Investigating Motives for the Spread of Covid-19 Vaccine Misinformation:

A Democratic Public Health Policy Challenge

Parviz Ali

https://www.linkedin.com/in/parviz-ali-b7549010/

parviza9999@gmail.com

Abstract

People are being misinformed about the COVID-19 vaccine by a large amount of inaccurate information posted online. Individuals are being influenced by information from outside the formal health care system. Hundreds and thousands of new COVID-19 cases have created a second surge of the coronavirus pandemic in the misinformed population with vaccine hesitancy. We examined a wide range of sources including Google, websites of television, YouTube transcript, and newspaper in English using the search terms "COVID-19 vaccine misinformation" or "vaccine misinformation benefits", "Anti-vaxx", and "misinformation", "false vaccine information" for top underlying motives behind generating and perpetuating misinformation. We processed and examined data from 29 sources. We identified the top 12 sources responsible for originating vaccine misinformation. Our hierarchical clustering analysis identified four motive clusters 1) Religious, 2) Political, 3) Conspiracy Theorist, and 4) Alternative Medicine Seller with the highest weight (>.01). The result strongly indicates monetizing misinformation to be the top motive for the spread of COVID-19 misinformation. We propose using proven tools such as topic modeling, artificial intelligence, and machine learning technologies to track and analyze wide-ranging media data in real-time and effectively disrupt the flow of COVID-19 misinformation.

Keywords: Misinformation, Vaccine, COVID-19, Hesitancy, Anti-vaccine.

Introduction

Misinformation and wrong information have been the greatest challenge in society (Cacciatore 2021). As COVID-19 vaccines are rolled out across the world, there are growing concerns about the roles that trust, belief in conspiracy theories, and spread of misinformation through social media play in impacting vaccine hesitancy (Jennings et al. 2021). It is adversely affecting public health and public health policy, which depends on a well-informed public (Cook, Ecker, and Lewandowsky 2015). False information is originating from a number of sources including rumors, literary fiction, mainstream media, corporate-vested interests, governments, and nongovernmental organizations.

Given the large amount of inaccurate information online, people can easily become misinformed. Unproven information such as eating apricot seeds will cure cancer is a misconception that can be found online (Milazzo and Horneber 2015). There is no scientific evidence to support the claim; in fact, it is well established that eating apricot seeds may even cause cyanide poisoning (Vogel, Sultan, and Ten Eyck 1981). Information obtained from outside the formal health care system have influenced individuals in making critical decisions, and health misinformation and disinformation are not new.

The consequences of misinformation or wrong information have had devastating effects on public health. The best laid public health policy has often failed and its intended impact defeated because the beneficiaries of this service are misled by individuals or corporate-vested interest group, with selfish ulterior motives. The classic example of misinformation in public health is the misconception that the measles, mumps, rubella (MMR) vaccine causes autism, a concept popularized by a 1998 study published in *The Lancet* (Wakefield et al. 1998). The link was immediately refuted by the scientific community (Taylor et al.1999), and eventually the

publication itself was retracted, with the lead author being barred from practicing medicine. However, the misconception has gained substantial currency with predictable negative societal impact. In 2019, the United States saw multiple declarations of public health emergencies due to measles outbreaks ("De Blasio Administration's Health Department Declares Public Health Emergency Due to Measles Crisis" 2019).

Richard Doll and A. Bradford Hill, published an article in the British Medical Journal in 1950 (Doll and Hill 1950) that confirmed the link between smoking and carcinoma of the lung, commonly known as 'lung cancer'. In January 1954, US tobacco manufacturers, a corporate-vested interest group, jointly sponsored an advocacy advertisement entitled "A Frank Statement to Cigarette Smokers" (Cummings, Morley, and Hyland 2002) which appeared in 448 newspapers in 258 cities reaching an estimated 43,245,000 Americans. The advertisement questioned research findings implicating smoking as a cause of cancer, promised consumers that their cigarettes were safe, and pledged to support impartial research to investigate allegations that smoking was harmful to human health. Millions of premature deaths could have been prevented with unfettered implementation of public health policy on smoking if the blatant misinformation had not been allowed to influence public perception.

Although the practice of perpetrating misinformation such as the one carried out by tobacco manufacturers has been curtailed over the decades with the help of complex regulations, the ultimate goal or motive of achieving monetary benefits by perpetrating misinformation remains unchanged.

In recent time, internet sites such as social media platforms have become a common source for health information. People have been using social media to improve their knowledge about the disease, transmission, and prevention mechanisms (Islam et al. 2020). Health

information circulating on online platforms are often amplified by rumors and conspiracy theories that are not always based on scientific evidence (Lavorgna et al. 2018). A 2020 study conducted by Stecula and colleagues found that people who were exposed to vaccine-related information on social media were more likely to be misinformed and become vaccine-hesitant (Stecula, Kuru, and Hall Jamieson 2020). The spread of misleading information about the virus has led the World Health Organization (WHO) to warn of an on-going "infodemic" or an overabundance of information— especially misinformation—during an epidemic (van der Linden, Roozenbeek, and Compton 2020). Study suggests rumors and conspiracy theories related to COVID-19 vaccine around the globe on internet news articles, social media narratives, online reports and/or blogs are precipitators for vaccine hesitancy (Islam et al. 2021). Prior studies showed political beliefs and attitudes towards vaccines are well connected (Group, n.d.; Kennedy 2019). Study also suggests COVID-19 vaccine has been used for economic and political interest (Galvão 2021). An established motive far too familiar that follows the pattern of the corporate-vested interest group witnessed in the past is seen as a trend from the results of the data gathered for this study. Individuals instead of corporate, however, stand to gain from ad revenue generated by internet site visits (Burki 2020).

Data collection

A wide range of sources including Google, published articles, journals, websites of television, YouTube video transcript, and newspaper were reviewed. Reports in English using the following search terms "COVID-19 vaccine misinformation" or "vaccine misinformation benefits", "Anti-vaxx", and "misinformation", "false vaccine information" were used. For each combination of terms searched on Google, screening was limited to the first 10 pages for

evaluation. All information reported by the websites related to vaccine misinformation were reviewed for contents, documented, and entered into the database.

Data extraction and preprocessing

Once the vaccine-related items were identified, documents were processed and stop words were removed using python code and Orange 3.29 Data Mining. A document corpus consisting selected txt documents was created and saved as a csv file.

Data processing and analysis method

With the help of open-source Orange: Data Mining Toolbox in Python version 3.29 (fig. 1) texts were tokenized and vectorized. TF-IDF was used to determine top text terms related to vaccine misinformation. We used Orange to embed documents from input corpus into vector space by using pretrained FastText model (Grave et al. 2018) as an unsupervised learning approach to better understand the document as a whole. We constructed clusters of terms based on the rumors, statements, interviews and conspiracy theories reported. The data were then categorized under predefined themes: rumor or conspiracy theory, vaccine related morbidity and mortality, political and economic motives.

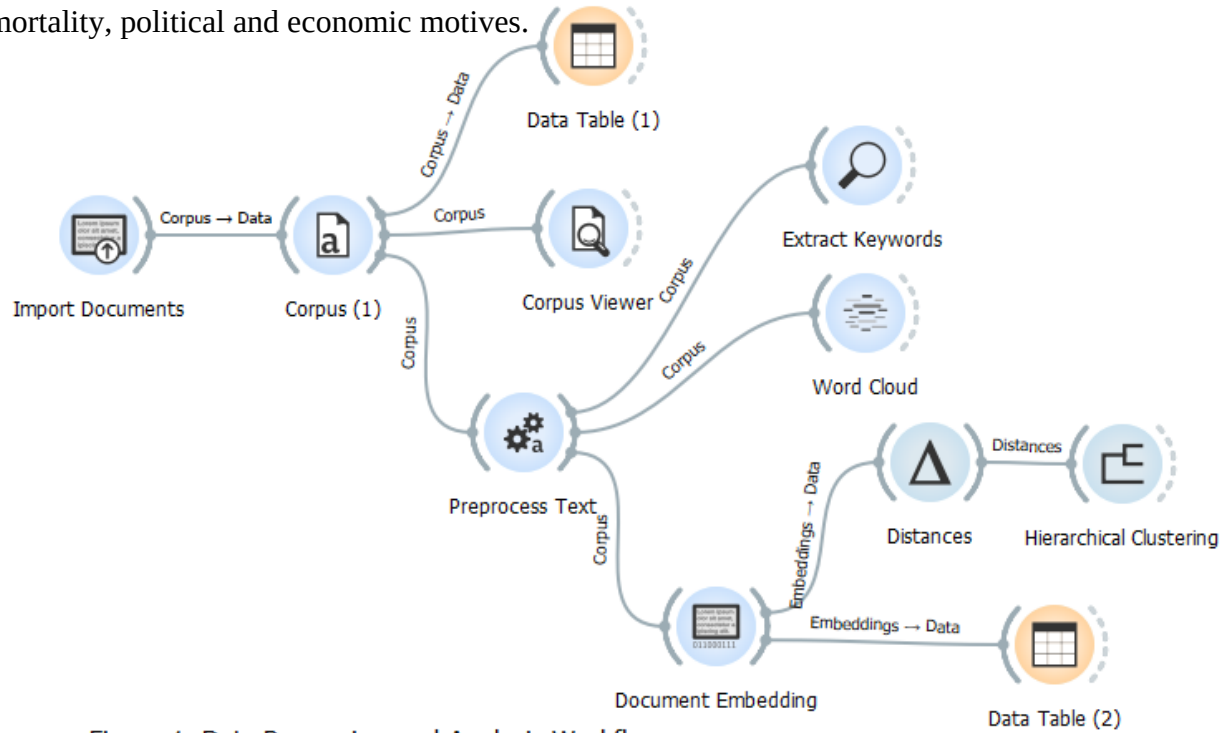

Figure 1. Data Processing and Analysis Workflow

Results

Total 29 documents comprised the text corpus. Using N-gram model with range 2 setting analysis Figure 2 shows the top themes of vaccine misinformation and the names of top spreaders of misinformation. A comparison of Rapid Automatic Keyword Extraction (RAKE)

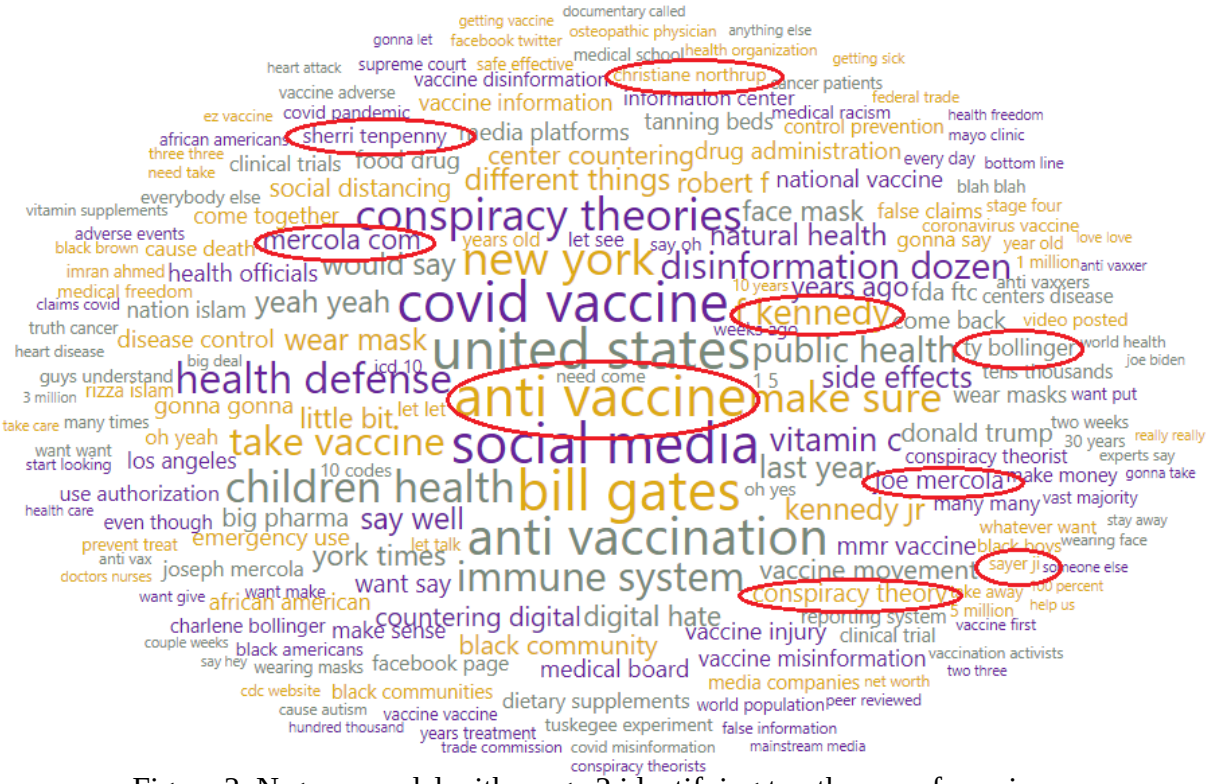

Figure 2. N-gram model with range 2 identifying top themes of vaccine misinformation and the names of top spreaders of misinformation.

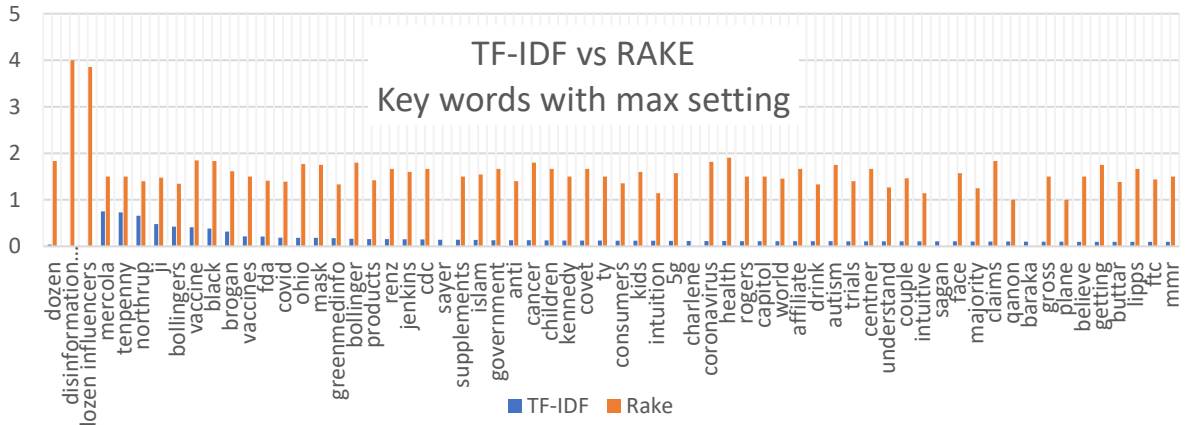

Figure 3. Comparison of TF-IDF and RAKE (Rapid Automatic Keyword Extraction) output.

and TF-IDF algorithms in Figure 3, for finding keywords in text, shows the top anti-vaccine spreaders and their key messages. The “disinformation dozen” (“The Disinformation Dozen” n.d.), as termed by the Center for Countering Digital Hate (CCDH) also appears in the RAKE term.

Based on the estimated number of followers analyzed by Alexa.com (Cooper n.d.) the projected ad-revenue of the top disinformation spreaders is as high as $36,000 from merchandise sale and product endorsements. Based on $10 income per 1000 followers estimated by Webfx.com (“Influencer Marketing Pricing: What Does It Cost in 2021?” n.d.) Figure 4 shows the estimated income of the top anti-vaccine social media influencers as found in our TF-IDF and RAKE analysis.

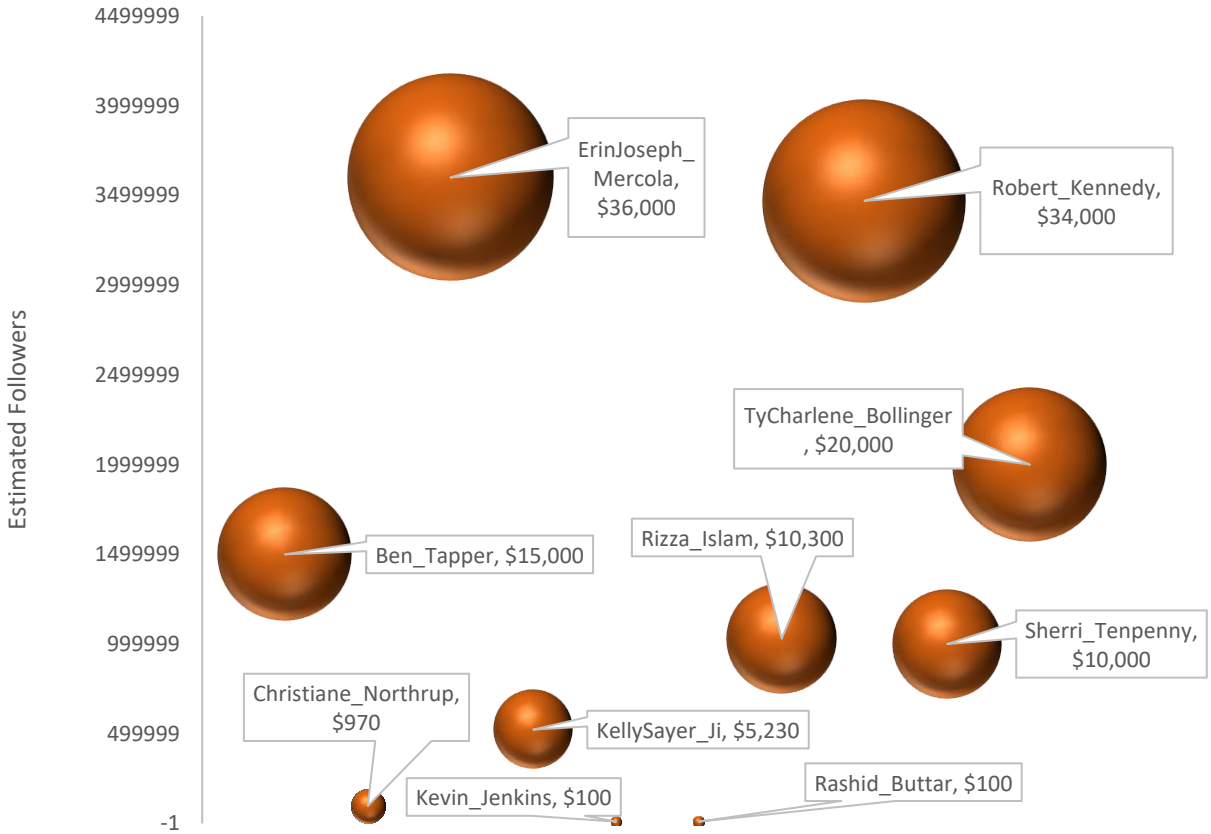

Figure 4. Anti-Vaccine Social Media Influencers with Estimated Income from Advertisement

Anti-vaccine social media influencers were categorized in four clusters based on reported data of their affiliation and statement (see Table 1).

Table 1. Affiliation and statement-based clusters

	Clusters
1	Religious
2	Political
3	Conspiracy Theorist
4	Alternative Medicine Seller

Corpus data was embedded with mean setting. Cosine distance metric was then applied to the data. Hierarchical Clustering analysis run with weighted linkage setting resulted in four distinct clusters emerge from the corpus data. Individuals identified in Figure 3 were ranked and placed in categories identified in Table 1. The result is shown in Figure 5.

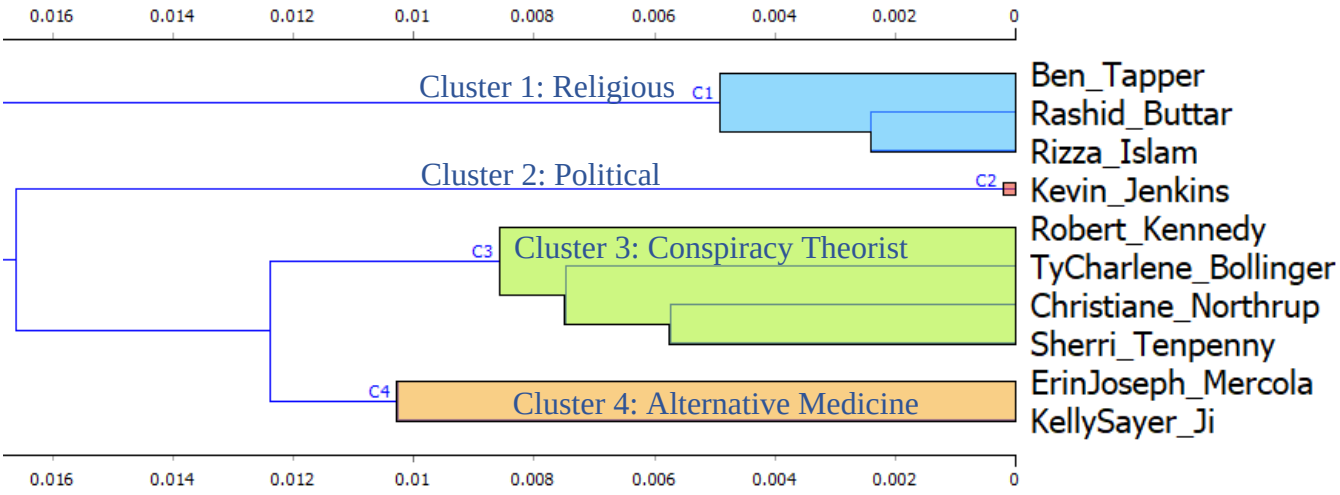

Figure 5. Clusters of vaccine misinformation motives.

Discussion

The study confirmed the top social media influencers identified by online reports. Twelve main social media influencers leading the COVID-19 vaccine misinformation campaign were detected in 2020 by the Center for Countering Digital Hate (CCDH) and identified in its report. This report led Facebook, Instagram, Twitter, and Google to remove direct statements related to COVID-19 vaccine misinformation posted by these influencers. One website "mercola.com" now deletes all contents after 48 hours as result of regulatory warning (Center for Food Safety, and Applied Nutrition. 2021). This limited our firsthand report text data and our corpus data to media report and video interviews. We extracted closed caption text data from interview videos and further processed them to eliminate noise such as "uh", "um" and "(__)" substituted for long pause which further limited our corpus size.

Four clusters emerged from hierarchical clustering of corpus (10 instances, 300 features) data which we believe are to be the top motives for COVID-19 vaccine misinformation. In cluster 1 or "Religious" cluster, we notice the model included Dr. Ben Tapper, Dr. Rashid Butter, and Rizza Islam. Islam is overlapping with cluster 2, or "Political" cluster. Cluster 2 has included Kevin Jenkins, a faith-based political activist who is overlapping with Cluster 3 or "Conspiracy Theorist" cluster. Cluster 3 includes Dr. Christiane Northrup, Ty and Charlene Bollinger (entrepreneur, alternative medicine sales), Robert Kennedy, and Dr. Sherri Tenpenny. Cluster 4 or "Alternative Medicine Seller" overlapping cluster 3 has the highest weight (>.01). Erin Elizabeth, Joseph Mercola, Kelly Brogan, and Sayer Ji. One of the highest grossing social media influencers (see Fig. 4), Erin Elizabeth and Joseph Mercola, belong to this cluster supporting our claim of top motives.

Conclusions

Misinformation related to COVID-19 is damaging with negative consequences. It is an important area to focus research on for public policy making. This study adds to extant knowledge by finding evidence of differential motives for COVID-19 narratives among US adults. The above analysis is not an exhaustive study of all websites pushing vaccine misinformation. It is only a small sample. Nonetheless, the analysis demonstrates different ways that anti-vaccination websites are building audiences and, in some cases, monetizing their sites. And as long as there are ways to make money or gather an audience off misinformation, the incentive to publish and push misleading content will remain.

In conclusion, our study provided a snapshot of misinformation circulating online that have built public distrust in COVID-19 vaccine. The diverse amount of circulating COVID-19 vaccine misinformation has undermined the universal rollout of the COVID-19 vaccine candidates. We suggest that traditional methods of risk communication and community engagement should be explored to track and fact-check misinformation as ways to immunize people against misinformation and thereby pre-empt potential vaccine program disruptions. Policymakers should consider these findings to devise risk communication and community engagement strategies to address these concerns with evidenced-based information. Additionally, topic modelling, artificial intelligence, and machine learning technologies are proven tools to track and analyze large media data in real-time. They have effectively disrupted the flow of COVID-19 misinformation as noted in our discussion. With this lesson learned, a permanent policy must be instituted to stem the tide of the spread and circulation of misinformation in social media at an early stage to ensure higher effectiveness of the public policy.

Directions for future work

This study has several limitations. We primarily relied on second-hand account websites, national and international news agencies, and TV channels. All of the direct claims related to COVID-19 misinformation social media platforms were found deactivated during our data collection period. Also, our study may not have detected all misinformation and conspiracy theories circulating online in other native languages. Therefore, the prevalent rumors and conspiracy theories detected through our surveillance could have underestimated the true motives and impacts. We were also unable to establish a direct correlation between an increase in social media followers and ad revenue increase. A topic we would like to explore in the future.

References

Burki, Talha. “The Online Anti-Vaccine Movement in the Age of COVID-19.” *The Lancet Digital Health* 2, no. 10 (October 1, 2020): e504–5. https://doi.org/10.1016/S2589-7500(20)30227-2.

“De Blasio Administration’s Health Department Declares Public Health Emergency Due to Measles Crisis.” 2019. Nyc.Gov. April 9, 2019. https://www1.nyc.gov/office-of-the-mayor/news/186-19/deblasio-.

Cacciatore, Michael A. 2021. “Misinformation and Public Opinion of Science and Health: Approaches, Findings, and Future Directions.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (15): e1912437117.

Center for Food Safety, and Applied Nutrition. 2021. “Mercola.Com, LLC - 607133 - 02/18/2021.” Fda.Gov. April 3, 2021. https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/mercolacom-llc-607133-02182021.

Cook, John, Ullrich Ecker, and Stephan Lewandowsky. 2015. “Misinformation and How to Correct It.” *Emerging Trends in the Social and Behavioral Sciences*. Wiley. https://doi.org/10.1002/9781118900772.etrds0222.

Cooper, Kim. n.d. “Keyword Research, Competitor Analysis, & Website Ranking.” Alexa.Com. Accessed August 22, 2021. https://www.alexa.com/.

Cummings, K. M., C. P. Morley, and A. Hyland. 2002. “Failed Promises of the Cigarette Industry and Its Effect on Consumer Misperceptions about the Health Risks of Smoking.” *Tobacco Control* 11 Suppl 1 (suppl 1): I110-7.

Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python, *Journal of Machine Learning Research* 14(Aug): 2349−2353.

Doll, R., and A. B. Hill. 1950. “Smoking and Carcinoma of the Lung; Preliminary Report.” *British Medical Journal* 2 (4682): 739–48.

Galvão, Jane. 2021. “COVID-19: The Deadly Threat of Misinformation.” *The Lancet Infectious Diseases* 21 (5): e114.

Grave, Edouard, Piotr Bojanowski, Prakhar Gupta, Armand Joulin and Tomas Mikolov. “Learning Word Vectors for 157 Languages.” *ArXiv* abs/1802.06893 (2018): n. pag.

Group, C. n.d. “A Future Vaccination Campaign against COVID-19 at Risk of Vaccine Hesitancy and Politicisation.” *The Lancet Infectious Diseases* 2020;20(7):769–70.

“Influencer Marketing Pricing: What Does It Cost in 2021?” n.d. Webfx.Com. Accessed August 22, 2021. https://www.webfx.com/influencer-marketing-pricing.html.

Islam, Md Saiful, Tonmoy Sarkar, Sazzad Hossain Khan, Abu-Hena Mostofa Kamal, S. M. Murshid Hasan, Alamgir Kabir, Dalia Yeasmin, et al. 2020. “COVID-19-Related

Infodemic and Its Impact on Public Health: A Global Social Media Analysis." *The American Journal of Tropical Medicine and Hygiene* 103 (4): 1621–29.

Jennings, Will, Gerry Stoker, Hannah Bunting, Viktor Orri Valgarðsson, Jennifer Gaskell, Daniel Devine, Lawrence McKay, and Melinda C. Mills. 2021. "Lack of Trust, Conspiracy Beliefs, and Social Media Use Predict COVID-19 Vaccine Hesitancy." *Vaccines* 9 (6): 593.

Kennedy, Jonathan. 2019. "Populist Politics and Vaccine Hesitancy in Western Europe: An Analysis of National-Level Data." *European Journal of Public Health* 29 (3): 512–16.

Lavorgna, L., M. De Stefano, M. Sparaco, M. Moccia, G. Abbadessa, P. Montella, D. Buonanno, et al. 2018. "Fake News, Influencers and Health-Related Professional Participation on the Web: A Pilot Study on a Social-Network of People with Multiple Sclerosis." *Multiple Sclerosis and Related Disorders* 25: 175–78.

Linden, Sander van der, Jon Roozenbeek, and Josh Compton. 2020. "Inoculating against Fake News about COVID-19." *Frontiers in Psychology* 11: 566790.

Milazzo, Stefania, and Markus Horneber. 2015. "Laetrile Treatment for Cancer." *Cochrane Database of Systematic Reviews,* no. 4: CD005476.

Stecula, Dominik Andrzej, Ozan Kuru, and Kathleen Hall Jamieson. 2020. "How Trust in Experts and Media Use Affect Acceptance of Common Anti-Vaccination Claims." *Harvard Kennedy School Misinformation Review* 1 (1). https://doi.org/10.37016/mr-2020-007.

Taylor, B., E. Miller, C. P. Farrington, M. C. Petropoulos, I. Favot-Mayaud, J. Li, and P. A. Waight. 1999. “Autism and Measles, Mumps, and Rubella Vaccine: No Epidemiological Evidence for a Causal Association.” *Lancet* 353 (9169): 2026–29.

“The Disinformation Dozen.” n.d. Counterhate.Com. Accessed August 22, 2021. https://www.counterhate.com/disinformationdozen.

Vogel, S. N., T. R. Sultan, and R. P. Ten Eyck. 1981. “Cyanide Poisoning.” *Clinical Toxicology* 18 (3): 367–83.

Wakefield, A. J., S. H. Murch, A. Anthony, J. Linnell, and D. M. Casson. 1998. “RETRACTED: Ileallymphoid- Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children.” *Lancet* 351: 637–41.