#Misinformation: COVID-19 Discourse on Twitter

Paras Ahuja
Whiting School of Engineering
Johns Hopkins University
Baltimore, United States
pahuja2@jh.edu

***Abstract*—With the recent growth in technology and automation, there has also been growth in inorganic accounts on social media. After the 2016 U.S. elections, it came to light that inorganic accounts on Twitter may have played a disproportionate role in spreading misinformation during the election year [1]. Social media users can be misled with the information that is in front of them. 260,087 global English tweets from January 21, 2020 to November 23, 2020 were analyzed to find if inorganic accounts were present in COVID-19 discourse, and spreading misinformation. Research suggests that inorganic accounts are present in the discourse, often posting links to low-credibility sources. The evidence shows vast majority of the inorganic accounts were created in 2020 – the year of the COVID-19 pandemic. Furthermore, the sentiment of a post by an inorganic account is mostly neutral to positive, whereas an organic account is mostly negative to neutral. LDA topic modeling was utilized to understand how inorganic account posts differ from organic accounts. We notice that inorganic accounts have overlap in their topics, whereas organic accounts are more varied.**

***Index Terms*—Social Media, COVID-19, Bots, SARS-CoV-2, Twitter, Coronavirus**

I. Introduction

The COVID-19 pandemic is expected to be the most expensive health crisis in the recent world history. Harvard economists, David M. Cutler and Lawrence H. Summers, state that COVID-19 is the greatest threat to prosperity and well-being that the U.S. has faced since the great depression. At the time of writing, a total of 436,944 deaths in North America have been attributed to COVID-19, out of which 264,858 deaths are in the U.S. [3].

According to Cutler and Lawrence, COVID-19's cost to the U.S. economy is expected to be $16 trillion USD [2]. In October 2020, IMF reported that the cost of COVID-19 to the world economy is expected to be $28 trillion USD [2]. In September 2020, PEW Research Center published that due to the economic fallout caused by the COVID-19 outbreak one-in-four adults is having trouble meeting their monthly financial obligations, a third have dipped into their savings or retirement accounts to make ends meet, and one-in-six have borrowed money from friends or family or acquired food from a food bank [4].

There is no doubt that COVID-19 has created a difficult economic and health environment. A swift economic recovery may require increased social cohesion. However, social media has widened the ideological chasm. National Bureau of Economic Research says that social media has contributed to more polarization. People seek out like-minded ideas and individuals, and the room for new thought narrows, which ultimately leads to echo-chambers [6]. In an interview to The New Yorker, former U.S. President Barack Obama stated, "the capacity to disseminate misinformation, wild conspiracy theories, to paint the opposition in wildly negative light without any rebuttal – that has accelerated in ways that much more sharply polarize the electorate and make it very difficult to have a common conversation", when discussing social media [7].

Conspiracy theories are not new to Twitter. In October 2020, The Harvard Kennedy School's Misinformation Review paper found widespread misinformation on Twitter in regards to the COVID-19 pandemic. The paper disclosed attempts by some actors in society to leverage pre-existing cynicism about topics like masks, vaccination, trustworthiness of information, and government officials to sow doubts about the pandemic [8].

Research shows that inorganic accounts were disproportionately responsible for spreading misinformation during the 2016 U.S. Presidential election [1]. This raises a few questions, which this work aims to answer:

– RQ1: Are inorganic accounts engaged in COVID-19 discourse on Twitter?
– RQ2: How do the posts by inorganic accounts differ from organic accounts?
– RQ3: Is there evidence of inorganic accounts making unsubstantiated claims about the pandemic?

This work examines aforementioned questions using pre-existing inorganic account detecting, and sentiment analysis APIs to help aid analysis of tweets.

II. Background

A. Related Work

There is evidence of previous work in the area, and more scientific research is underway to better understand the role of social media in spreading information about COVID-19. Existing work highlights perils of social media misinformation, how to detect inorganic accounts, hashtags being used by inorganic accounts, and how to stop inorganic accounts from spreading misinformation. For instance, research conducted by Dr. Emilio Ferrara relies upon a combination of machine learning and manual validation to identify inorganic accounts. After identifying inorganic accounts, Dr. Ferrara uses statistical analysis to describe inorganic account behaviour. Research by Dr. Matteo Cinelli and colleagues demonstrates that information and misinformation have similar diffusion patterns.

Furthermore, research by Dr. Kathleen M. Carley analyzes spread of COVID-19 misinformation on Twitter. Dr. Carley's research characterizes two competing COVID-19 factions – informed and misinformed users.

B. COVID-19

Early cases of COVID-19 were reported in Wuhan, China in December of 2019. Shortly thereafter, fatalities were reported in January 2020. The U.S. reported its first case on January 21, 2020 [9]. On March 11, 2020, The WHO officially declared COVID-19 a pandemic [10]. At the time of writing there are 62,004,643 confirmed cases of COVID-19 around the world and 1,448,647 lives have been claimed by the virus [4]. Soon after the first case was reported in the U.S., social media misinformation started with everyone chiming in and offering their unsolicited advice on the matter. The situation was so grave that in February of 2020, WHO called it an 'infodemic' [11]. Gordon Pennycook and colleagues wrote: "in the case of COVID-19, this misinformation comes in many forms – from conspiracy theories about the virus being created as a biological weapon in China to claims that coconut oil kills the virus" [12].

C. Inorganic accounts Detection

An inorganic account is an account whose posts are generated by an algorithm, most commonly known as a bot. These days it is relatively easy to create a bot, as the instructions can easily be found by a simple Google search. However, detecting a bot is becoming an increasingly difficult task. Advancements in Machine Learning, NLP, and multilingual models are helping in the creation of human-like content [13]. This allows for the creation of advanced bots, which makes it harder to detect on a platform like Twitter [14]. For instance, Dr. Emilio Ferrara and colleagues found: "social bots can search the Web for information and media to fill their profiles, and post collected material at predetermined times, emulating the human temporal signature of content production and consumption – including circadian patterns of daily activity and temporal spikes of information generation" [15].

Work by Dr. Emilio Ferrara and colleagues attempts to understand inorganic account behavior. Dr. Ferrara highlights that the concern with inorganic accounts is not that they spread misinformation, but that they carry false information to popularity, giving a false sense of accuracy and endorsement [15].

There are many APIs available that help detect inorganic accounts online. One such example is Botometer, hosted by Indiana University's Observatory on Social Media. However, this work relies on **tweetbotornot2** by Dr. Michael W. Kearney of University of Missouri, which is available as an R package [21].

D. Sentiment Analysis

Sentiment analysis techniques have gotten better over the years. Growth in social media usage has also lead to growth in interested parties seeking ways to understand user sentiment. Companies such as Twitratr, Tweetfeel, and Social Mention are a few in the field of Twitter sentiment analysis [16]. However, Google, Amazon, IBM, and Microsoft also provide sentiment analysis APIs that perform the same task. This work relies on Valence Aware Dictionary and Sentiment Reasoner (VADER) Sentiment Analysis. VADER Sentiment is a lexicon and rule-based tool that is specifically attuned to sentiments expressed in social media [17].

III. Dataset

For this research, global tweets from January 21, 2020 to November 23, 2020 were analyzed for trends and patterns. The data was originally collected, and maintained by researchers at the University of Southern California (USC) [18]. The overall dataset contains approximately 840 million tweets. Following is the breakdown of the top 5 languages in the dataset:

TABLE I
Dataset Language Breakdown

Language	**ISO**	**No. Tweets**	**% total Tweets**
English	en	607,346,491	67.75%
Spanish	es	105,648,005	11.79%
Portuguese	pt	29,723,763	3.32%
Undefined	und	24,991,150	2.79%
French	fr	24,172,504	2.7%

The dataset consists of .txt files of tweet IDs. Researchers from USC made the dataset publically available on GitHub, and it was obtained from there.

A. Keywords

The dataset consists of tweet IDs that were collected using 76 keywords. These keywords were added as they came in usage. For instance, prior to President Trump calling the virus 'Kungflu', the term was not in usage, however tweets from June 2020 started to use the term [19]. Some of the keywords used are: Coronavirus, Koronavirus, Corona, CDC, Wuhancoronavirus, Wuhanlockdown, Ncov, Wuhan, N95, Kungflu, Epidemic, outbreak, Sinophobia, China, covid-19, corona virus, covid, covid19, sars-cov-2, COVID-19, COVD, pandemic, coronapocalypse, canceleverything, Coronials, SocialDistancingNow, Social Distancing, and SocialDistancing.

B. Data Preparation

Files for each month were merged into a single .txt file. Since each month has hundreds of files, the task was automated with a combination of R and Python. This resulted in 11 .txt files. Then, a random sample of 40,000 from each of the 11 files was selected. Hydrator application was used in order to download information associated with the tweet IDs in the sample dataset. Information obtained from Hydrator was in .csv format – 11 .csv files were obtained.

C. Data Cleaning

Each of the 11 .csv files contained information like time of tweet, hashtags, media shared, urls associated with the account, tweet ID, if the tweet is in reply to a username, user screen name, text of tweet, etc. Only the English language tweets were kept. Some of the tweets were no longer available because they were either deleted by the account holder, or the account itself was deleted or suspended. In total 260,087 tweets were analyzed over a period of 11 months from January 21, 2020 to November 23, 2020.

IV. Finding Inorganic Accounts

With regards to RQ1, **tweetbotornot2** was used to determine the probability of an account being inorganic. The package returns probabilities associated with each account. The closer a probability, $P(B)$, is to 1, the higher the chances of an account being inorganic. In this research accounts with $P(B) > 0.5$ were treated as inorganic.

A. Types of Inorganic Accounts

Inorganic accounts come in all shapes and sizes. Some inorganic accounts self-identify as such and post meaningless phrases, while others simply retweet information shared by other sources. Some inorganic accounts were verified accounts. Verified inorganic accounts most often belong to corporations, or corporate accounts that mimic bot behavior for various reasons. Only unverified inorganic accounts were examined for this research.

B. Inorganic Accounts in the Dataset

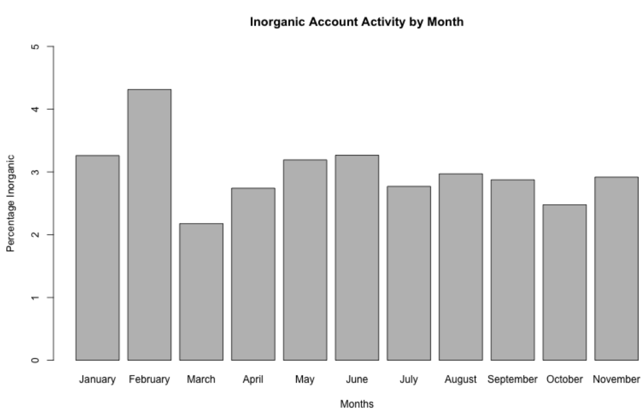

Fig. 1. Inorganic account activity by months of 2020

Inorganic accounts are present in the dataset, and on Twitter. Over the 11 month period in question, 6,430 bots were identified, or approximately 3% of the data used for this research.

Fig. 1 shows that every month approximately 2% to 5% of the accounts involved in COVID-19 discourse are inorganic. Furthermore, most of the inorganic account activity was in February, where approximately 5% of the accounts were classified as inorganic.

C. Inorganic Account Creation Year

Another trend that is rather interesting to see is when the accounts were created. Research shows that the vast majority of inorganic accounts related to COVID-19 keywords used for this research, were created in 2020.

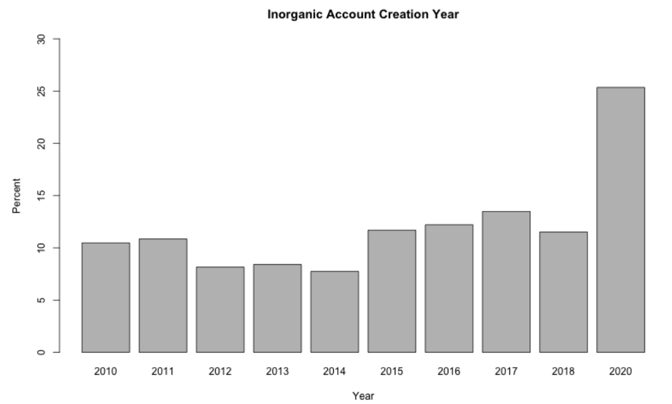

Fig. 2. Inorganic account creation year

In Fig. 2 we notice that initially, inorganic accounts were decreasing; less-and-less were created. However, there is a clear inflection point between 2014 to 2015. The data does not reflect any inorganic accounts from 2019. More than 25% of the accounts identified as inorganic were created in 2020.

Research suggests that societally or globally important events may lead to an increase in inorganic accounts on social media. On June 17, 2015 CNN reported that Donald Trump was running for president [22]. In the dataset, we notice that 68 inorganic accounts were created in June of 2015 – more than any other month of 2015. This suggests a correlation between the event and increase in inorganic accounts. We notice an even greater increase in inorganic accounts in 2020 – the year of the COVID-19 pandemic. There is a 120% increase in inorganic accounts as compared to the previous available year.

V. Inorganic Accounts vs. Organic Accounts

With regards to RQ2, inorganic and organic accounts differ in some key ways. First, we notice that inorganic accounts more often share a link in the posts as compared to organic accounts. Second, inorganic accounts more often share links to untrusted sources than organic accounts.

A. Links

Research shows that inorganic accounts are approximately three times as likely to share a link, with 70% inorganic accounts, as compared to 24% organic accounts. However, the credibility of these links is in question because mos of the links were shared through a link shortener. In manual tests link shorteners linked to obscure websites. Links like blogspot are highly unreliable since neither the source nor the information presented, in many cases, can be readily verified. Out of all the inorganic accounts that shared a link, close to 2.5% shared bit.ly link, followed by Facebook and blogspot, whereas the organic accounts most often shared links to Instagram, followed by YouTube, and Facebook.

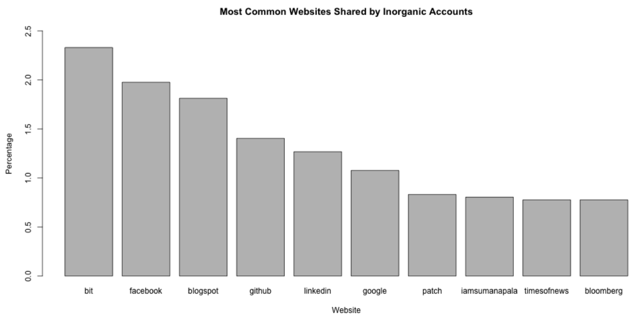

Fig. 3. 10 most common websites shared by inorganic accounts

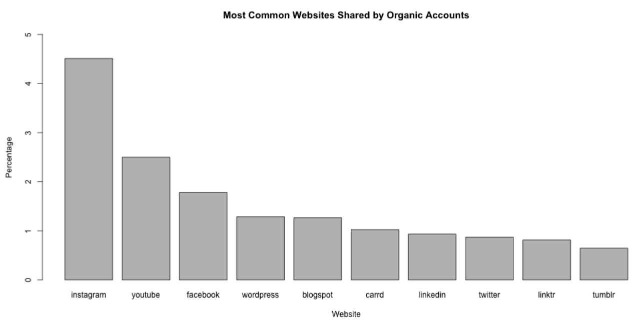

Fig. 4. 10 most common websites shared by organic accounts

B. Retweets

While inorganic accounts more often share links about COVID-19, organic accounts more often retweet about COVID-19. Research shows that approximately 23% of inorganic accounts retweeted, as compared to approximately 76% of organic accounts. Organic accounts most often retweeted posts from 72,458 accounts. Note that some of these accounts were retweeted more often than others. However, out of these 72,458 accounts 6,374 were classified as inorganic accounts.

C. Tweet Sentiment

VADER Sentiment was used to obtain the sentiment of a tweet. VADER Sentiment provides a compound score as follows:

- Positive sentiment: compound score ≥ 0.05
- Neutral sentiment: $-0.05 <$ compound score < 0.05
- Negative sentiment: compound score ≤ -0.05

Based on above criteria we notice that inorganic accounts mostly present neutral to positive sentiment, while organic accounts mostly present negative to neutral sentiment in COVID-19 tweets. However, this is not say that negative sentiment does not exists in posts by inorganic accounts – it is simply that the sentiment is mostly neutral or positive.

D. LDA Topic Modelling

We notice a difference between topics discussed by inorganic accounts as compared to organic accounts.

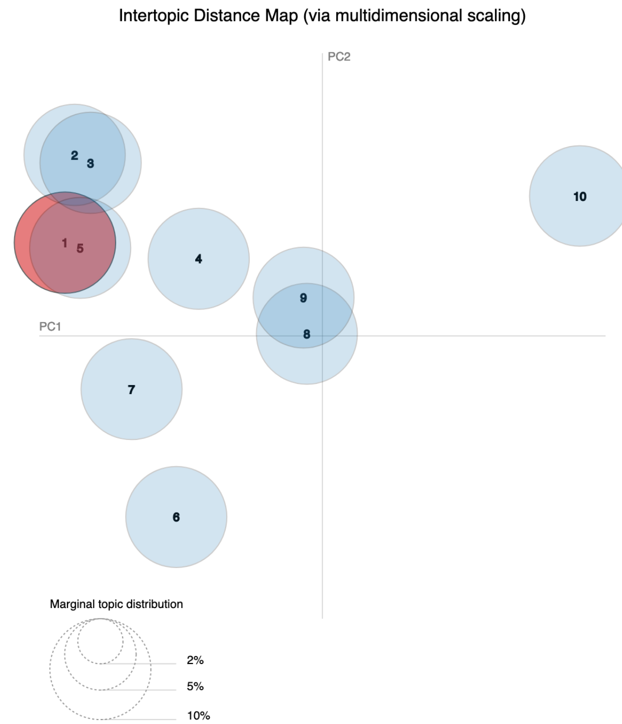

Fig. 5. Inorganic accounts intertopic distance map

Fig. 6. Organic accounts intertopic distance map

In Fig. 5 and Fig. 6, each circle represents a topic. Distance between the circles is indicative of the topic relatedness. In Fig. 5 the inorganic accounts have some overlap in topics and closeness. However, in Fig. 6 this is not the case with the organic accounts – there is minimal overlap and closeness. This shows the inorganic accounts are repeatedly discussing similar or same topics.

VI. Misinformation

With regards to RQ3, there is evidence of inorganic accounts making unsubstantiated claims. This is not specific to the pandemic alone, these accounts are also making varying degree of politically polarizing posts. However, for this research attention is on COVID-19 specific misinformation or unsubstantiated claims.

Consider **@o_newsroom**, which has a $P(B) > 0.99$. At the time of writing, this account made the following posts seen in Fig. 7 and Fig. 8:

Fig. 7. Dec 7 post by @o_newsroom classified as an inorganic account

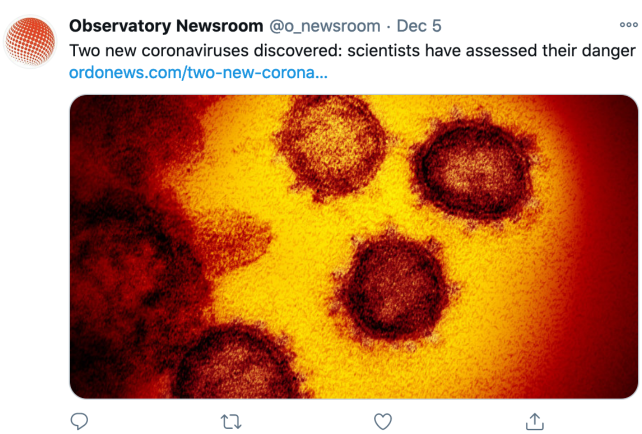

Fig. 8. Dec 5 post @o_newsroom classified as an inorganic account

The account in question refers to its own website as a source, and it has made other posts like: "NASA experts say aliens warned people about COVID-19 pandemic". Each time the account refers to its own website as a source, and does not provide a trusted third party who can verify the legitimacy of the claims.

Another account, which was also classified as an inorganic account, made varying degree of political posts. Some of the posts question the legitimacy of the COVID-19 vaccine, while others question the legitimacy of the 2020 U.S. Presidential Election. Account, **@CensoredToday**, which also has a $P(B) > 0.99$, made the following posts as seen in Fig. 9 and Fig. 10:

Fig. 9. Dec 8 post by @censoredtoday classified as an inorganic account

Fig. 10. Dec 5 post by @censoredtoday classified as an inorganic account

For this research a specific collection of hashtags were used. It is entirely possible to choose a different – more polarizing – collection of hashtags and find that inorganic accounts are involved to an even greater degree. For instance, Dr. Kathleen Carley and Shahan Memon found inorganic accounts on twitter using the following keywords: bleach, vaccine, acetic acid, steroids, essentialoil, saltwater, ethanol, children, kids, garlic,alcohol, chlorine, sesame oil, conspiracy, 5G,cure, colloidal silver, dryer, bioweapon, co-caine, hydroxychloroquine, chloroquine, gates, immune, poison, fake, treat, doctor, sennamakki, and senna tea. Dr. Carley and Shahan Memon found as many as 19% inorganic accounts in one set of users. A different set of keywords were used for this research, and yet there is still evidence of inorganic accounts making unsubstantiated claims on Twitter.

VII. Conclusion

6,430 unverified inorganic accounts were found in a dataset of 260,087 COVID-19 related tweets ranging from January 21, 2020 to November 23, 2020. Accounts were classified as inorganic if the probability was greater than 0.5 based on the probabilities obtained from **tweetbotornot2**.

Research shows that the vast majority of inorganic accounts related to COVID-19 were created in 2020. Inorganic accounts more often share links, however organic accounts more often retweet. In the dataset, organic accounts retweeted 72,458 accounts, and out of these 6,374 accounts were classified as inorganic accounts. The links shared by the inorganic accounts are usually from untrusted sources.

The sentiment of tweets for the inorganic accounts is mostly neutral to positive, however organic accounts tend to have more negative to neutral tweets. Topic modeling shows that the inorganic accounts have some overlap and closeness implying the discussion is focused on similar topics. Organic accounts have more variation in their topics of discussion.

This work demonstrates that there is evidence of misinformation being spread by the inorganic accounts, and the links that are posted by the inorganic accounts are mostly self-referential sources, which are presented as facts, but lack verification by other trusted sources.

More research is needed to understand the role of inorganic accounts in spreading misinformation in regards to COVID-19 and its societal ramifications. Moving forward, an important research question to consider should include the relationship between COVID-19 misinformation being spread by inorganic accounts and their link to positive cases.

References

[1] Shao, C., Ciampaglia, G. L., Varol, O., Yang, K., Flammini, A., &; Menczer, F. (2018). The spread of low-credibility content by social bots. Nature Communications, 9(1). doi:10.1038/s41467-018-06930-7.

[2] Transcript of October 2020 World Economic Outlook Press Briefing. (2020, October 13). Retrieved November 27, 2020, from https://www.imf.org/en/News/Articles/2020/10/13/tr101320-transcript-of-october-2020-world-economic-outlook-press-briefing

[3] Cutler, D. M., &; Summers, L. H. (2020). The COVID-19 Pandemic and the $16 Trillion Virus. Jama, 324(15), 1495. doi:10.1001/jama.2020.19759.

[4] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis; published online Feb 19. https://doi.org/10.1016/S1473-3099(20)30120-1.

[5] Parker, K., Minkin, R., &; Bennett, J. (2020). Economic Fallout From COVID-19 Continues To Hit Lower-Income Americans the Hardest. Pew Research Center.

[6] Boxell, L., Gentzkow, M., &; Shapiro, J. (2017). Is the Internet Causing Political Polarization? Evidence from Demographics. doi:10.3386/w23258

[7] Remnick, D., MacFarquhar, L., &; Collins, L. (2016, November 28). Obama Reckons with a Trump Presidency. Retrieved November 27, 2020, from https://www.newyorker.com/magazine/2016/11/28/obama-reckons-with-a-trump-presidency

[8] Kearney, M. D., Chiang, S. C., &; Massey, P. M. (2020). The Twitter origins and evolution of the COVID-19 "plandemic" conspiracy theory. Harvard Kennedy School Misinformation Review. doi:10.37016/mr-2020-42

[9] World Health Organization. (2020, April 23). Coronavirus disease 2019 (COVID-19) Situation Report –94. Retrieved November 28, 2020, from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf

[10] Archived: WHO Timeline - COVID-19. (n.d.). Retrieved November 28, 2020, from https://www.who.int/news/item/27-04-2020-who-timeline—covid-19

[11] World Health Organization. (2020, February 02). Novel Coronavirus (2019-nCoV) Situation Report –13. Retrieved November 28, 2020, from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf

[12] Pennycook, G., Mcphetres, J., Zhang, Y., Lu, J. G., &; Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention. doi:10.31234/osf.io/uhbk9

[13] Mctear, M. (2020). Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots. Synthesis Lectures on Human Language Technologies, 13(3), 1-251. doi:10.2200/s01060ed1v01y202010hlt048

[14] Alarifi, A., Alsaleh, M., &; Al-Salman, A. (2016). Twitter turing test: Identifying social machines. Information Sciences, 372, 332-346. doi:10.1016/j.ins.2016.08.036

[15] Ferrara, E., Varol, O., Davis, C., Menczer, F., &; Flammini, A. (2016). The rise of social bots. Communications of the ACM, 59(7), 96-104. doi:10.1145/2818717

[16] Kouloumpis, E., Wilson, T., & Moore, J. (2011, July). Twitter sentiment analysis: The good the bad and the omg!. In Fifth International AAAI conference on weblogs and social media.

[17] Gilbert, C. H. E., & Hutto, E. (2014, June). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth International Conference on Weblogs and Social Media (ICWSM-14). Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. vader. hutto. pdf (Vol. 81, p. 82).

[18] Chen E, Lerman K, Ferrara E Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set JMIR Public Health Surveillance 2020;6(2):e19273 DOI: 10.2196/19273 PMID: 32427106

[19] Nakamura, D. (2020, June 25). With 'kung flu,' Trump sparks backlash over racist language - and a rallying cry for supporters. Retrieved December 01, 2020, from https://www.washingtonpost.com/politics/with-kung-flu-trump-sparks-backlash-over-racist-language–and-a-rallying-cry-for-supporters/2020/06/24/485d151e-b620-11ea-aca5-ebb63d27e1ff_story.html

[20] Memon, S. A., & Carley, K. M. (2020). Characterizing covid-19 misinformation communities using a novel twitter dataset. arXiv preprint arXiv:2008.00791.

[21] Kearney, M. W., Dr. (2020, February 08). Detect Twitter Bots. Retrieved November 25, 2020, from https://tweetbotornot2.mikewk.com/

[22] Diamond, J. (2015, June 17). Donald Trump is running for president in 2016 - CNNPolitics. Retrieved December 12, 2020, from https://www.cnn.com/2015/06/16/politics/donald-trump-2016-announcement-elections/index.html