# Tracking Evolution of Coordinated Activity on Twitter

**Master's Thesis**

*submitted to*

**Indian Institute of Science Education and Research Tirupati**

*in partial fulfillment of the requirements for the*

BS-MS Dual Degree Programme

by

**Padinjaredath Suresh Vishnuprasad**

Roll Number: 201701003

Supervisor: Prof. Dr Silvia Giordano

Istituto Sistemi Informativi e Networking, SUPSI, Lugano, Switzerland

Co-Supervisor: Dr Luca Luceri

Istituto Sistemi Informativi e Networking, SUPSI, Lugano, Switzerland

April, 2022

# Certificate

This is to certify that the MS Thesis entitled **Tracking Evolution of Coordinated Activity on Twitter**, submitted towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research Tirupati, represents the work carried out by **Padinjaredath Suresh Vishnuprasad** at Istituto Sistemi Informativi e Networking, SUPSI, Lugano, Switzerland under the supervision of Prof. Silvia Giordano, Istituto Sistemi Informativi e Networking, during the academic year 2021-2022 and the same has not been submitted for any other degree/diploma of any other university or institute.

Name of student: Padinjaredath Suresh Vishnuprasad

Roll Number: 201701003

Signature of student:

Date : April 12, 2022

Name of Supervisor:**Prof.Silvia Giordano**

Signature of Supervisor:

Date: April 12, 2022

Name of Co-Supervisor:**Dr.Luca Luceri**

Signature of Co-Supervisor:

Date: April 12, 2022

# Declaration

I declare that the matter presented in the MS thesis entitled **Tracking Evolution of Coordinated Activity on Twitter**, are the results of the work carried out by me at Istituto Sistemi Informativi e Networking, under the supervision of Prof. Silvia Giordano.

The contents are expressed in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic ethics and integrity and have not fabricated or falsified any idea /data /fact /image /source in my submission.

I understand that violation of the above will lead to disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained.

Name of student: Padinjaredath Suresh Vishnuprasad

Roll Number: 201701003

Signature of student:

Date : April 12, 2022

**Endorsed by:**

Name of Supervisor:**Prof.Silvia Giordano** Name of Co-Supervisor:**Dr.Luca Luceri**

Signature of Supervisor:

Date: April 12, 2022

Signature of Co-Supervisor:

Date: April 12, 2022

# Acknowledgements

I want to thank every one who supported me in this journey but especially these folks:

Silvia and Luca for their guidance.
The Mercatus Center for enabling this work.
My colleagues for all the meetings, *pause con caffè*, and collaboration.
My family and friends for all their care and love.
Nisha Rao for her unwavering support.

# Abstract

Social media impacts the way we communicate have have real-life implications beyond online discourse. This thesis deals with the online events on Twitter that preceded the violent attack on Capitol Hill in the US on January 6th and the events in the 2020 elections that preceded.

We use computational social science methods to look into two forms of social media manipulation — rapid retweet and copypasta. We see strong evidence for astroturfing in both methodologies and suggest concerted efforts to amplify the hashtag #stopthesteal. Further, we performed spatiotemporal analysis in a phase space and came up with an evolutionary model for hashtag evolution. This is in line with journlistic and OSINT results and this study complements them.

Finally, this also encompasses the pipelines used in the project S4CH: Search 4 Computational Hate and is funded by the Mercatus Center.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Social Media and Online Discourse

With the meteoric rise in the use of social media since the early 2010s, the landscape of civil discourse and communication has been altered irrevocably.[16] Social media impacts not just the way we communicate but also may have real-life implications beyond social media discourse. This is particularly clear when there is heated public debate, both influencing and influenced by offline events and crises. While one may be tempted to disregard the influence of social media on far-reaching socio-political events, there is mounting evidence that shows that not only does social media influence current events, but also there are unprecedented political consequences to online activity.[23]

The Arab Spring, a series of pro-democracy movements that spread across the Arab world in the early 2010s, has proved to be an early example of social media conducting socio-political events. The democratised nature of the Internet allows users to access information that may otherwise be unavailable due to censorship. Platforms like Facebook and Twitter give activists access to mass organisational capabilities within weeks or months. This inspired trans-boundary revolutions against authoritarian regimes throughout the Arab world, bringing international attention to the power of social media.[21]

Since then, there has been an increased focus and research on social media and its capacity to drive social events. The increased data consumption rate led to the automation of content aggregation by the social media platforms by recommender system-based algorithms. This is an environment where automated scripts, called bots, and human agents can thrive simultaneously, widely referred to as Web 3.0 (as opposed to Web 2.0 featuring just social media networks). High levels of automation and human-code interaction have fostered a space where social media can be manipulated to influence the narrative online. Bad-faith actors have known to promote disinformation and propaganda using various tactics, even influencing electoral processes.[2][13] Computational propaganda has now become a major concern and threat to democracy worldwide.

In contrast to the Arab Spring, the 49th quadrennial US Elections in 2020 showcased an example of social media playing a vital role in undermining the results of a democratic election, resulting in a violent insurrection at the country's seat of government.[11][12][14] While social media has played crucial roles in the American electoral process since 2008, its significance as an agent of political influence and space for discourse has increased since 2016. The revelations from the Cambridge Analytica whistle-blower and the role social media may have played in swaying voter sentiments in crucial swing states highlighted some of the issues within the system.[3] This arguably reached a tipping point when various users ranging from conservative politicians to far-right fringe groups, and even the incumbent president, promoted the unsupported claim that the 2020 US presidential election was "stolen" from them and people should revolt against the electoral system.[14] This resulted in the followers of these conspiracy theories storming Capitol Hill on 6th January 2021 in a violent attempt at "stopping the steal", resulting in widespread damage to property, injuries, and up to nine deaths.

## 1.2 The 2020 US Elections and The Insurrection at Capitol Hill

The 2020 US elections took place against the backdrop of the global COVID-19 pandemic following a controversial presidency by the Trump administration. Significant issues debated were health care, the COVID-19 era economy, racial unrest following the Black Lives Matter movement, and climate change. After primaries by August 2020, the nominees were the 47th Vice President Joe Biden and California Senator Kamala Harris for the Democratic Party and President Donald Trump and Vice President seeking reelection for the Republican Party. The polling took place on 3rd November 2020.

Due to the ongoing COVID-19 pandemic, social distancing and postal voting made this election unique. During crucial months leading up to the polls, the incumbent president Donald Trump challenged the integrity of the elections by falsely alleging widespread voter fraud and encouraging conspiracy theories about the polls. He alleged that mail-in ballots and absentee ballots were rife with fraudulent votes and that they would shift the balance against him. This was later proven to be a fraudulent claim and was seen as an attempt "to upend the constitutional order".

The ballot counting began soon after the polling ended on 3rd November. Trump started claiming victory in multiple states prematurely. On 4th November at 2:30 AM

EST, he falsely announced that he had won the election and that all counting happening after that was fraudulent. He tweeted "STOP THE COUNT!" and mounted legal attacks on the election results with the help of Republican-led states. Even after all the major news networks declared Biden to have won the election, Trump refused to concede and even delayed the transition process by weeks. The Supreme Court eventually dismissed all litigation, and Biden was acknowledged as the winner on November 23rd.


Figure 1.1: Trump tweets "STOP THE COUNT!".
Source: Twitter

Trump and his supporters continued to engage in conspiracy theories claiming victory for himself. They rallied up several groups of people ranging from far-right white supremacist organisations to conservative voters to fight for an election they believed to have been stolen from them. They used the slogan "STOP THE STEAL" and the same hashtag online to garner support. This is now seen as an exemplary instance of a "Big Lie", a propaganda technique used in Nazi Germany to propagate "a lie so colossal that no one would believe that someone could have the impudence to distort the truth so infamously"[7].

The first example of an offline manifestation of the online campaign to overturn the elections was the protest outside the Maricopa County Elections Department on 5th November by Trump supporters to overturn the results and stop the ballot counting. Trump and various conservative politicians, television personalities, and activists encouraged multiple conspiracy theories, including the QAnon conspiracy theory.[23] Activism from fringe groups gained traction online in the months following the elections, calling for violence against those who opposed Trump's claims. On 6th January 2021, the two houses of the legislature, The Senate and The House of Representatives of the United States, convened to ratify the electoral college votes at the Capitol, Washington DC. Trump held a rally in parallel near Capitol Hill and called upon his followers to gather en masse.[25] As the legislature counted the electoral votes, Trump called upon his audience to "March for America", which was seen as a call to action. At 12.53 PM EST on January 6, 2021, a mob of Trump supporters attacked the United States Capitol and breached the Capitol Building. The joint session of Congress was recessed, and the mob successfully delayed the ratification of the election — one of their primary goals. The violence on 6th January resulted in at least five immediate deaths and multiple injuries. There have also been numerous concerns surrounding Capitol Hill's security, cybersecurity, and legal aspects of elections. This insurrection was seen as an attempted coup d'état.[27]

The attack was followed by various socio-political repercussions. Donald Trump was impeached as the president on 13th January 2021 but failed to be convicted with a two-thirds majority in the Senate. The FBI undertook a large-scale criminal investigation, and more than 615 people have been charged with federal crimes. Apart from the five people who died in the riots, four law enforcement officers took their own lives in the days and months following the insurrection. Since the violence was organised and promoted online, large-scale deplatforming ensued with Twitter, Instagram, and Snapchat permanently banning Donald Trump. Many of his followers flocked to alt-tech social media platforms such as Gab and Parler after being suspended from Twitter.


Figure 1.2: Insurrection at the Capitol Hill, January 6th 2021.
Uploaded to WikiMedia Commons by Tyler Merbler. Attribution 2.0 Generic (CC BY 2.0) [6]

The January 6th Insurrection was an attempt at undermining the results of a free and fairly conducted democratic election and a failed autocoup. With diverse value systems and backgrounds, Donald Trump assembled his followers and a violent assault on the seat of the government. An amalgamation of ideas and group identities played an essential role in this, and people came together to achieve a singular goal. A large part of this occurred online on social media platforms populated by organic users and automated bots, often with vested interests, making this a computational social science issue as well.

## 1.3 Objective and Scope of this Work

This study deals with the online events on Twitter leading up to January 6th, 2021. We aim to deconstruct the tweet space to investigate the evolution of ideas across time. We construct a broad picture of the relationships between the various movements that coalesced to incite the insurrection. This is best viewed as a work of data journalism with concrete roots in mathematical methods and data science. It functions as a complementary study that goes hand-in-hand with investigative journalistic studies. This is, however, not at the expense of scientific rigour or integrity; the tools we crafted are novel techniques that further state of the art in computational social science.

We do not look into the complex socio-political climate that led to the events we study — that remains outside the scope of our work. Importantly, our methodologies remain agnostic to the actual events we study. While we acknowledge the peer-reviewed academic work that characterises it by citing relevant sources, this study endorses a neutral point of view. It is only concerned with the methodology, results, and scientific interpretation. This allows the adaptation of our methods into a pipeline that we can use to study online social networks in other domains and contexts (see 2.7: Pipeline Building).

We have employed various pipelines to analyse and dissect the tweet space from the elections to the riots. These methodologies can be broadly viewed as two orthogonal yet complementary axes of study: the spatial and temporal axes. The spatial axis deals with the topology and underlying structure of the information landscape and can show us the nature of discourse on Twitter. The temporal axis studies the evolution of the landscape over time, allowing us to understand the narratives constructed surrounding the elections. Furthermore, the pipelines are intended to look into two kinds of social media manipulation techniques usually employed in coordinated campaigns — rapid retweets and copypasta — and one to broadly map the tweet landscape over time. Together, these pipelines helped us to characterise the tweet space for the 2020 US Elections.

# Chapter 2

# Methodology

## 2.1 Data Collection

We used the first public Twitter dataset on the 2020 US Presidential elections, a large tweet dataset collected by Emily Chen, Ashok Deb, and Emilio Ferrara.[4][5] This was compiled using three data points: mentions and accounts related to candidates running to be nominated as their party's nominee for president of the United States, some general election-related keywords and some hashtags. They used the Twitter API and Tweepy API, a python-based end-point, to streamline the tweet collection procedure. It was collected from January 2020 to July 2021, and it contains above 1 billion tweets, with release v1.16 containing 1.35 billion tweets spanning January 2020 to February 2021.[4]

To collect data from Twitter, one needs developer credentials represented by a secret key and a bearer token. Researchers and academics associated with universities and research institutions can get elevated access to additional features if they can provide evidence for their work. This is under the conditions stipulated by Twitter's Developer Agreement & Policy. This clearly states that publically released datasets cannot reveal any data specific to individual tweets except for a tweet's Tweet ID. This makes it necessary to "hydrate" the Tweet IDs to reconstruct the dataset.[24]

We used custom-made pipelines to collect the data based on the publically available Tweet IDs. The code uses asynchronous functional programming to streamline the process of Tweet collection. It makes the code highly modular and reusable. The code is fully automated, and it outputs comma-separated value files (CSV files) with data regarding the tweet, metadata like hashtags, URLs, and mentions, as well as user data as columns. This allows us to have a broad overview of the context of each tweet. Since we are dealing with sensitive data pertaining to users, we follow the ethical guidelines set forth by the FAIR Principles for Findability, Accessibility, Interoperability, and Reuse of digital assets.[9] We also abide by Twitter's Developer Agreement & Policy and will only release the Tweet IDs of the tweets used in the dataset.

We use a subset of the US Election dataset — tweets from July 2020 to January 2021. This dataset takes up 515GB of disk space and contains a total of about 500 million tweets. Of this, retweets — tweets that simply re-share another tweet by another author or self — were 52% of the tweets, original tweets — tweets with original content produced by the author — were 16% of the tweets, replies — tweets that were written as a response to another tweet — were 19% of the tweets, and quotes — tweets that re-share a tweet with additional original content — were 13% of the tweets.

Due to its large size, we work on subsets of this dataset. We sample tweets from this dataset to contextualise and compare our results with the general corpus wherever necessary. Additionally, tracking patterns in the sampled dataset allows us to account for noise in some results, lending additional clarity to our interpretations. More information on this is given in the relevant sections.

## 2.2 Exploratory Data Analysis

### 2.2.1 Fringe Hashtags

We want to study specifically the events leading up to the January 6th Attack. We looked into multiple preliminary analyses and reports on the attack [23][26] and election-related discourse on the Internet. We isolated nineteen hashtags associated with three broad schools of conspiracy theories. They belonged to qanon, election fraud, and COVID-19 related conspiracy theories.

QAnon conspiracy theories are related to the fringe political movement that believes an anonymous user claiming to have a high-level security clearance in the US Government, Q, on various online spaces and message boards. They claim that a cult of Satanic, cannibalistic sexual abusers of children operating a global child sex trafficking ring conspired against Donald Trump. The conspiracy theory was born in the imageboard 4chan and then migrated to 8chan. While this was in no way the first "anon" that claimed to be a high-level government employee revealing secrets, QAnon was the first that managed to gain popularity beyond internet culture and was soon seen in election rallies by Donald Trump. This was when mainstream media started reporting on them. QAnon often uses the slogan, "Where we go one, we go all", giving rise to the hashtag #wwg1wga. Other hashtags they use are #pizzagate (one of the first claims by Q was that a pizzeria in Washington DC was where several high-profile Democrats were sexually abusing children, which led to an

| Election Related | QAnon-related | COVID 19 |
|---|---|---|
| obamagate | pizzagate | plandemic |
| sharpiegate | qanon | |
| qsnatch | qarmy | |
| stopthesteal | taketheoath | |
| dobbs | wwg1wga | |
| dominionsoftware | projectveritas | |
| dominion | thegreatawakening | |
| hammer | civilwar | |
| scorecard | | |
| hammerandscorecard | | |

Table 2.1: Table of Fringe Hashtags

**Election-related:** Conspiracy theories that are attributed to the elections and primarily allege election fraud. It was mainly popularised by conservative politicians, talk-show hosts, and influencers.

**QAnon-related:** Conspiracy theories that come under the umbrella of the QAnon conspiracy theory propagated on message boards like 4chan and 8chan by the anonymous user called Q.

**COVID19:** Conspiracy theories related to the COVID-19 pandemic. #plandemic alleges that the COVID-19 pandemic is a planned biological attack by 'the global elite'.

armed attack on the establishment), #taketheoath (a "digital soldier oath" that QAnon followers were encouraged to take by high-profile followers), #projectveritas (a far-right activist group that has engaged in disinformation campaigns), #thegreatawakening (refers to the biblical trope of a Great Awakening drawing parallels between the supposed fight between Trump and "the cabal of satanic cannibalistic paedophiles" and the canonical battle between good and evil; this particular belief has led to may experts characterising QAnon as a neo-religious cult), and #civilwar (calling for a second American civil war, often tied to white-supremacist movements).

Donald Trump made claims of voter fraud well before the November 3rd elections; it played a significant role in setting the stage for the election fraud claims he made soon after polling ended. This fuelled narratives of election fraud among Trump supporters starting the hashtag #stopthesteal. This was accompanied by two main kinds of conspiracy theories: voting machine-related and ballot counting related. The former gave rise to #hammer, #scorecard, and #hammerandscorecard, alleging a software called Hammer would break the protected networks and Scorecard would alter the total vote tally against Trump. #dominion was also widespread since its followers believed that Dominion Voting Systems, a producer of voting machines, "deleted" millions of votes cast for Trump. #qsnatch was a hoax that claimed a malware was "stealing the election" in multiple states. The latter comprised two hashtags reignited during the election period, namely #obamagate and #sharpiegate. #obamagate (alluding to the Watergate scandal) was a claim that Trump made alleging former President Barack Obama of wiretapping him. #sharpiegate was a hashtag referring to the claim that Donald Trump altered a map showing the trajectory of Hurricane Dorian with a Sharpie (permanent marker) in September 2019 to fit his earlier false claims that it would hit the southern state of Alabama. Around November 2020, the hashtag became popular again, claiming that legal ballots with stray markings or bleed-through from Sharpies were being thrown out. Stories surfaced on social media claiming that election workers purposely had voters mark their ballots with a Sharpie so their vote would not be counted. This was later disproven, and Maricopa County announced that Sharpies do not result in disenfranchisement. The Trump campaign dropped a related lawsuit on 13th November.

Notably, the one hashtag that stood out was #plandemic, a COVID-19 conspiracy theory alleging the global pandemic to be an orchestrated hoax by the pharmaceutical industry and entrepreneur and philanthropist Bill Gates, and it has been a launchpad for various anti-vax conspiracy theories. The data suggest that this intersects with several other election-related conspiracy theories since the pandemic and the Trump administration's handling of it was an important issue debated in the run-up to the elections. The list of hashtags used is given in Table 2.1.

### 2.2.2 Analysis

We extracted all tweets containing at least one of the nineteen "fringe" hashtags. We used parallelisation to speed up the process and utilised multiple cores of a MacBook Pro with a 2.3GHz 8-Core Intel Core i9 processor with 32 GB DDR4 memory. The code ran through about 515GB of raw data to output a file of approximately 458K unique tweets from 156K unique users in about 10 hours. This gave us a set of tweets, named the "fringe dataset", with which we performed most of the analysis in this study.

We grouped the fringe tweets by time and looked at the spread across months. The tweets were primarily retweets throughout the months, except in October 2020, when reply tweets took over the retweet count. That was also the month with the lowest number of tweets in the dataset — around 8000 tweets. November 2020 had the highest number of tweets, with more than 100,000 tweets. A log-linear plot of the number of tweets in the fringe dataset is given in Figure 2.1.

Next, we looked at the distribution of hashtags in the dataset. #stopthesteal was the most used hashtag, with more than 250,000 instances being used. About 50% of the tweets contain #stopthesteal — a highly unusual number. We also found that #qanon, #obamagate, #wwg1wga, #dominion, #plandemic are the top six hashtags in the dataset.

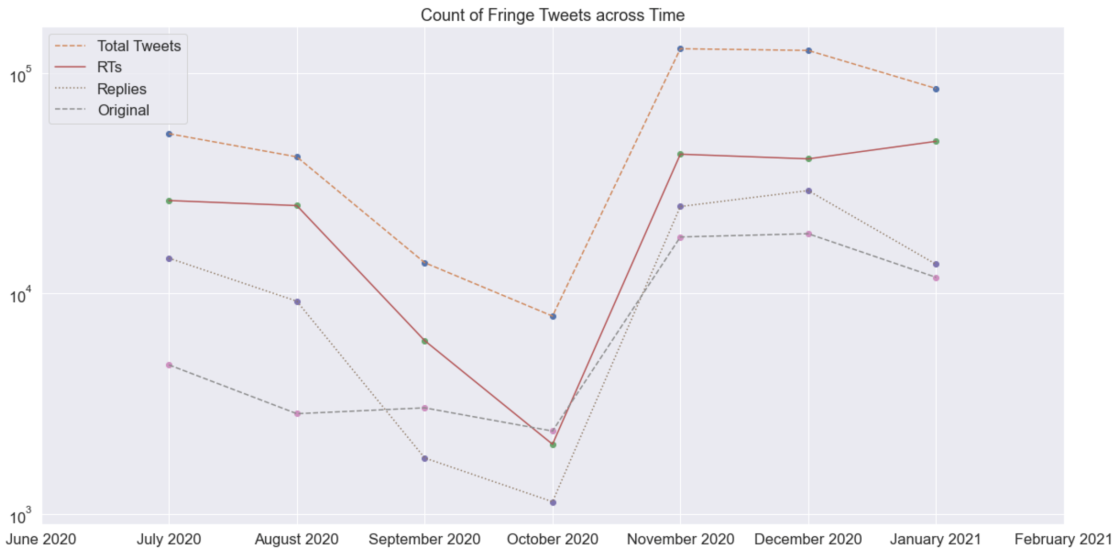


Figure 2.1: The count of tweets in the filtered dataset by month.

Interestingly, they represent hashtags from all three groups of hashtags mentioned above. This is visualised in the log-bar graph in Figure 2.2.

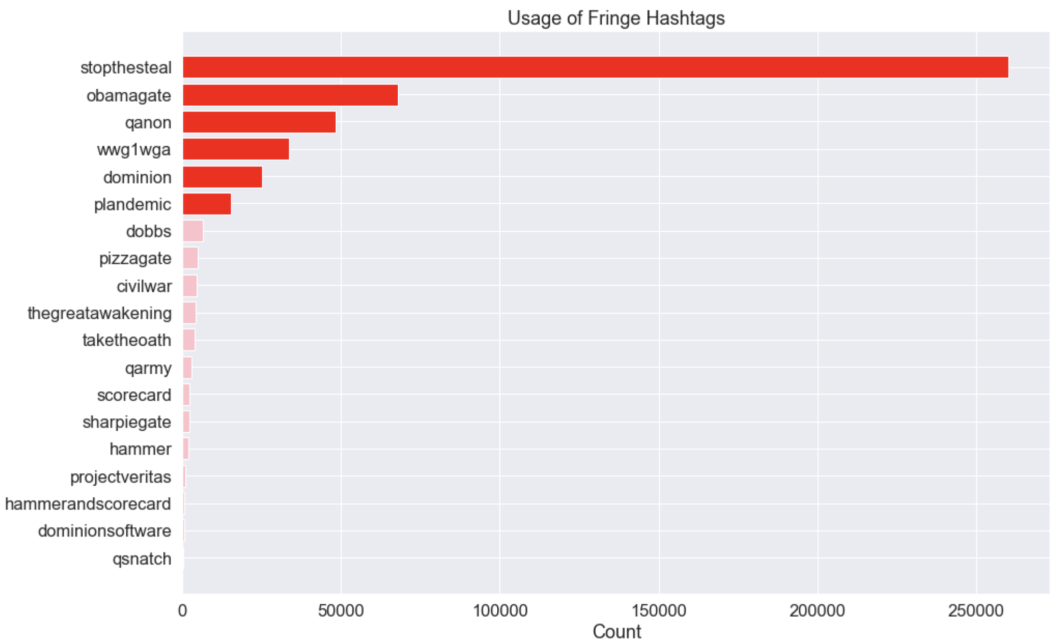


Figure 2.2: The count of hashtags used in the fringe dataset.
Top six hashtags are highlighted

Looking at the temporal spread of tweets by the top six hashtags, we see that they became popular at different times. For instance, #obamagate was used throughout the entire period, along with #qanon, #wwg1wga. This contrasts with the spread of #stopthesteal and #dominion, which became popular only in November after the elections. #plandemic became popular around September. This aligns with the qualitative understanding we have of the hashtags and their origin. Figure 2.3 represents this with a temporal kernel density estimate.

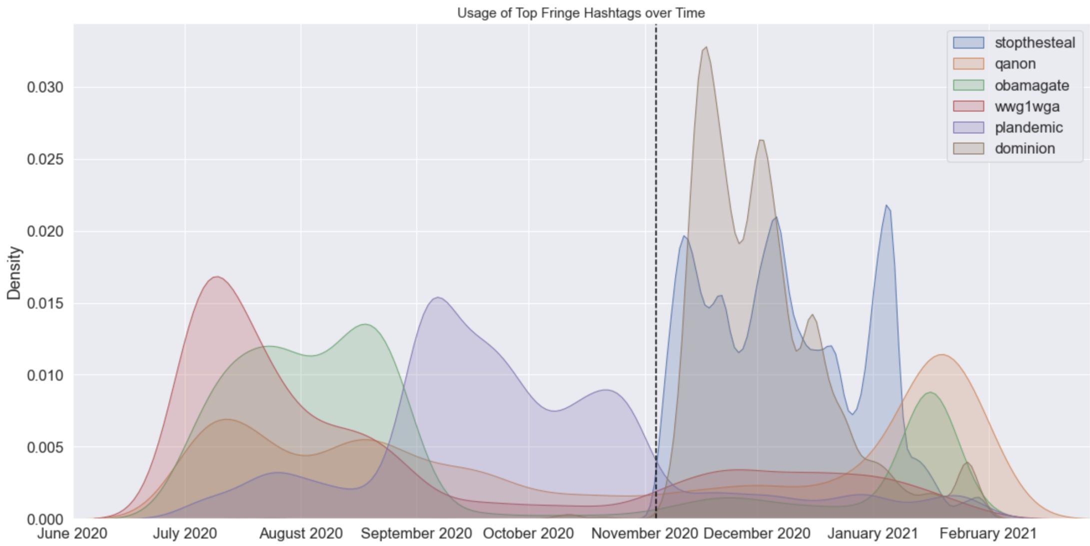


Figure 2.3: Spread of tweets with the top six hashtags used over time
The early hashtags used correspond to QAnon The hashtags used in between correspond to the #plandemic COVID-19 conspiracy theory The late hashtags used correspond to election disinformation post-November 3rd

This exploratory analysis gives us some context on the temporal spread of tweets, allowing us to design our methodology coherently. It also allows us to contextualise the fine-grained results from the detailed analysis that follows with elementary inferences. This is also why the “Methodology” section contains some preliminary results.

## 2.3 Spatiotemporal analysis

It is necessary to look at the spatial and temporal aspects of the data to understand the events leading up to the events on January 6th. We used three methods: two for tracking social media manipulation and one that tracks the spatiotemporal evolution of the tweet space.

First, we looked into the structure and topology of the tweet space by building two different networks: the rapid retweet and the tweet similarity network. This allows us to look into two social media manipulation tactics observed on Twitter, rapid retweet cascades and copypastas, which also represent well-known signals of coordinated inauthentic behaviours. For the evolution of the tweet space, we looked at the set of hashtags (“tagsets”) used in each tweet to characterise every tweet. We compared the tweets based on this characterization, we mapped the temporal evolution of the tweets. This gives us relational tweet networks that infer the higher level topology of a network of hashtags.

We also compared some results with a "baseline dataset" to contextualise the results from the fringe dataset. This was constructed by randomly sampling the raw data. This baseline dataset contains 70K tweets (42845 RTs, 15994 Replies, 15181 Quotes, 27155 Originals) and accounts for the noise in the signal we observe.

To feed into the pipelines, we split the dataset into two datasets — ORI (containing original tweets, replies, and quotes) and RT (having retweets and quotes). We consider quote tweets as an intersection of a retweet and an original tweet. This allows us to include quote tweets in both the subsets since it is essentially a retweet with a new idea expressed to go along with the quoted tweet. We use the RT dataset for the rapid retweet pipeline while the ORI for the other two. This is motivated by the following rationales:

- For rapid retweet analysis, we can consider quote tweets as a kind of retweet that adds commentary and can be used to manipulate the tweet landscape as much as regular retweets;
- For tweet similarity, we need to exclude all retweets since they have the same text as the retweeted tweet and will add noise to our data
- For landscape analysis, we are looking at the evolution of ideas through tweets. This is best represented by tweets that contain text, including quotes.

This gives us two intersecting datasets: ORI and RT. ORI contains 265740 tweets from 66305 unique users and RT contains 302378 tweets from 121490 unique users. They were then passed on to the pipelines.

### 2.3.1 Rapid Retweet Network

We define a rapid retweet as a retweet between two users that happened within ten minutes of each other.[17] We include a user in our analysis if they have done at least two instances of a rapid retweet of another user. This way, we have excluded all accidental rapid retweets, that is when a user sees a tweet and retweets it as soon as it is posted by chance, from our analysis. This means that we assume a rapid retweet happening more than once is not a chance event but may be a sign of coordinated activity to promote the tweets of a particular user. This is adapted from previous work in the literature.[17]

We do this by taking all the tweets in the RT dataset and calculating the time difference between the columns "created_at" and "rt_created_at" or "quote_created_at", as needed depending upon the type of tweet. This is done using the timedelta function in the python module pandas. This is then filtered based on its value being ¡ 600 seconds. This gives us a list of retweets and quotes within 10 minutes of its parent tweet. Next, we group the dataset using the columns "user_screen_name" and "rt_user_screen_name" or "quote_user_screen_name", and the number of instances is counted. This gives us a dataset with the parent and child users against the number of times they have interacted with retweets and quote tweets. All pairs with counts $< 2$ are removed, leaving us with the dataset of rapid retweets. The order of magnitude of the threshold values used is inspired by literature.[17]

This dataset is then exported as a comma-separated values file for visualisation with Gephi, a widely-used network visualisation software written in Java, with the parent user (retweeted users) as a "Source" and the child user (retweeter) as "Target". The count of interactions between them is saved as the "Edge weight". This is then post-processed by removing small, disconnected subgraphs. This is done using the python module networkx and its connected component function. This function outputs a list of disjoint subgraphs of the network, and we filter it based on its length being $> 5$. This cleans up our network and better represents the nodes of interest, and it is now ready for visualisation.

Furthermore, we characterise the accounts in the network by finding out if they are a bot or not. This is done using the tool Botometer from Indiana University.[29] We query the system with the Botometer API and identify the different bot scores it has for each user. We use the "botscore_overall" metric to characterise if the users are bots or not.

### 2.3.2 CopyPasta Network

In the Internet culture, a copypasta is a block of text copied and pasted across the Internet by individuals through online forums and social networking websites. Copypastas are often spam as they are used to disrupt online discourse.[28] However, Twitter's algorithm is built to boost engagement, and it will promote tweets, keywords and hashtags that are popular. This may cause tweets that are identical or nearly identical to be popularised. Malicious actors often take advantage of this to spread propaganda by repeatedly copy-pasting the same or highly similar pieces of content to simulate grassroots support and synthetically popularise an idea.[10] [18] We define such tweets as CopyPasta Tweets. To detect such activity, we make a pair-wise comparison of the text of tweets in the ORI dataset. We use the ORI dataset since retweets, by definition, have the exact text and cannot be called CopyPasta. Additionally, highly similar or identical text in tweets would indicate the same idea being conveyed by manipulation.

Calculating pairwise similarity between 265K tweets is a significantly time-consuming task that makes a blanket application of the algorithm infeasible. So, we compare tweets that happen within a neighbourhood of each other. We arrange all tweets in chronological order and consider a sliding window of ten tweets. We construct a $10\times10$ matrix of tweet similarity scores for each window. We use the Ratcliff/Obershep[19] algorithm to compare strings of tweet text with python's difflib library. For every tweet pair in the window, we record the similarity, the step-distance, and the TimeDelta between them to qualitatively infer the structure of the tweet space.[17]

This dataset is exported as a comma-separated values file for visualisation. The CopyPasta tweets are identified by plotting a histogram of similarity measures and are compared with the similarity distribution of a background dataset (as described before). The CopyPasta network is the network of tweets with a similarity score above a certain threshold. This is also post-processed by removing small, disconnected subgraphs using the connected components function. We filter it based on its length being $> 15$.

Furthermore, we plot the co-occurrence of hashtags within the CopyPasta network. We map the Tweet IDs of each tweet in the network to the hashtags they contain. This dataset is then decomposed into pairwise comparisons of the 19 hashtags to get an inferred hashtag network describing the CopyPasta tweet space with the hashtags as the frame of reference.

We also plotted the user interactions within the CopyPasta network. Similar to the algorithm above, we map the Tweet IDs of each tweet in the network to their authors and to the hashtags they used the most, giving us the usage of hashtags within the community. We also characterise their bot scores and plot the CopyPasta user network.

### 2.3.3 Tweet Landscape

In order to take a broad look at the tweets that form the fringe conspiracy theories, we construct a "landscape" of tweets. This is essentially a phase space where we map out the evolution of the conspiracy theories. We assume that the hashtags associated with the tweet represent the "idea" conveyed by the tweet. Additionally, the evolutionary landscape is a dynamic space; we aim to infer an evolutionary trajectory between the hashtags from the landscape. We employ an algorithm named "Tagset Analysis", where we map out the relationship between hashtags based on their co-occurrence in the conversational space.

For every hashtag, we define its "tag-dataset" to be the set of all tweets containing the hashtag in the dataset. Once we extract out a tag-dataset, we arrange it in reverse-chronological order. This implies that for every tweet in the tag-dataset, there is a "closest relative" tweet for a given distance metric. We calculate the upper-triangular pairwise-distance matrix, which compares every tweet with every other tweet that occurred before it. In this analysis, we obtained the similarity between the two tweets by comparing the set of hashtags they have, called "tagset". [20] Since the analysis is done solely within the confines of a tagset, the similarity metric will always be non-zero. We do not consider tweets with the exact same set of hashtags to make the concept of "closest predecessor" that we use later meaningful. The metric we used was a normalised cosine similarity given by:

$$S_{1,2} = \frac{\texttt{counter}(l_1) \cdot \texttt{counter}(l_2)}{\|\texttt{counter}(l_1)\| \times \|\texttt{counter}(l_2)\|} \cdot \frac{\texttt{min}(\texttt{len}(l_1), \texttt{len}(l_2))}{\texttt{max}(\texttt{len}(l_1), \texttt{len}(l_2))}$$

Where $l_1$ and $l_2$ are the hashtag sets of the tweets being compared, and `counter` refers to the python method `collections.Counter`. Counter takes a input of a set and gives an output of a dictionary of the counts of each unique element in the list. This can be represented as a vector of $\texttt{len}(l_1)$ + $\texttt{len}(l_1)$ dimensions. So, the first part of the equation now refers to a cosine similarity in $\texttt{len}(l_1)$ + $\texttt{len}(l_1)$-dimensional space. The second part of the equation is a normalisation term for length so that lists that are represented by subspaces in the $(l_1 + l_2)$-space don't converge to 1.

Once we have an upper-triangular matrix, each of its rows is scanned. We take the tweet pair with the maximum similarity as nodes and join them with a directed edge going from the older tweet to the newer one. Once we do this for every tweet in a tag-dataset, we obtain an evolutionary tree for that hashtag. We repeat the algorithm for all the 19 hashtags of interest. This gives us a set of temporally directed tree graphs that connects tweets containing each hashtag in phase space.

Since tweets can have multiple hashtags and will belong to multiple tagset graphs, we need to infer the relationship between the hashtags by looking at the relationship between each tagset tree. This is achieved by constructing a bipartite graph with the tweets in the first layer interacting with each other by the evolutionary trees and the hashtags in the second, inferred, layer interacting with each other. The edges in the second layer are obtained by looking at the co-occurrence of hashtags in tweets in the first layer; each tweet can be thought of as having a latent edge to the hashtags it contains and the dynamics of the first layer influence the second.

Practically, this is achieved by concatenating all the evolutionary trees together and then decomposing each interaction between the tagsets (with multiple hashtags) into interactions between individual hashtags. This essentially "explodes" each interaction into a series of interactions given by all permutations possible between the elements in their respective tagsets. This is done using the pandas function `explode` and `groupby` and it results in a pairwise comparison of all hashtags. This forms the edges in the second layer. A graphical representation is given in Figure 2.4.

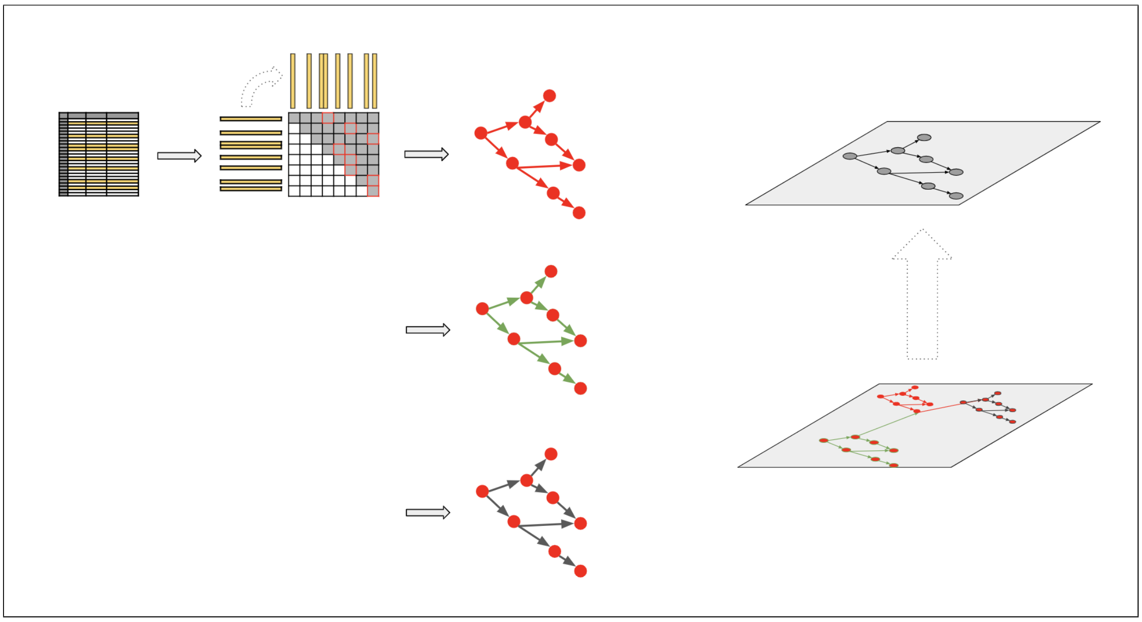

Figure 2.4: Tweet Landscape Algorthm [20]: The steps from, left to right, are tagset identification, similarity scoring, network building, network collation and hashtag network inference

## 2.4 Parallelisation and Functional Programming

Running the algorithms to run the code sequentially is time-consuming and does not allow all the available resources to be utilised. We have a MacBook Pro with a 2.3GHz 8-Core Intel Core i9 processor with 32 GB DDR4 memory and 8GB of dedicated GPU. It was soon evident that the bottleneck would be the read/write speed, even with SSD drives with a high I/O rate on board. So, we reworked all python notebook code to fit into a wrapper algorithm that parallelised the tasks with multiple processors while sequentially reading chunks of the data. This allowed for runtimes to go down from tens of hours to minutes. We used the python library `concurrent.futures` to run a `ProcessPoolExecutor`. It automatically manages the number of workers working with the available resources for you. This allows us to write short and modular code to run most processes in parallel.

To read large amounts of data, we employed a `try-except` argument that iteratively takes in data with the chunksize `**kwarg` in pandas `read_csv`. This manages each instance of data to be fed into the `ProcessPoolExecutor` automatically, while the latter handles multi-core processing. After each "chunk" is dealt with, it processes the data with a functionalised version of the algorithm that returns the results in a `dtype object` that can be compiled with the outputs from each chunk.

This methodology has multiple benefits over sequential processing, including lowering the time complexity of the algorithm. Functionalising the algorithm allows collaboration since a function can be seen as a mobile entity that can be applied in different contexts. It also allows the code to be published as a tool for researchers for them to employ and adapt for their own research. Open-source programming and reproducibility form the core tenants of the project, `S4CH: Search 4 Computational Hate`, of which this is a crucial component. See the next section for more details.

## 2.5 Pipeline Building

This work forms an integral part of the funded project `S4CH: Search 4 Computational Hate`. The project's primary output will deploy a web app that collates computational hate across Indian internet spaces. The end-user will be able to map out the spatio-temporal spread of the information they provide. The pipelines discussed here form the basis of this tool. It outputs visualisation-ready files from the initial data provided by the end-user. All the algorithms that form this tool are included in this project. As of writing this monograph, the data collection components are being appended to the pipelines and the tool will be released as an open-source project on GitHub.

# Chapter 3

# Results and Discussion

The results from the pipelines are presented separately here, and their inferences are discussed in length. However, one can draw overarching inferences from contextualising the individual results with each other that make this study especially powerful. Combining the empirical outputs with exceptional journalistic works on the events that transpired on January 6th 2020, builds a compelling argument. This makes our work a venture into data journalism while being a novel study in computational social science. This is also discussed in Chapter 4.

## 3.1 Rapid Retweet Network

For the rapid retweet network, the network was plotted using `gephi v0.9.2` with an inbuilt energy-based algorithm called ForceAtlas2. ForceAtlas2 implements a force-directed algorithm of network visualisation where the equilibrium state of a network is found by solving a system of Hamiltonian equations that assume the system to be a complex spring-mass system. This can be computationally intensive to be iterated sequentially. So, we use a Graphical Processing Unit (GPU), or a "graphics card" to run the matrix manipulations that the algorithm uses. It can take up to a few hours for the network to be resolved.

Here, the nodes represent the users in the retweet network, and the edges represent the presence of a rapid retweet, as defined in Chapter 2, happening between the two users. The node size is proportional to the in-degrees (or the number of times they have been retweeted), and the edge weights are the number of times a rapid retweet has occurred in the node pair. The node colour is bot score values going from green (zero) to red (one). This is represented in Figure 3.1.

We noticed that the network shows star topology, that is, some major nodes with lots of nodes attached to it, here, seen around some prominent actors with a lot of followers. This forms an inter-linked network with some important users forming subgraphs that shows evidence of being a part of an amplifying subsystem. Importantly the users that have a remarkable following are verified accounts from various conservative public figures — Tom Fitton, the president of Judicial Watch, a conservative think-tank (remarkably, also present in the network with a high bot score) and Lou Dobbs, a television personality who has a history of spreading conspiracy theories and vaccine misinformation are both examples of this pattern.

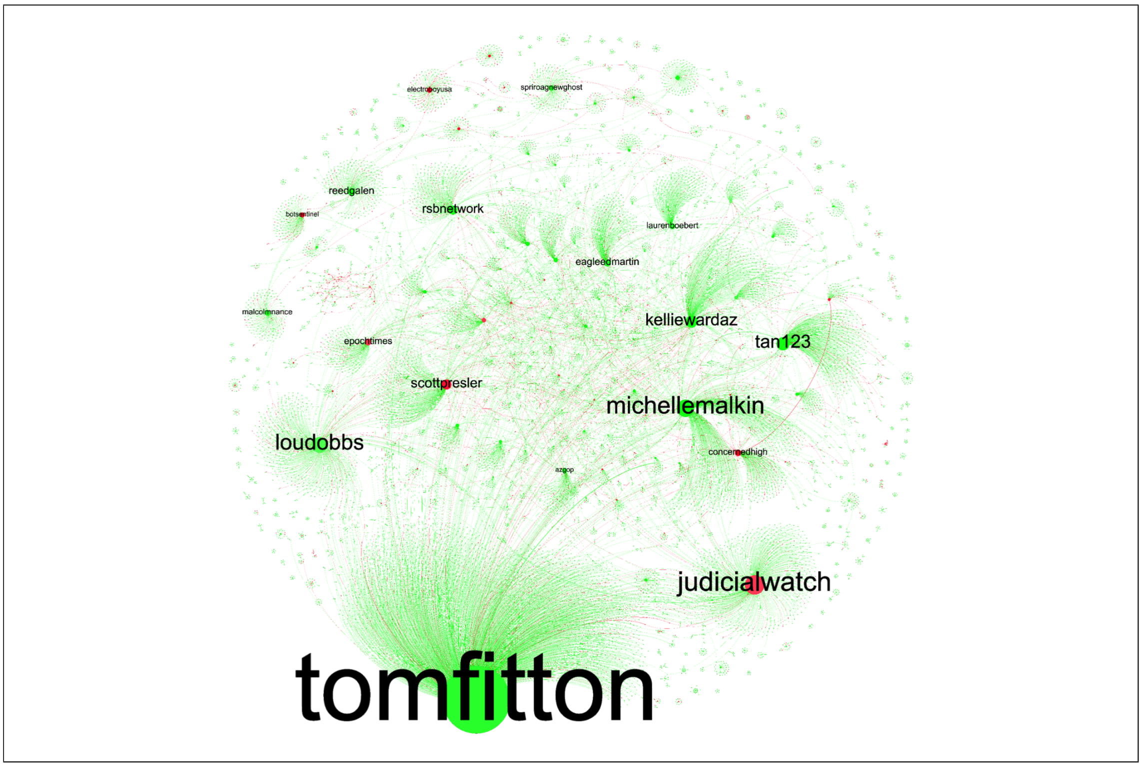


Figure 3.1: The network of users involved in rapid retweet
**Node Size:** In degree, or number of times retweeted
**Node Colour:** Bot Score

Another feature is that most of the important nodes have a low bot score, scored with Botometer's "botscore_overall". Users with higher bot scores seem to be sparsely distributed within the network. The most prominent nodes also correlate to those with a high eigenvector centrality [15], also known as the "prestige score", meaning that they play a central role in the structure of the network. This is represented by the node labels — only the nodes with a high prominence are labelled with label size proportional to the in degree, to match with the node size.

From this result, we can infer that there is some evidence for social media manipulation using rapid retweets. The retweeted users seem to be already high-profile users, and the network shows a pattern of preferential attachment towards them. This is backed by the eigencentrality measures of the prominent nodes. In the context of the US Election, Tom Fitton and Judicial Watch both being present is interesting since they are affiliated with each other, and the latter is an independent think-tank. Amplification of the tweets by these users is indicative of astroturfing, [18] the deceptive practice of presenting an orchestrated campaign to feign grassroots support. Additionally, the distributed topology of the network suggests some degree of interdependence and overlap in the interaction space — something that would be useful in a concerted disinformation campaign. Furthermore, the fact that the bot scores of most users qualify them to be organic users; that is, the human followers of these users are likely amplifying their tweets, and not automated scripts — indicative of human intervention.

## 3.2 CopyPastoid Network

Within the fringe dataset, we found empirical evidence that considerable amounts of copypasta activity is happening by plotting the similarity score distribution and comparing it with a sample non-fringe dataset. In a set of random tweets, it is expected that the number of pairs of tweets will go down with the similarity score — it is generally unlikely for the exact same text to appear in multiple tweets. This is what we observe for the sample dataset. However, when it comes to the fringe dataset, we see an initial downwards trend in the plot, but then it reaches a minima around 0.7 and then goes up to form a local maxima. This is evidence that an unusually high number of tweet pairs have a highly similar (similarity $>$ 0.7) tweet text. We call these tweets "copypasta tweets" and perform further analysis on them. [17]

We then plotted the cumulative distribution of the time difference between copypasta tweet pairs and compared it against the time difference between normal tweets with any similarity. This showed us that copypasta tweets tend to happen earlier than the baseline. In fact, 80% of copypasta tweets happen in the first 200 seconds while only $<$ 50% of the tweets have occurred in that time span. In the lighter shades in Figure 3.2, we also plotted the cumulative distributions of both the tweet pair subsets for every incremental distance within the 10-tweet window we use for this methodology (see Chapter 2).

Furthermore, we characterized the kind of tweets prevalent in the copypasta tweets and the ORI dataset. We observe that for the ORI tweets, the number of original tweets is much more in proportion than for the CopyPasta tweets. However, for replies and quote tweets, copypasta is more in number. This may be evidential for malicious manipulation of social media since quote tweets and replies count towards "tweet interactions" while orig-

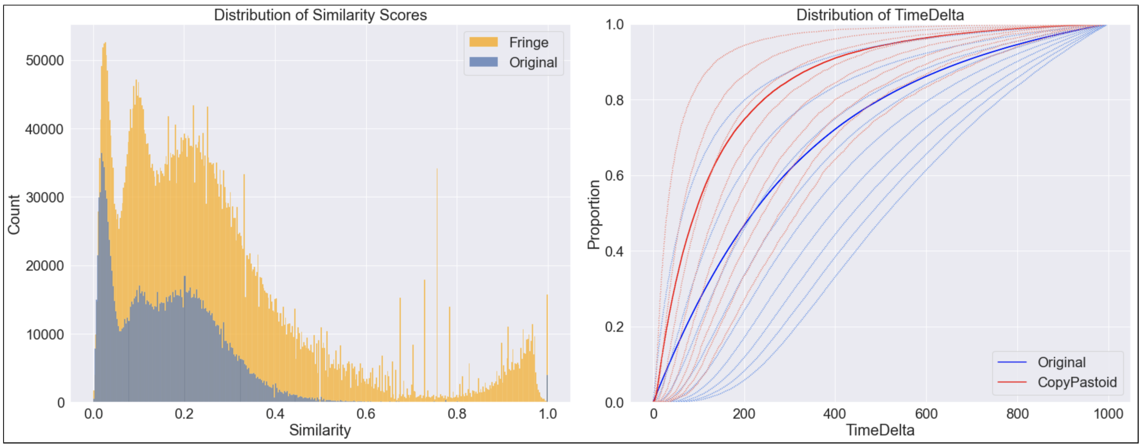


Figure 3.2: Tweet Similarity metrics

**Left:** The distribution of Similarity scores indicate that fringe tweets show copypasta activity while normal tweets sampled from the corpus do not. We can surmise that tweets with a similarity score > 0.7 are copypasta.
**Right:** The cumulative distribution of TimeDelta in the first 1000s indicate that copypasta tweets tend to be happening much faster than other fringe tweets. Lighter shades are incremental distances (see Chapter 2).

inal tweets do not; by interacting with copypastas, the copypasta is algorithmically boosted.

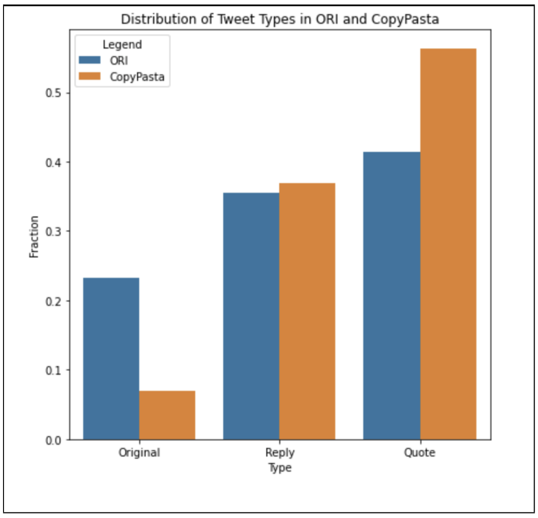


Figure 3.3: Tweet Types in CopyPasta network

These results allow us to construct a copypasta network with tweets with similarity < 0.7. This is given in Figure 3.4, with node size proportional to eigenvector centrality and plotted using the force-directed algorithm ForceAtlas2. We observe that the tweets form chains or meshes of tweets that are independent of each other. Each subgraph in the network is a set of tweets that have the same or very similar text associated with it — each a cluster of copypastas within the dataset. We then colour coded nodes by the hashtags involved in the tweet they represent and an overwhelming majority (76.25%) of them were

#stopthesteal tweets. This is strong evidence for #stopthesteal being highly manipulated with CopyPastas. The second-highest hashtag manipulated was #qanon with 4% of the share.

Further, we translate the copypasta tweet network to a user network. This is done by replacing the tweets with the corresponding authors and grouping them together by source and target of each edge. These users are then tagged with the most popular hashtag by

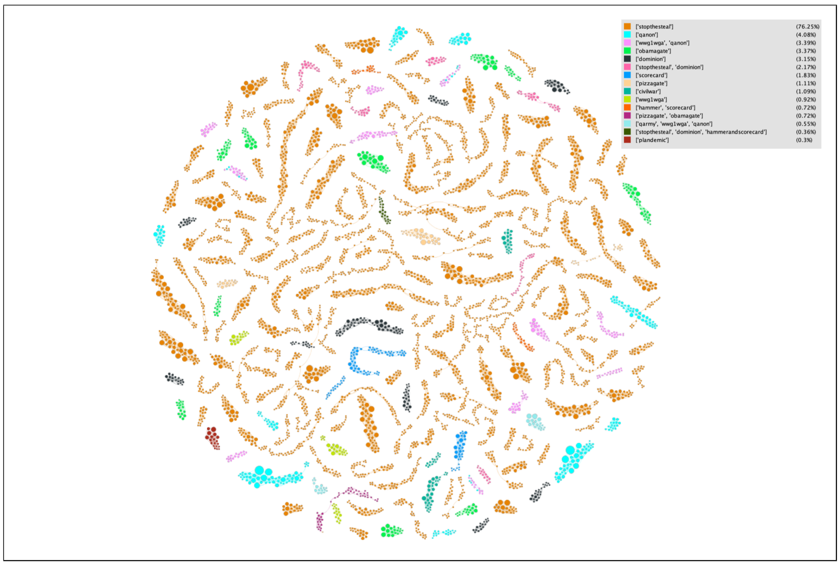


Figure 3.4: Hashtag-labelled copypasta tweet network.
#stopthesteal is highly copypasta-ed hashtag in the network by a large margin. This suggests that the hashtag was popularised by manipulation involving copypasta

that user in the copypasta dataset and visualised with the ForceAtlas2 in gephi. This network is shown in Figure 3.5. We observe a prominent community that has most of the people active in copypasta activity. Additionally, 96.4% of users are involved in copypasta activity involving #stopthesteal. This is an important result since it strongly supports the hyopthesis that #stopthesteal was intensely used in CopyPasta campaigns.

Finally, we translate the tweet network to a hashtag network by constructing the inferred bipartite graph from the co-occurrence of hashtags. This gives us the network in Figure n. The node size is proportional to eigencentrality. We performed an in-built community detection algorithm in Gephi that is an implementation of the Louvain algorithm,[1] a community detection algorithm that measures the density of links inside communities compared to links between communities to assign a modularity score for each node. We observe two communities in the space corresponding to QAnon-related hashtags (orange) and Election disinformation hashtags (violet). This suggests a separation between the co-occurrence of the election disinformation and the larger, more fringe conspiracy theories that were occurring in the space. This is in line with our current understanding that there were multiple

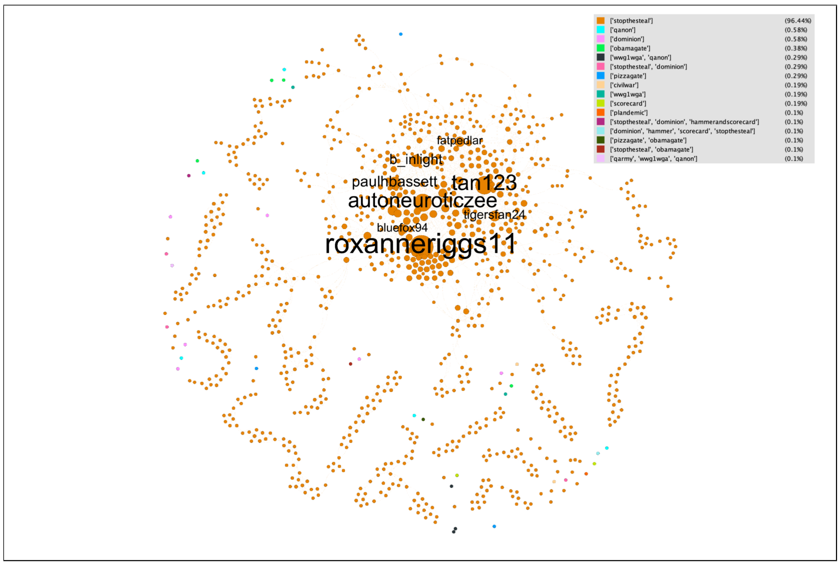


Figure 3.5: The network of users involved in CopyPasta
**Node Size:** Eigencentrality
96.4% of tweets contain #stopthesteal

groups of varying degrees of extremist beliefs at the Insurrection, with most people just supporting claims of election fraud and more radicalised groups promoting hate and violence to a larger degree.

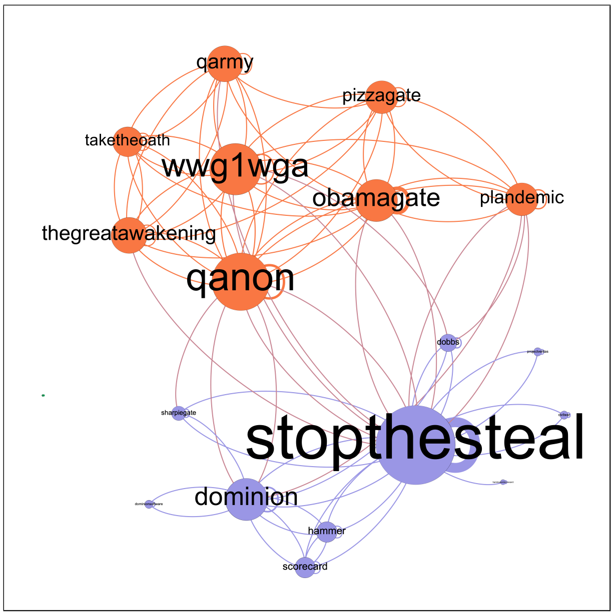


Figure 3.6: CopyPasta hashtag network
**Node Size:** Eigencentality value
**Node Colour:** Modularity: Orange is QAnon-related and Violet is election-related hashtags

These results are also evidence that indicates that #stopthesteal was manipulated using

CopyPastas. This could have played a crucial role in the popularity of the hashtag, allowing more people to be recruited to their cause by feigning grassroots support, or in other words, astroturfing.[18]

## 3.3 Tweet Landscape

For constructing a tweet landscape evolution that represents the hashtags, and the conspiracy theories they endorse, we use the tagset methodology on the ORI dataset elaborated in Chapter 2. The tweets are mapped to the respective hashtags, and the landscape is inferred by co-occurrence. We use a force-directed algorithm, ForceAtlas2, to plot the graph in Gephi. The node size is proportional to the eigencentrality, and the edge weight represents the number of inferred interactions between the two tweets. We also look at the temporal spread of the tweets containing each of these hashtags. We find the median timestamp of the creation date of each of the tweets to obtain a chronological sequence for hashtags. This is represented in the network as the node colour goes from green to yellow to red. The graph is shown in Figure 3.7

First, we observe a small degree of clustering as seen with the copypasta tweets as well. On running the Louvain Algorithm[1] on this, we got similar clustering suggesting that this pattern is indicative of topology beyond the manipulation techniques. The most important nodes are #stopthesteal, #qanon, #wwg1wga, and #dominion. Taking the temporal axis into account, we see that the QAnon-related hashtags were prevalent much earlier than the Election-related ones. In fact, almost all the election-related hashtags were popular soon after November, almost simultaneously, unlike the QAnon-related ones that seem to be staggered. This is in line with the backgrounds they have and the conspiracy theories they convey — QAnon theories predate the Elections and the Election-related conspiracy theories were often promoted by QAnon.

The close correlation between the hashtags are also explained by their backgrounds as decribed in Chapter 1. #wwg1wga is the slogan than many #qanon members used. The hashtags cooccurring with #stopthesteal are mostly specific election fraud conspiracy theories advocated by its followers, namely #dominion, #dominonsoftware, #scorecard, #hammer, and #hammerandscorecard.

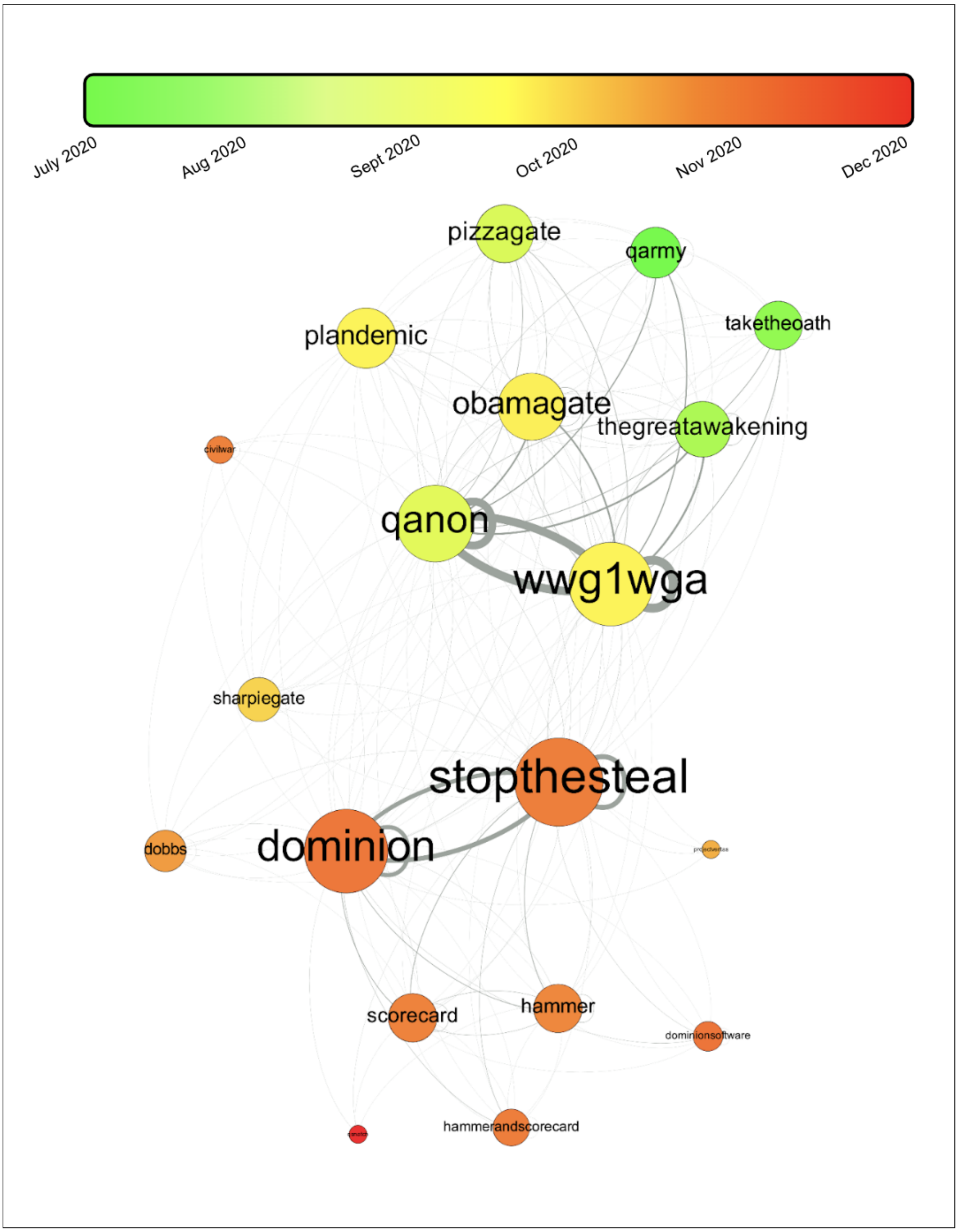


Figure 3.7: Tweet Landscape network
**Node Size:** Eigencentality
**Node Colour:** Median time of popularity

# Chapter 4

# Conclusion

## 4.1 Overview

The work done so far has resulted in novel points of view on disinformation during the 2020 US Election. From the results discussed here, we have three main takeaways. They are powerful on their own, but also has added value in the context of each other.

First, the rapid retweet network tells us that there has been significant amplification of tweets from prominent users. Even though the manipulation seems to be done by mostly organic users, or not bots, it still suggests astroturfing from the point of view of the tweet space. Generating apparent grassroot support is especially powerful on Twitter since the reccomender-systems will favour those points of views to people with views adjacent to those leading to radicalisation.

Second, the copypasta network strongly suggests that the hashtag #stopthesteal was amplified using this mode of manipulation. The fringe dataset showed strong evidence of copypasta activity while the genaral sampled dataset showed the expected similarity score distribution in the absence of manipulation. More than 76% of the copypasta tweets were about #stopthesteal and more than 96% of users involved in copypasta were talking about #stopthesteal. This could possible qualify the hashtag #stopthesteal to be the subject of a concerted disinformation campaign in an attempt to overturn a major election — an activity that have social implications and maybe even severe legal implications.

Finally, the tweet landscape weaves a narrative regarding the evolution of conspiracy theories online from the beginning of the election campaigns in the US. It tells us, backed by data, that tweets concerning QAnon conspiracy theories gained traction before the election and as soon as the polling was done in early November, the narrative shifted to that of election disinformation. The existence of two hashtag sub-communities that are separated in the phase space and temporally is indicative of this shift in the narrative.

The uniqueness of this work lies in the fact that these results, while being born out of data and mathematical analysis, complements journalistic and Open Source Intelligence (OSINT) reports. There are even analogies between the results we see here and the actual events on January 6th as they transpired, and how various fringe groups were involved in the violence and how they were seeking various outcomes. This is discussed in length here.

## 4.2 Data-Journalistic Interpretation

Since the events on January 6th transpired, all major journalistic agencies have covered it and hundreds of reports have been compiled.[8] A US House Select Committee was also launched to investigate the events. The New York Times published an in-depth investigative study with the minute-by-minute footage uploaded online, radio communication transcripts, and police footage.[22][25] This NYT Visual Investigation, spread over 6 months produced a 40-minute visual report on the events. This reconstruction of the actual insurrection can be seen as an analogy to our results backed by data.

The NYT investigation looks at several groups of people who gathered in Washington DC, 30 minutes away from the Capitol Hill for a Save America rally organised by Trump. People from all 50 US States were represented there, with various backgrounds. Some of the remarkable ones are:

- Most of the people present are Trump supporters who believed that the election was stolen from them. They were not affiliated with any militia, or violent hate groups.
- QAnon conspiracy theorists who have a far-right fringe view on the events and they believe that Trump is fighting against a cabal of satanic paedophiles who run the government.
- The Oath Keepers are a far-right libertarian paramilitary group who has military training, equipment, and were a highly organised militia who aimed to take government officials into custody.
- The Proud Boys are a far-right nationalist, white-supremacist group with a history of street violence. They came into the mainstream after Trump mentioned them by name in a presidential debate and asked them to *stand back and stand by.*

Some of these groups were calling for violence well before the insurrection and Trump's speech. They managed to incite the mob further and interpreted Trump's speech as a call-to-action. Of them, Proud Boys instigated the storming before most of the other groups reached the Capitol Building. There were two distinct groups - one from the West side of the Capitol Building and one from the East Side - the former representing mostly Proud Boys and QAnon supporters and the latter is mostly Trump supporters. Most of the protesters arrive at the West side just minutes before some of the Proud Boys breach the actual building. The breach on the East side was also prompted by some of the protesters moving from the West to the East. This clearly shows a pattern of leadership from the armed militias involved. They came prepared for a violent attack on the Capitol while most of the Trump supporters followed suit.

Our results suggest a similar lead-in to the election conspiracy theories. Led by the various far-right groups. The clustering in the phase space, in this context, implies co-occurrence in the data, and we see that a similar overlap was there in the goals and the actions of the rioters. The temporal scale also supports this and we see a parallel where members of the Proud Boys led the rioters into the building, similar to the far-right conspiracy theories leading up to #stopthesteal.

Another important point relevant to this study was the fact that law-enforcement underestimated the possibility of an attack. Even though there were multiple warnings of an attack being organised online, it was deemed not credible enough to warrant action. Computational social science tools need to be developed to assist OSINT investigations into online disinformation campaigns. The tool, `S4CH: Search 4 Computational Hate` that this work is a part of is a stride towards that goal.

## 4.3 Implications and Further Work

This thesis forms a part of the project `S4CH: Search 4 Computational Hate` funded by an Emergent Ventures India grant (EV#1802) from The Mercatus Center at George Mason University at Washington D.C. The project's primary output will to deploy an online tool that collates computational hate across Twitter. The end-user will map out the spatio-temporal spread of the information they provide.

The pipelines described in this thesis form the cornerstone of the tool. Users can input the Twitter API queries, and with their own Twitter Developer Credentials,[24] collect data and replicate the steps in this analysis. This would be published as open-source code on Github and researchers can use adapt the methodology to fit their context.

However, to deploy this resource, we needed to build innovative models that fare well in well-researched systems, such as the US Election twitter space, which has been researched since 2016. What works well in anglophone internet spaces can be modified to fit the needs like India.[2] This is because the social media manipulation tactics remain the same throughout the world.

This work will be done in the months following and will be published in the form of papers and open-source repositories.

We have used a publicly available dataset on the US 2020 Elections[4] that has followed all the protocols and principles regarding handling tweet data. In our research, we have also followed the FAIR Principles for Findability, Accessibility, Interoperability, and Reuse of digital assets.[9] Our data is public, so fully accessible for everybody with open, free and universally implementable protocol and access, as requested by FAIR Principles of Accessibility. Our research includes only publicly-available data and does not endanger the rights, dignity, safety, and privacy of the involved participants. Our study was conducted at an aggregate level and has not focused on specific individuals, only a specific event. Results disseminated consider only the collective behaviours and preserve personal information. For these reasons, and in compliance with social media policies (Twitter's Developer Agreement & Policy)[24] and the conventional research community's research approach, we need not acquire ethical clearance of data usage."

I would like to thank my supervisors, Silvia Giordano and Luca Luceri, my colleagues whose collaboration made this work possible, Gianluca Nogara and Felipe Cardoso, and Stefano Cresci and Savvas Zannetou for their feedback and expert reviews.

# Bibliography


[1] Vincent D Blondel et al. "Fast unfolding of communities in large networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2008 (Oct. 2008), P10008. DOI: 10.1088/1742-5468/2008/10/p10008.

[2] Ualan Campbell-Smith and Samantha Bradshaw. *Global Cyber Troops Country Profile: India*. 2019. URL: https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/05/India-Profile.pdf.

[3] Rosalie Chan. *The Cambridge Analytica whistleblower explains how the firm used Facebook data to sway elections*. Business Insider, Oct. 2019. URL: https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10?r=US&IR=T.

[4] Emily Chen. *2020 US Presidential Election Tweet IDs*. GitHub, Mar. 2022. URL: https://github.com/echen102/us-pres-elections-2020 (visited on 03/31/2022).

[5] Emily Chen, Ashok Deb, and Emilio Ferrara. "Election2020: the first public Twitter dataset on the 2020 US Presidential election". In: *Journal of Computational Social Science* (Apr. 2021). DOI: 10.1007/s42001-021-00117-9.

[6] Creative Commons Corporation. *Creative Commons Legal Code*. URL: https://creativecommons.org/licenses/by/2.0/legalcode (visited on 04/10/2022).

[7] *Definition of BIG LIE*. Merriam-webster.com, 2020. URL: https://www.merriam-webster.com/dictionary/big%20lie.

[8] Steve Doig. *It is difficult, if not impossible, to estimate the size of the crowd that stormed Capitol Hill*. The Conversation, Jan. 2021. URL: https://theconversation.com/it-is-difficult-if-not-impossible-to-estimate-the-size-of-the-crowd-that-stormed-capitol-hill-152889.

[9] *FAIR Principles - GO FAIR*. GO FAIR, 2017. URL: https://www.go-fair.org/fair-principles/.

[10] Camille Francois, Ben Nimmo, and Shawn C Eib. *The IRA CopyPasta Campaign*. graphika.com, Oct. 2019. URL: https://graphika.com/reports/copypasta/ (visited on 03/31/2022).

[11] Kevin Freking. *Trump tweets words 'he won'; says vote rigged, not conceding*. AP NEWS, Nov. 2020. URL: https://apnews.com/article/donald-trump-tweets-he-won-not-conceding-9ce22e9dc90577f7365d150c151a91c7.

[12] Anisa Holmes and NBC Washington Staff • •. *Trump Supporters Breach Capitol in Jan. 6 Riot*. NBC4 Washington, Jan. 2021. URL: https://www.nbcwashington.com/news/local/photos-trump-rallies-crowds-of-supporters-in-nations-capital/2531368/ (visited on 03/31/2022).

[13] David M. J. Lazer et al. "The science of fake news". In: *Science* 359 (Mar. 2018), pp. 1094–1096. DOI: 10.1126/science.aao2998. URL: https://www.science.org/doi/abs/10.1126/science.aao2998.

[14] Stuti Mishra. *Trump tweets 'I WON THE ELECTION' and is immediately corrected by Twitter*. The Independent, Nov. 2020. URL: https://www.independent.co.uk/news/world/americas/us-election-2020/trump-tweet-election-flagged-biden-b1723487.html.

[15] M. E. J. Newman. "Mathematics of networks". In: Networks. 2010, pp. 109–167. DOI: 10.1093/acprof:oso/9780199206650.003.0006. URL: http://www-personal.umich.edu/~mejn/papers/palgrave.pdf (visited on 04/10/2022).

[16] Esteban Ortiz-Ospina. *The rise of social media*. Our World in Data, Sept. 2019. URL: https://ourworldindata.org/rise-of-social-media.

[17] Diogo Pacheco, Alessandro Flammini, and Filippo Menczer. "Unveiling Coordinated Groups Behind White Helmets Disinformation". In: *Companion Proceedings of the Web Conference 2020* (Apr. 2020). DOI: 10.1145/3366424.3385775.

[18] Gordon Pennycook and David G. Rand. "Research note: Examining false beliefs about voter fraud in the wake of the 2020 Presidential Election". In: *Harvard Kennedy School Misinformation Review* 2 (Jan. 2021). DOI: 10.37016/mr-2020-51. URL: https://misinforeview.hks.harvard.edu/article/research-note-examining-false-beliefs-about-voter-fraud-in-the-wake-of-the-2020-presidential-election/.

[19] John W. Ratclif. *Pattern Matching: the Gestalt Approach*. Dr. Dobb's, July 1988. URL: https://www.drdobbs.com/database/pattern-matching-the-gestalt-approach/184407970?pgno=5 (visited on 03/31/2022).

[20] Hiroki Sato et al. "Exploration in Evolutionary Space by Hashtag Evolution on a Social Network Service". In: *The 2021 Conference on Artificial Life* (2021). DOI: 10.1162/isal_a_00464. (Visited on 03/31/2022).

[21] Ekaterina Stepanova. *The Role of Information Communication Technologies in the " Arab Spring " IMPLICATIONS BEYOND THE REGION*. 2011. URL: http://pircenter.org/kosdata/page_doc/p2594_2.pdf.

[22] New York Times. *Day of Rage: How Trump Supporters Took the U.S. Capitol — Visual Investigations*. www.youtube.com. URL: https://www.youtube.com/watch?v=jWJVMoe7OY0.

[23] Terri L Towner and Jody C Baumgartner. *The Internet and the 2020 campaign*. Lexington Books, 2021, pp. 3–23.

[24] Twitter. *Docs*. Twitter.com, 2019. URL: https://developer.twitter.com/en/docs.

[25] "U.S. Capitol Riot". In: *The New York Times* (). URL: https://www.nytimes.com/spotlight/us-capitol-riots-investigations.

[26] © Stanford University, Stanford, and California 94305. *The Long Fuse: Misinformation and the 2020 Election*. purl.stanford.edu. URL: https://purl.stanford.edu/tr171zs0069 (visited on 04/10/2022).

[27] District of Columbia 1100 Connecticut Ave NW Suite 1300B Washington and Dc 20036. *PolitiFact - Is this a coup? Here's some history and context to help you decide*. @politifact. URL: https://www.politifact.com/article/2021/jan/06/coup-heres-some-history-and-context-help-you-decid/.

[28] *What is Copypasta? - Definition from Techopedia*. Techopedia.com. URL: https://www.techopedia.com/definition/31470/copypasta.

[29] Kai-Cheng Yang et al. "Scalable and Generalizable Social Bot Detection through Data Selection". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 34 (Apr. 2020), pp. 1096–1103. DOI: 10.1609/aaai.v34i01.5460. (Visited on 09/22/2020).