularity of images defined by the number of retweets within the *promoter* cluster. After grouping images by the same pHash value, we present in Figure 4 the top three images that have been retweeted in the *promoter* cluster. Also note that despite the high values of cluster retweets, all these “popular” images appeared in only a few original tweets in our data. For example, image (a) appeared in 15 tweets. These tweet’s total retweet count based on the Twitter metadata (as returned from the API) counts add up to 24,399. The images were retweeted (as recorded in our dataset) from users in the *promoter* cluster



(a)



(b)

Figure 5: (a) The cumulative number of repeated images by hash matches. (b) The most tweeted image.

18,020 times. We note that images (a) and (b) were also the top two images retweeted by users in the *suspended* users set, with 5,547 and 3,122 retweets in that set, respectively (recall that almost all suspended users belong to the *promoter* cluster).

As expected, the pattern of retweeted images in the *detractor* cluster is quite different. The three most retweeted images in the *detractor* cluster (not included for lack of space) have somewhat lower spread, appearing in tweets that were retweeted 10743, 6425, and 3411 times (based on metadata). The top image is a screenshot of the NY Times front page of Nov 11th, 2020 reporting that top election officials across the country have not identified any fraud.

The analysis presented above can be easily extended with less-strict image similarity matching by calculating the

<i>promoter</i> cluster		<i>detractor</i> cluster	
Domain	Retweets	Domain	Retweets
pscp.tv	51,822	washingtonpost.com	11,220
youtube.com	44,031	rawstory.com	9,267
thegatewaypundit.com	35,967	cnn.com	4,139
davidharrisjr.com	18,793	independent.co.uk	3,882
foxnews.com	17,332	nytimes.com	3,746
theepochtimes.com	15,297	newsweek.com	3,496
thedcpatriot.com	14,958	news.yahoo.com	2,899
thefederalist.com	13,288	deadstate.org	2,409
djhjmedia.com	11,816	theguardian.com	2,232
justthenews.com	11,149	politicususa.com	2,032

Table 2: Top 10 domains being retweeted in the *promoter* and the *detractor* clusters respectively, as well as the number of retweets by users in these clusters.

Hamming distance between a pair of perceptual hash values. In this initial analysis, we used a strict sense of similarity, treating images as similar only when they share the exact same perceptual hash values.

URLs. We conduct preliminary analyses of the external links that have been included in the *VoterFraud2020* tweets. Table 2 lists the top 10 domains that have been shared inside the *detractor* and *promoter* clusters respectively. Most of the links shared by users in the *detractor* clusters are to mainstream news media, such as the Washington Post, CNN, and the New York Times. The rest are other news-related websites. The links shared by users in the *promoter* cluster mostly point to less authoritative news sources. Through manual inspection, we find that 50% of the top domains shared by the *promoter* cluster were evaluated as low-quality news sources by either MediaBiasFactCheck⁷ or by Grinberg et al. (2019).

The most shared domain in the *promoter* cluster is pscp.tv, a live video streaming app owned by Twitter. YouTube stands out as the second most retweeted domain among the *promoter* users. This trend is reflected in multiple news reports, warning of the significant role that YouTube plays in spreading false information related to voter fraud claims (Frenkel 2020). The majority of the top 10 most retweeted videos by the *promoter* users falsely claim evidence of widespread election fraud. The users spreading these videos had significant overlap with the subsequent suspension action by Twitter. For eight of the top 10 most retweeted videos by the *promoter* cluster, 29%-42% of the retweets of tweets sharing those videos were by accounts later suspended by Twitter.

A scan of the top 10 YouTube channels retweeted in the *promoter* cluster shows that they were relatively large (millions of subscribers), though there are also several smaller channels. For example, the most retweeted channel, Precinct 13, has only 3.67K subscribers, but has a video that appeared in 88 tweets and has been retweeted over 9K times.

Despite YouTube’s announcement that it will take actions against content creators who falsely claim the existence of

widespread voter fraud⁸, as of Jan 11th, the top 10 channels and videos listed in our tables are still available on YouTube.

6 Related Work and Datasets

We review prior work using Twitter data analysing politically related events, with an emphasis on those that have released a public dataset.

In particular, prior research had heavily used and published Twitter data to study U.S. elections. Using tweets collected during the 2016 U.S. election, researchers have analysed information operations run by social bots (Rizoïu et al. 2018), characterized the dissemination of misinformation (Vosoughi, Roy, and Aral 2018) and the exposure of American voters to misinformation (Grinberg et al. 2019). Work in Hua, Naaman, and Ristenpart (2020); Hua, Ristenpart, and Naaman (2020) characterized adversarial interaction against political candidates during the 2018 U.S. general election and shared 1.7M tweets interacting with political candidates.

Focusing on the U.S. 2020 election, research studied false claims regarding mail-in ballots (Benkler et al. 2020) before the election as the COVID-19 pandemic made it hard to vote in person. Closest to our work is the *#Election2020* dataset (Chen, Deb, and Ferrara 2020), which streamed a broad set of Twitter data for both political candidates’ tweets and related keywords. As discussed above, although some of the voter fraud related keywords were included in their data collection process, our *VoterFraud2020* dataset contains more than 2.3 times as much of the related data in *#Election2020*, for the ‘voter fraud’ keyword, presumably because of our more focused stream. Our stream also included a broader set of fraud-claim related keywords.

In order to help understand the dissemination of misinformation across platforms, Brena et al. (2019); Hui et al. (2018) used news articles as queries and released the tweets pointing to these articles. In 2018, Twitter published a list of accounts that the platform suspects to be related to Russia’s government controlled Internet Research Agency (Twitter 2018). This release enabled a number of studies that deepened our understanding of foreign information manipulation

⁷mediabiasfactcheck.com

⁸see:twitter.com/YouTubeInsider/status/1347231471212371970

in the U.S. (Arif, Stewart, and Starbird 2018; Im et al. 2020; Badawy, Ferrara, and Lerman 2018).

Most of the previous works that released Twitter datasets only included the tweet IDs, in accordance with Twitter’s Terms of Service. We keep to that practice, and augment the data without sharing tweet content, as detailed above, making our multi-modal dataset more accessible and useful to the research community.

7 Discussion and Conclusions

The unsubstantiated voter fraud claims spread on Twitter and elsewhere around the U.S. 2020 presidential elections are likely to form one of the most consequential misinformation campaigns in modern history. It is critical to allow a diverse set of researchers to provide a deeper understanding of this effort, which will continue to have national and global impact for years to come. To enable that contribution, it is important to provide a public and accessible archive of this campaign on various social media platforms, including Twitter as we do in *VoterFraud2020*.

The *VoterFraud2020* dataset has the potential to benefit the research community, and to further inform the public regarding both Twitter user activities around the voter fraud claims, as well as Twitter’s own response. Yet, the data has some limitations. We could not possibly capture the full extent of the voter fraud claims on Twitter, as our dataset was constructed by using matching keywords. Further, as analyzed above, we do not have full coverage even for the keywords we tracked, though we estimate that we have a majority of the tweets with those keywords. Nevertheless, the breadth of the data enables various types of investigation using both the tweet data, as well as the aggregated data of URLs, videos and images used in the campaign. We propose three major categories of such investigation.

First, researchers can use the dataset to study the spread, reach, and dynamics of the voter fraud campaign on Twitter. Researchers can describe and analyze the participants, including the activities of political candidates using information from orthogonal datasets of candidate accounts⁹, or the interaction between public figures and other accounts spreading claims and promoting certain narratives. Further, the data can help expose how different public figures spread different claims, for example the claims regarding the Dominion voting machines, what kind of engagement such narratives received. The data can also be used to understand the role of bots and other coordinated activities and campaigns in spreading this information. In general, the dataset can provide for analysis of the distribution of attention to these claims and how it spreads – via images, tweets, URLs – including comparison among different pre-computed communities and clusters.

Second, we include auxiliary data – URLs including YouTube links, and image hashes – that can help researchers examine *other* sources of information and their role in spreading the voter fraud claims. For example, using the

image hash values that were encoded using publicly available algorithms, researchers can easily map images not just within the Twitter data, but also from the larger web and media ecosystem. Researchers may combine our dataset with datasets that are collected from other social media platforms to examine how visual misinformation spread across platforms (e.g., (Zannettou et al. 2018; Moreira et al. 2018)).

A third potential area of investigation is Twitter’s response to this campaign of spreading voter fraud claims. Twitter’s civic policy and in particular its approach to misinformation around the election results and claims of voter fraud had been rated as “non-comprehensive” and was generally not applied consistently and transparently (Election Integrity Partnership 2021). The *VoterFraud2020* dataset can help better understand Twitter’s response. A specific question is the characterization of the suspended users whom as we shown above are primarily part of a specific community, related to QAnon. Researchers can use the data to both understand Twitter’s non-public response and its effectiveness, or even simulate the effectiveness of hypothetical earlier bans of the same population or users. As noted above, while Twitter’s terms of service do not allow us to publicly sharing full data for the suspended users—the *VoterFraud2020* tweets for these users are no longer available on Twitter by their ID—we will make these tweets available privately to published academic researchers, as we believe these tweets are of immense and justified public interest.

The publicly released *VoterFraud2020* data was collected and made available according to Twitter’s Terms of Service for academic researchers, following established guidelines for ethical Twitter data use (Rivers and Lewis 2014). By limiting to the Tweet IDs as the main data element, the dataset does not expose information about users whose data had been removed from the service. The only content in our data that is directly tied to a Tweet ID is the hash of the images for tweets that included them. Even though that hash, theoretically, can be tied to an image from another source, in absence of the original tweet the image will not be associated with any user account. We believe that this minor disclosure risk is justified given the potential benefits of this data.

Acknowledgements

We thank the anonymous reviewers for their helpful feedback. This research was funded by NSF research grant CNS-1704527.

⁹For example, <https://github.com/vegetable68/Midterm-2020-candidates>

Appendix

Community 0		
Handle	Active Status	Retweets
kylegriffin1	active	76,302
mmpadellan	active	74,393
donwinslow	active	69,796
BernieSanders	active	60,961
AriBerman	active	58,222
a) Community 1		
realDonaldTrump	suspended	1,560,373
LLinWood	suspended	1,057,805
SidneyPowell1	suspended	633,273
GenFlynn	suspended	334,197
CodeMonkeyZ	suspended	274,210
b) Community 2		
DonnaWR8	suspended	38,388
zeusFanHouse	suspended	36,347
LeahR77	suspended	33,352
TheRISEofROD	suspended	32,992
Bubblebathgirl	active	27,787
c) Community 3		
ganaha_masako	active	12,480
KadotaRyusho	active	6,890
yamatogokorous	active	5,716
mei98862477	active	5,347
kohyu1952	active	5,244
d) Community 4		
FernandoAmandi	active	4,217
POTUS_Trump_ESP	active	2,981
TDN_NOTICIAS	active	2,459
1VAFI	active	1,802
Gamusina77	active	1,638

Table 3: Top 5 Users in each community sorted by retweets from other users.

Seed list

#abolishdemocratparty
 #ballotharvesting #ballotvoterfraud
 #cheatingdemocrats #democratvoterfraud
 #gopvoterfraud #ilhanballotharvesting
 #ilhanomarballotharvesting
 #ilhanomarvoterfraud #mailinvoterfraud
 #stopvoterfraud #voterfraud #voterfraudbymail
 #voterfraudisreal

Filtered out

#abolishdemocratparty

Generated from the seed list

#ballotharvesting #voterid
 #ilhanomarforprison #stopgopvoterfraud
 #ilhanomar #nancypelosiabusingpower
 #nancypelosimustresign #junkmailballots
 #traresforcongress #immigrationfraud
 #votebyemailfraud #ballotfraud #exposed
 #votersuppression #ilhanresign #voteinperson
 #votebyemail #video #lockherup #nomailinvoting
 #ilhanomarelectionfraud #taxfraud
 #ballotharvesting #massivemailinballots
 #arrestilhanomar #obamagate
 #ilhanomarlockherup #buyingvotes
 #2020election #campaignfraud #homewrecker
 #voteinperson #minneapolis #absenteeballots
 #trump2020 #arrestilhanomar #absenteeballot
 #darktolight #wwg1wga #terrorist
 #daveygravyspirualsavage #trump #fraud #liar
 #pizzagate #republicans #qproof #theawakening
 #voteatthepolls #marriedherbrother
 #glasshouses #sheepnomore #voteyouout
 #cheater #georgesoros #georgia #vote
 #walkaway #thegreatawakening #qanon #evil
 #savethechildren

Keywords list 10/24

#ballotfraud #ballotharvesting
 #ballotvoterfraud #cheatingdemocrats
 #democratvoterfraud #ilhanomarballotharvesting
 #ilhanomarvoterfraud #mailinvoterfraud
 #nomailinvoting #stopgopvoterfraud
 #stopvoterfraud #votebyemailfraud #voterfraud
 #voterfraudisreal

Added on 10/31

#discardedballots #electionfraud
 #electioninterference #electiontampering
 #gopvoterfraud #hackedvotingmachines
 'destroyed ballots' 'discarded ballots'
 'election fraud' 'election interference' 'election
 tampering' 'hacked voting machine' 'pre-filled
 ballot' 'stolen ballots' 'ballot fraud' 'ballot
 harvesting' 'cheating democrats' 'democrats
 cheat' 'harvest ballot' 'vote by mail fraud' 'voter
 fraud'

Added on 11/03

#stopthesteal

Table 4: Hashtags and keywords related to election fraud.

References

- Allcott, H.; and Gentzkow, M. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31(2): 211–36.
- Arif, A.; Stewart, L. G.; and Starbird, K. 2018. Acting the part: Examining information operations within# BlackLivesMatter discourse. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW): 1–27.
- Badawy, A.; Ferrara, E.; and Lerman, K. 2018. Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 258–265. IEEE.
- Bastian, M.; Heymann, S.; and Jacomy, M. 2009. Gephi: an open source software for exploring and manipulating networks. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, volume 3.
- Benkler, Y.; Tilton, C.; Etling, B.; Roberts, H.; Clark, J.; Faris, R.; Kaiser, J.; and Schmitt, C. 2020. Mail-In Voter Fraud: Anatomy of a Disinformation Campaign. Technical report, Berkman Center Research Publication. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3703701.
- Berlinski, N.; Doyle, M.; Guess, A. M.; Levy, G.; Lyons, B.; Montgomery, J. M.; Nyhan, B.; and Reifler, J. 2021. The Effects of Unsubstantiated Claims of Voter Fraud on Confidence in Elections. Technical report. URL <https://cpb-us-e1.wpmucdn.com/sites.dartmouth.edu/dist/5/2293/files/2021/03/voter-fraud.pdf>.
- Bisenius, P.; Bergamini, E.; Angriman, E.; Meyerhenke, H.; and Oct, D. S. 2018. Computing Top- k Closeness Centrality in Fully-dynamic Graphs. *Proceedings of the Twentieth Workshop on Algorithm Engineering and Experiments (ALENEX) SIAM*: 1–26.
- Bohlin, L.; Edler, D.; Lancichinetti, A.; and Rosvall, M. 2014. Community Detection and Visualization of Networks with the Map Equation Framework. In *Measuring Scholarly Impact: Methods and Practice*, 3–34. Springer International Publishing.
- Brena, G.; Brambilla, M.; Ceri, S.; Di Giovanni, M.; Pierri, F.; and Ramponi, G. 2019. News sharing user behaviour on twitter: A comprehensive data collection of news articles and social interactions. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13.
- Chen, E.; Deb, A.; and Ferrara, E. 2020. #Election2020: the first public Twitter dataset on the 2020 US presidential election. Technical report. URL <https://arxiv.org/abs/2010.00600>.
- Cherepnalkoski, D.; and Mozetič, I. 2016. Retweet networks of the European Parliament: evaluation of the community structure. *Applied Network Science* 1(1): 1–20.
- Conger, K. 2021. Twitter, in Widening Crackdown, Removes Over 70,000 QAnon Accounts. <https://nytimes.com/2021/01/11/technology/twitter-removes-70000-qanon-accounts.html>. [accessed 12-Jan-2021].
- Election Integrity Partnership. 2021. The Long Fuse: Misinformation and the 2020 Election. <https://purl.stanford.edu/tr171zs0069>. [v1.2.0; accessed 30-Mar-2021].
- Ferrara, E.; Chang, H.; Chen, E.; Muric, G.; and Patel, J. 2020. Characterizing social media manipulation in the 2020 US presidential election. *First Monday* 25(11).
- Frenkel, S. 2020. How Misinformation ‘Superspreaders’ Seed False Election Theories. <https://nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html>. [accessed 4-Jan-2021].
- Goel, S.; Meredith, M.; Morse, M.; Rothschild, D.; and Shirani-Mehr, H. 2020. One person, one vote: Estimating the prevalence of double voting in US presidential elections. *American Political Science Review* 114(2): 456–469.
- Golovchenko, Y.; Buntain, C.; Eady, G.; Brown, M. A.; and Tucker, J. A. 2020. Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube During the 2016 US Presidential Election. *The International Journal of Press/Politics* 25(3): 357–389.
- Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; and Lazer, D. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363(6425): 374–378.
- Highfield, T.; and Leaver, T. 2016. Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice* 2(1): 47–62.
- Hua, Y.; Naaman, M.; and Ristenpart, T. 2020. Characterizing twitter users who engage in adversarial interactions against political candidates. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.
- Hua, Y.; Ristenpart, T.; and Naaman, M. 2020. Towards Measuring Adversarial Twitter Interactions against Candidates in the US Midterm Elections. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14.
- Hui, P.-M.; Shao, C.; Flammini, A.; Menczer, F.; and Ciampaglia, G. L. 2018. The Hoaxy misinformation and fact-checking diffusion network. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12.
- Hussein, E.; Juneja, P.; and Mitra, T. 2020. Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW): 1–27.
- Im, J.; Chandrasekharan, E.; Sargent, J.; Lighthammer, P.; Denby, T.; Bhargava, A.; Hemphill, L.; Jurgens, D.; and Gilbert, E. 2020. Still out there: Modeling and identifying russian troll accounts on twitter. In *12th ACM Conference on Web Science*, 1–10.
- Kaplan, A. 2020. YouTube has allowed conspiracy theories about interference with voting machines to go viral. <https://mediamatters.org/google/youtube-has-allowed-conspiracy-theories-about-interference-voting-machines-go-viral>. [accessed 5-Jan-2021].

- Moreira, D.; Bharati, A.; Brogan, J.; Pinto, A.; Parowski, M.; Bowyer, K. W.; Flynn, P. J.; Rocha, A.; and Scheirer, W. J. 2018. Image provenance analysis at scale. *IEEE Transactions on Image Processing* 27(12): 6109–6123.
- Morstatter, F.; Pfeffer, J.; Liu, H.; and Carley, K. M. 2013. Is the sample good enough? Comparing data from twitter’s streaming API with Twitter’s firehose. *Proceedings of the International AAAI Conference on Weblogs and Social Media*.
- NPR. 2020. Poll: Just A Quarter Of Republicans Accept Election Outcome. <https://npr.org/2020/12/09/944385798/poll-just-a-quarter-of-republicans-accept-election-outcome>. [accessed 5-Jan-2021].
- Okamoto, K.; Chen, W.; and Li, X.-y. 2008. Ranking of Closeness Centrality for Large-Scale Social Networks. In *Frontiers in Algorithms*, 186–195. Springer Berlin Heidelberg.
- Papadamou, K.; Zannettou, S.; Blackburn, J.; De Cristofaro, E.; Stringhini, G.; and Sirivianos, M. 2020. “It is just a flu”: Assessing the Effect of Watch History on YouTube’s Pseudoscientific Video Recommendations. *arXiv preprint arXiv:2010.11638*.
- Paris, B.; and Donovan, J. 2019. Deepfakes and Cheap Fakes. Technical report, Data & Society. URL <https://datasociety.net/library/deepfakes-and-cheap-fakes/>.
- Petrov, D. 2017. Wavelet image hash in Python. <https://fullstackml.com/wavelet-image-hash-in-python-3504fdd282b5>. [accessed 15-Jan-2021].
- Rivers, C. M.; and Lewis, B. L. 2014. Ethical research standards in a world of big data. *F1000Research* 3(38).
- Rizoiu, M.-A.; Graham, T.; Zhang, R.; Zhang, Y.; Ackland, R.; and Xie, L. 2018. #debatenight: The role and influence of socialbots on twitter during the 1st 2016 us presidential debate. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12.
- Schaul, K.; Rabinowitz, K.; and Mellnik, T. 2020. 2020 turnout is the highest in over a century. <https://washingtonpost.com/graphics/2020/elections/voter-turnout/>. [accessed 5-Jan-2021].
- Tollefson, J. 2021. Tracking QAnon: how Trump turned conspiracy-theory research upside down. *Nature* 590: 192–193. doi:10.1038/d41586-021-00257-y.
- Twitter. 2018. Update on Twitter’s review of the 2016 US election. https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html. [accessed 7-Jan-2021].
- Twitter. 2019. Twitter Standard API. <https://developer.twitter.com/en/docs/tweets/filter-realtime/overview>. [accessed 15-Jan-2019].
- Twitter. 2021. An update, following the riots in Washington, DC. https://blog.twitter.com/en_us/topics/company/2021/protecting--the-conversation-following-the-riots-in-washington--.html. [accessed 12-Jan-2021].
- Vitak, J.; Zube, P.; Smock, A.; Carr, C. T.; Ellison, N.; and Lampe, C. 2011. It’s complicated: Facebook users’ political participation in the 2008 election. *CyberPsychology, behavior, and social networking* 14(3): 107–114.
- Vosoughi, S.; Roy, D.; and Aral, S. 2018. The spread of true and false news online. *Science* 359(6380): 1146–1151.
- Wakabayashi, D. 2020. Election misinformation continues staying up on YouTube. <https://nytimes.com/2020/11/10/technology/election-misinformation-continues-staying-up-on-youtube.html>. [accessed 5-Jan-2021].
- Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3.
- Wilson, T.; and Starbird, K. 2020. Cross-platform disinformation campaigns: lessons learned and next steps. *Harvard Kennedy School Misinformation Review* 1(1).
- YouTube. 2021. YouTube Data API — Google Developers. <https://developers.google.com/youtube/v3>. [accessed 8-Jan-2021].
- Zannettou, S.; Caulfield, T.; Blackburn, J.; De Cristofaro, E.; Sirivianos, M.; Stringhini, G.; and Suarez-Tangil, G. 2018. On the origins of memes by means of fringe web communities. In *Proceedings of the Internet Measurement Conference 2018*, 188–202.
- Zannettou, S.; Caulfield, T.; Bradlyn, B.; De Cristofaro, E.; Stringhini, G.; and Blackburn, J. 2020. Characterizing the Use of Images in State-Sponsored Information Warfare Operations by Russian Trolls on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14.
- Zauner, C.; Steinebach, M.; and Hermann, E. 2011. Rihamark: perceptual image hash benchmarking. In *Media watermarking, security, and forensics III*, volume 7880, 78800X. International Society for Optics and Photonics.