



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Short Communication

## An analysis of AstraZeneca COVID-19 vaccine misinformation and fear mongering on Twitter

D. Jemielniak <sup>a, b, \*</sup>, Y. Krempovych <sup>a</sup><sup>a</sup> Kozminski University, Management in Networked and Digital Societies (MINDS), Warsaw, Poland<sup>b</sup> Harvard University, Berkman-Klein Center for Internet and Society, Cambridge, MA, USA

## ARTICLE INFO

## Article history:

Received 4 May 2021

Received in revised form

27 July 2021

Accepted 23 August 2021

Available online 30 August 2021

## Keywords:

AstraZeneca  
COVID-19 vaccine  
Disinformation

## ABSTRACT

**Objectives:** The objective of this study was to analyse the media discourse about the AstraZeneca COVID-19 vaccine on Twitter.**Study design:** The study design used in this study is data scraping, media analysis, social network analysis, and botometer.**Methods:** We collected 221,922 tweets containing '#AstraZeneca' from 1 January 2021 to 22 March 2021. From 50,080 tweets in the English language, we analysed the linked media sources and conducted a network detection study.**Results:** We found that the most frequently retweeted tweets were full of negative information, and in many cases came from media sources that are well-known for misinformation. Our analysis identified large coordination networks involved in political astroturfing and vaccine diplomacy in South Asia but also vaccine advocacy networks associated with European Commission employees.**Conclusions:** The results of this study show that Twitter discourse about #AstraZeneca is filled with misinformation and bad press, and may be distributed not only organically by anti-vaxxer activists but also systematically by professional sources.

© 2021 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

## Introduction

Vaccination debates are prone to misinformation and this has been especially true during the COVID-19 pandemic.<sup>1,2</sup> A particularly interesting case is the AstraZeneca COVID-19 vaccine, which has had significant media coverage related to the tensions over EU-UK exports, as well as an alleged link to very rare cases of blood clots.

To study media coverage of the AstraZeneca COVID-19 vaccine, we analysed tweets containing '#AstraZeneca' and links in the most frequently retweeted tweets. Additionally, we investigated co-tweet networks and revealed bot activity related to the discourse.

## Methods

We conducted a Thick Big Data<sup>3</sup> analysis of the AstraZeneca hashtag on Twitter. Thick Big Data is a mixed-method analysis, combining large data set computational analysis with qualitative

insight into specific tweets (selected based on quantitative criteria). We used a Python GetOldTweets3 script to collect 221,922 tweets with '#AstraZeneca' from 1 January 2021 to 22 March 2021. We focused on 50,080 tweets in the English language. Furthermore, we established the final URLs from these tweets by extracting them from popular link shorteners.

We focused on time periods before and after 7 March 2021 because this was the date when Austrian authorities took a precautionary step to suspend vaccinating with a batch of AstraZeneca's vaccine.

Furthermore, we studied which media sources were most often linked in frequently retweeted tweets (more than 10 times for both time periods). We also analysed the most liked and most retweeted tweets qualitatively, which enabled us to determine which content receives the largest public exposure and backing.

Finally, we concluded Coordination Detection analysis<sup>4</sup> to identify cooperative efforts to propagate certain tweets. To detect coordination, we conducted similarity analysis of tweet corpuses using the Min-Hash<sup>5</sup> local-sensitive hashing method that is useful to process large sets of textual data. As a result, we were able to identify similar items based on the Jaccard similarity between sets of strings' hashtag n-grams – a commonly used statistical language model that can be used to distinguish text strings. We adopted a 0.8

\* Corresponding author. Kozminski University, Management in Networked and Digital Societies (MINDS), Warsaw, Poland.

E-mail addresses: [djemielniak@cyber.harvard.edu](mailto:djemielniak@cyber.harvard.edu) (D. Jemielniak), [ykrempovych@kozminski.edu.pl](mailto:ykrempovych@kozminski.edu.pl) (Y. Krempovych).

Jaccard similarity threshold as proposed in the method introduced by Pacheco et al.<sup>4</sup>

To project each tweet's coordination network, we drew an edge between two accounts with matching tweet text corpus using co-occurrence analysis functionality from the quanteda package for R. In the output, we reached a graph with 4200 nodes representing tweets with a similar or identical corpus, which accounted for >11,000 unique Twitter user handles.

We also performed bot identification with Botometer, a machine-learning platform that computes a bot likeliness ranking. Botometer measures the score by comparing an account against tens of thousands of labelled inputs in its database.<sup>5</sup> Considering prior coordination detection analysis that we performed, a Complete Automation Probability score was set to  $\geq 0.76$ .

### Results

Before 7 March 2021, in the top ten linked sources among the tweets retweeted more than 10 times, four Western news media sources were identified as follows: AFP (ranked no. 18,806 most popular website according to Alexa traffic ranking), Politico.eu (no. 20,905 in Alexa), the Telegraph (no. 1730 in Alexa), and the Guardian (no. 172 in Alexa).

After 7 March 2021, 7 NewsRand, a relatively unpopular Nigerian news website (ranked no. 126,271 in Alexa), was more frequently linked than the four Western news media sources above, which all dropped out of the top ten linked sources. More interestingly, GreatGameIndia (ranked no. 126,482 in Alexa), an Indian website that has previously been described as spreading disinformation and in particular COVID-19 fake news,<sup>7</sup> was number 5 before 7 March 2021, increased to number 2 in the period after, and included tweets linking AstraZeneca to eugenics.

RT (formerly Russia Today, ranked no. 312 in Alexa) had a clear lead in both time periods. RT is a state-owned media outlet that has been described as supporting Russian diplomatic goals as an information warfare tool.<sup>8</sup> A qualitative analysis of RT links showed that they are predominantly negative when referring to the AstraZeneca vaccine, even though they repeat actual news rather than pure disinformation.

Outside media links, the most retweeted (2656 retweets as of 26 March 2021) tweet overall from the first time period was one by Robert Kennedy Jr, a known anti-vaxxer advocate, whose account on Instagram was terminated in February 2020 because of COVID-19 disinformation. The tweet was discrediting AstraZeneca vaccine as 'controversial', 'heavily invested in by Bill Gates' and 'being rejected over widespread concerns'. The most retweeted (2015 retweets as of 26 March 2021) tweet from the second period was one by Disclose.tv, a site described as involved in disinformation.<sup>9</sup>

Tweet coordination analysis has revealed 10,728 instances in the coordination carried out by 1137 unique handles, of which 2278 instances and 616 unique handles are related to automatic bot accounts, according to Botometer Complete Automation Probability score of  $\geq 0.76$ . The largest coordination network had 451 coordination instances by 111 accounts, of which 74 scored above our threshold on the Botometer. The second-largest network consisted of 37 accounts responsible for 47 coordination instances; of these 37 accounts, 13 accounts can be considered automated (see Fig. 1).

The first coordination network connects accounts that tweet about current vaccination situations in Bangladesh, praising the Bangladesh Prime Minister, Sheikh Hasina, for fighting the pandemic with the '#ThankYouPM' hashtag, which is often used on Twitter to refer to India's Prime Minister, Narendra Modi.

Another topic presented in this coordination network is large vaccine donations by India, one of the largest vaccine producers, to

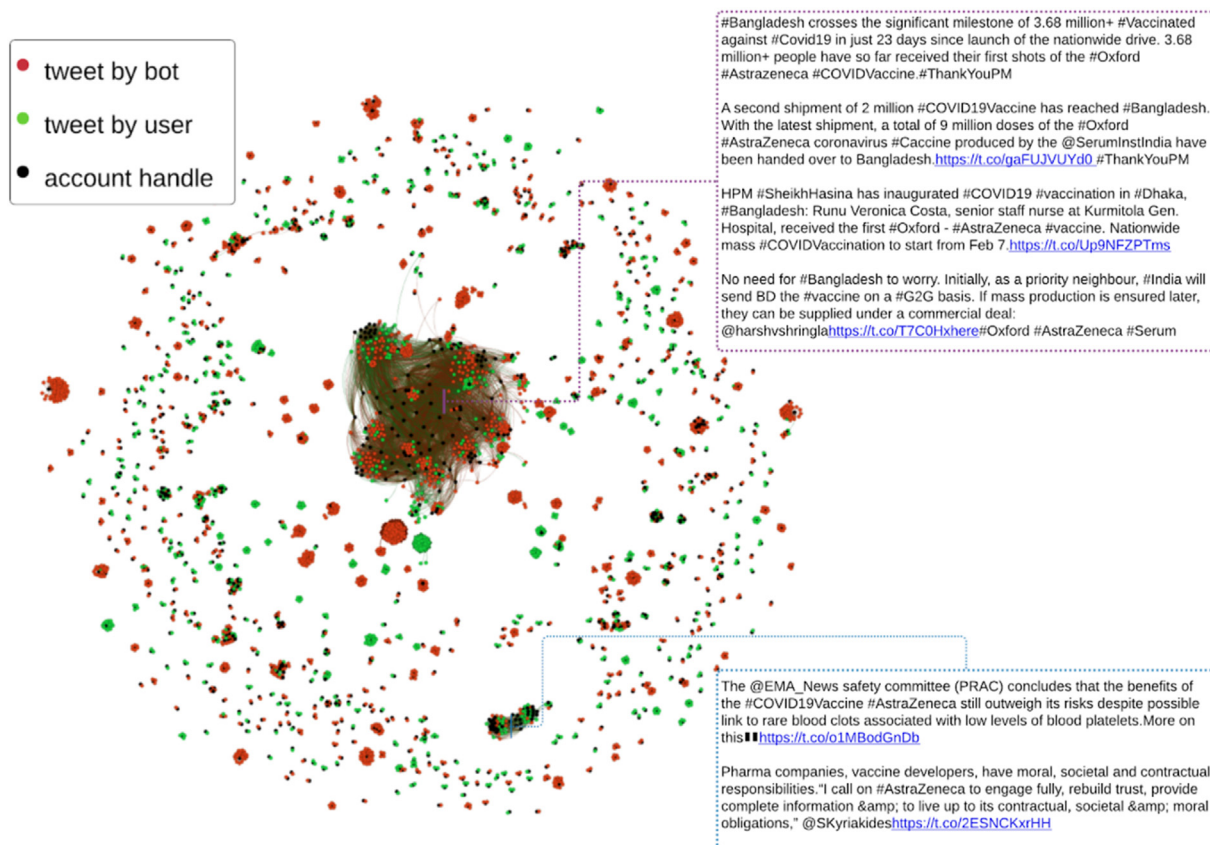


Fig. 1. Co-tweets networks discovered in the data set.

Bangladesh. The Indian government is using vaccine distribution to strengthen its ties through vaccine diplomacy.<sup>10</sup>

In both cases presented above, automated bot accounts are engaged in an activity of political astroturfing<sup>11</sup> and also involved a co-tweet coordination network. Automated bot accounts spread identical or very similar texts to amplify political message reach: in this case, vaccine diplomacy.

When it comes to the second-largest coordinated network, the ratio of accounts involved to the number of coordination instances is much smaller while representing only two coordinated messages. Those messages referred to the official stand of European Medicines Agency regarding safety and efficacy of the AstraZeneca vaccine.

This is a network composed predominantly of European Commission's employee's accounts and high-level officials, even though 35% of these accounts are considered in this analysis to be automated to some extent. Moreover, research of shared links indicated that all the messages connected to the article 'Remarks by Commissioner Stella Kyriakides on vaccines'.<sup>12</sup> All these messages used shortened URL from the domain 'smh.re' that indicates the use of employee advocacy software Smarp, which helps employees to coordinate corporate messages through their private accounts. Considering the facts mentioned earlier, one can safely assume that this network presented a centralised health advocacy communication campaign coordinated by the European Commission's employees to address increased criticism of AstraZeneca vaccines.

Astroturfing, in addition to coordination of activities in social media, sway public opinion.<sup>13</sup> While those actions can be used to disseminate political propaganda, as in the case of coordination activity in Bangladesh, they can also be helpful for health advocacy campaigns, as in the case of European Commission employees. The methods, goals, and degree of automation can be different – in both cases, coordination techniques were used while aiming to appear as organic content of Twitter users. These findings bring new insights into the use of coordinated activity in social media in the context of diplomacy, politics, and health advocacy during the global COVID-19 pandemic.

## Discussion

Our results focus on a short time period and only on one vaccine-related hashtag. Nevertheless, the picture that emerges is deeply troubling. Twitter discourse about #AstraZeneca abounds in misinformation and reputable media news sources representation is, at best, on par with the misinforming sources, and, at worst, significantly smaller.

Popular fear-mongering tweets are spread not only by individual powerful activists and conspiracy websites but also by state-owned media, supported by bot networks. Given the fact that Russia has a heavy interest in promoting its own Sputnik-V vaccine, both for economic and political reasons, the activity of RT in posting often negative information about #AstraZeneca may be perceived as part of a larger campaign, potentially aimed at discrediting the vaccine. Clearly, India and the EU also see the potential in public health online campaigning.

Our research highlights the following three points: first, professional vaccine misinformation more frequently relies on disproportionate reporting of negative news rather than brute

disinformation. Second, coordinated networks and bots are routinely used for vaccine communication and, without regulation or clear counteractions from social network, the socially damaging misinformation arm race is unavoidable. Third, future research should focus on other vaccines as well as analysing other topics that are covered by the identified coordination networks. Our research is limited to one case on one social network site; thus, it is essential for future studies to have a wider scope. Nevertheless, it is already evident that developing new interventions, such as continuous monitoring of coordinated networks, is required to detect and reduce misinformation in the public health discourse.

## Author statements

### Ethical approval

Not applicable.

### Funding

This work was possible thanks to a grant no. GOSPOSTRATEG-II/0007/2020-00 from Polish National Center for Research and Development.

### Competing interests

None declared.

## References

1. Carrion-Alvarez D, Tijerina-Salina PX. Fake news in COVID-19: a perspective. *Health Promot Perspect* 2020;**10**(4):290–1.
2. Ashton J. COVID-19 and the anti-vaxxers. *J R Soc Med* 2021;**114**(1):42–3.
3. Jemielniak D. *Thick Big Data: doing digital social sciences*. Oxford: Oxford University Press; 2020.
4. Pacheco D, Hui P-M, Torres-Lugo C, Truong BT, Flammini A, Menczer F. Uncovering coordinated networks on social media. *arXiv [cs.SI]* 2020. Available at: <http://arxiv.org/abs/2001.05658>.
5. Broder AZ. On the resemblance and containment of documents. In: *Proceedings. Compression and complexity of SEQUENCES 1997 (Cat. No. 97TB100171)*. IEEE; 1997. p. 21–9.
6. Sayyadharikandeh M, Varol O, Yang K-C, Flammini A, Menczer F. Detection of novel social bots by ensembles of specialized classifiers. In: *Proceedings of the 29th ACM international conference on information & knowledge management*. New York, NY, USA: Association for Computing Machinery (CIKM '20); 2020. p. 2725–32.
7. Deutch G. *How one particular coronavirus myth went viral, wired, 19 March 2020*. Available at: <https://www.wired.com/story/opinion-how-one-particular-coronavirus-myth-went-viral/>. [Accessed 26 March 2021].
8. Golovchenko Y, Hartmann M, Adler-Nissen R. State, media and civil society in the information warfare over Ukraine: citizen curators of digital disinformation. *Int Aff* 2018:975–94. <https://doi.org/10.1093/ia/iiy148>.
9. Guarino S, Pierri F, Di Giovanni M, Celestini A. Information disorders during the COVID-19 infodemic: the case of Italian Facebook. *Online Soc Network Media* 2021;**22**:100124.
10. Analytica O. Vaccine exports will boost India's neighbourhood ties. *Emerald Expert Briefings* 2021;**39**:88.
11. Keller FB, Schoch D, Stier S, Yang J. Political astroturfing on Twitter: how to coordinate a disinformation campaign. *Polit Commun* 2020;**37**(2):256–80.
12. "Press Corner." n.d. [https://ec.europa.eu/commission/presscorner/detail/en/SPEECH\\_21\\_267](https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_21_267). [Accessed April 2, 2021].
13. Zerback T, Töpfl F, Knöpfle M. The disconcerting potential of online disinformation: persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media Soc* 2021:1080–98. <https://doi.org/10.1177/1461444820908530>.