Vaccination trials on hold: malicious and low credibility content on Twitter during the AstraZeneca COVID-19 vaccine development

Sameera Horawalavithana[1] · **Ravindu De Silva**[2] · **Nipuna Weerasekara**[2] · **N G Kin Wai**[1] · **Mohamed Nabeel**[3] · **Buddhini Abayaratna**[2] · **Charitha Elvitigala**[2] · **Primal Wijesekera**[4] · **Adriana Iamnitchi**[5]

Abstract

The development of COVID-19 vaccines during the global pandemic that started in 2020 was marked by uncertainty and misinformation reflected also on social media. This paper provides a quantitative evaluation of the Uniform Resource Locators (URLs) shared on Twitter around the clinical trials of the AstraZeneca vaccine and their temporary interruption in September 2020. We analyzed URLs cited in Twitter messages before and after the temporary interruption of the vaccine development on September 9, 2020 to investigate the presence of low credibility and malicious information. We show that the halt of the AstraZeneca clinical trials prompted tweets that cast doubt, fear and vaccine opposition. We discovered a strong presence of URLs from low credibility or malicious websites, as classified by independent fact-checking organizations or identified by web hosting infrastructure features. Moreover, we identified what appears to be coordinated operations to artificially promote some of these URLs hosted on malicious websites.

Keywords AstraZeneca vaccine · Coordinated URL promotion · Misinformation

1 Introduction

Throughout the COVID-19 pandemic that started in early 2020, social media platforms have played a critical role in widely spreading information, regardless of its veracity (Rosenberg et al. 2020). As COVID-19 vaccines moved through development, a

✉ Sameera Horawalavithana
    sameera1@usf.edu

✉ Adriana Iamnitchi
    a.iamnitchi@maastrichtuniversity.nl

Extended author information available on the last page of the article

significant surge in misinformation and anti-vaccine narratives became evident on Twitter (Smith et al. 2020; Bagherpour 2020). Twitter discussions about AstraZeneca, a vaccine developed at Oxford University, were no exception to this as the vaccine has been mired in controversy ever since its inception (Jemielniak and Krempovych 2021).

AstraZeneca's vaccine development faced numerous challenges due to a temporary halt of its trials in September 2020 caused by an unexplained illness in one of the participants (Reuters 2021; Robbins et al. 2020). Details surrounding the event were not thoroughly disclosed, which lead scientists to question the transparency of the vaccine development efforts (Cyranoski and Mallapaty 2020). The temporary interruption of the AstraZeneca vaccine trials also sparked political debates as government officials, especially in the United States, had been trying to fast-track its approval and roll-outs (Forbes 2020). As AstraZeneca quickly resumed its clinical trials and entered its final phase of development, safety issues and concerns about its efficacy intensified (Cyranoski and Mallapaty 2020; PBS 2020).

This paper investigates the information shared on Twitter in a period that marked an important step in the development of the AstraZeneca vaccine. We ask several research questions in this study. *To what extent is the presence of low credibility and malicious information about Astrazeneca vaccine related discussions in Twitter due to its global halt/suspension? What are the properties of the URLs that share low credibility and malicious information? Is there any coordinated behavior for promoting bad quality information, if so, how persistent is such behavior?*

We focus on two time periods, before and after the temporary interruption of the vaccine development on September 9, 2020. While our focus in on AstraZeneca-related discussions, we also detected tweets that mentioned SputnikV, Moderna and Pfizer vaccines with stories that cast doubt and fear towards the safety and efficacy of the vaccines. A deeper analysis shows that vaccine-related stories are promoted by both low credibility and malware-hosted websites. In contrast to the low credibility websites that often promote vaccination myths and conspiracy stories, the malware hosted on other websites can be used to trace the web searching activities of the Twitter audience interested in COVID-19 vaccines. For example, the properties of the most popular malicious URLs in two time periods are quite similar, such as short-lived newly created domains, self-signed certificates, content management system and hosting geographical location.

Some user accounts use this information to target specific communities with stories promoting certain vaccines favorably, while questioning others. For example, Russian-sponsored media outlets took advantage of the halt of trials to push narratives discrediting the AstraZeneca vaccine and boost the promotion of SputnikV instead. We discovered persistent groups of users engaged in the rapid propagation and artificial inflation of a particular URL through simultaneous tweeting. This behavior suggests potential signs of coordination to promote malicious and low-credibility URLs on Twitter.

2 Related work

Along factual information, misinformation and disinformation related to COVID-19 regularly circulate in social media (Huang and Carley 2020). As vaccines became available, (mis)disinformation surrounding them started to increase and persist on Twitter (DeVerna et al. 2021). Combating the spread of (mis)disinformation and conspiracy theories regarding COVID-19 on social media has become a global concern as, even in the absence of any scientific evidence, it has increased the number of people hesitant to get vaccinated (Wilson and Wiysonge 2020; Chadwick et al. 2021).

Numerous research efforts aimed to assess the prevalence of misinformation related to COVID-19 in social media. Pranesh et al. (2021) quantified the magnitude of misinformation presence in COVID-19 related tweets across different languages. Karami et al. (2021) showed that vaccine discussions on Twitter are evolving with negative and non-negative attitudes on different issues. Their analysis revealed that tweets on topics related to vaccination sites, getting vaccines, and vaccine effectiveness exhibit positive sentiment, while tweets about government strategies related to vaccination, vaccination hesitancy, vaccine immunity, and social distancing often sparked negative discussions. Yang et al. (2020) and Sharma et al. (2022) investigate URLs embedded in tweets to evaluate the credibility of the content at domain level. These studies highlight that there is a comparable presence of both low-quality sources and sources from mainstream media in Twitter. In addition, they found that social bots are most likely to engage with the promotion and amplification of low credibility information.

Other studies analyzed the extent to which different users engage with misinformation. Teng et al. (2022) highlight that social bots' contribution to misinformation is surprisingly low. The authors identified two user groups who engage with misinformation: the strong-adherent users, who aim at supplying or/and promoting misinformation for certain purposes, and the weak-adherent individuals, who are occasionally exposed to misinformation but are easily triggered to further spread misinformation. Echoing the same sentiment, Silva et al. (2020) showed that the significant majority of both factual and misinformation tweets were generated by real users (not bots) even though bot accounts tweeted more misinformation (8.5%) relative to factual information (5.2%). Contrary to these findings, Yang et al. (2020) observed a higher-than-normal participation rate of bot accounts in both posting and amplifying low-credibility content. Their results also show that bot-like tweeters attract more bot-like retweeters than human-like tweeters.

This paper builds on our previous analysis of Twitter discussions during a period of uncertainty in the AstraZeneca vaccine development (Horawalavithana et al. 2021). In contrast to many previous research, we characterize low quality content via both malicious and low credibility URL analysis to identify tweets which escalate community concerns with regards to vaccine efficacy. We also analyze the coordinated link-sharing behavior by revealing groups of users who share the same URL within unusually short time.
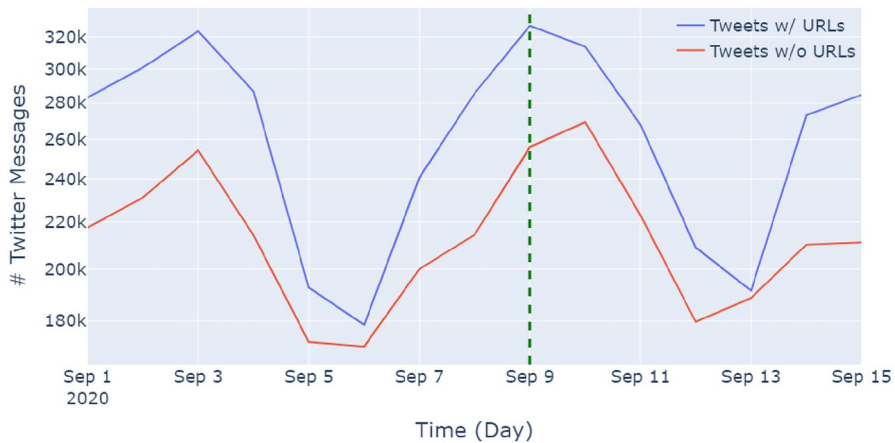
Fig. 1 Number of Twitter messages over time. AstraZeneca vaccination trial halted date is highlighted in the green dashed vertical line

3 Data collection and processing

The mistrust in what was perceived as rushed vaccine development during the COVID-19 pandemic has been reported to undermine people's intention to receive the COVID-19 vaccines (Vivek 2021). In addition, concerns related to the side effects reported in vaccine clinical trials raised questions about their safety (Nuzhath et al. 2020). One such major incident was reported in September 2020, when an unexplained illness appeared during the clinical trials for the AstraZeneca vaccine.

We collected tweets using the Twitter API from 1st of September 2020 to 15th of September 2020. The original keywords used to collect this dataset are *AstraZeneca, "Astra Zeneca", AZD1222, COVID, vaccine, immunity, "herd immunity", Barrington,* and *"focused protection".* We used the same keyword list used to collect a similar dataset released as part of the 2021 Grand Challenge of the North American Social Network Conference (NASN) (NASN 2021). Our data collection cover many missing tweets in the NASN dataset after September 9th 2021 (Horawalavithana et al. 2021). We used the vaccine-related keywords in the regex condition: *pfizer OR astrazeneca OR moderna OR (sputnik AND vaccine)* to select tweets relevant to our study.

We grouped the Twitter messages into two disjoint subsets based on their inclusion of URLs. The first subset contains 3,212,586 Twitter messages (by 1,788,788 users) without URLs. The second subset consists of 3,958,864 messages (55% of all messages) that contain at least one URL. These messages are shared by 1,267,873 users and cite 3,247,946 distinct URLs from 167,186 distinct web domains. The number of such messages per day are presented in Fig. 1.

We pre-processed this dataset to identify hashtags, user mentions, and URL domains. We eliminated the URLs which link to other tweets. In addition, the external links (e.g., a tweet mentioning a YouTube video, or an external website
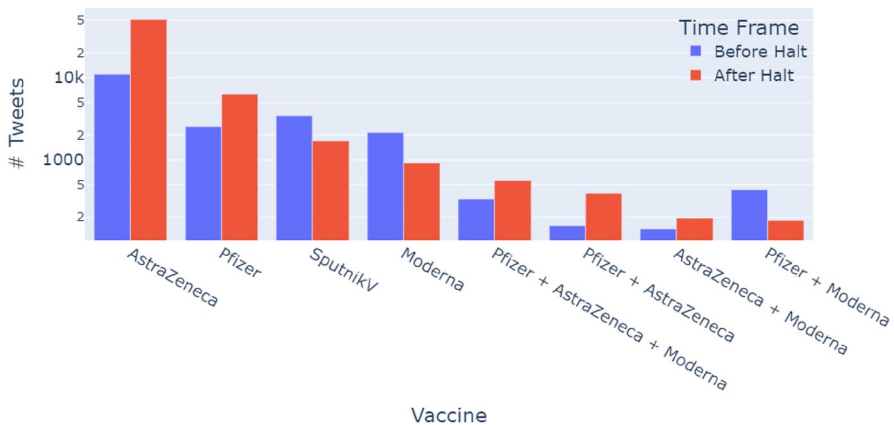
Fig. 2 Number of tweets with vaccine mentions

domain) mentioned in messages are pre-processed as follows: The shortened URLs are expanded, and HTML parameters are removed from the URLs. The YouTube URLs are resolved to the base video URLs if they include a parameter referencing a specific time in the video. We represent the URLs by the parent domain when multiple child domains exist (e.g., `fr.sputniknews.com`, `arabic.sput-niknews.com`, etc., are renamed as `sputniknews.com`). This pre-processing code of resolving URLs is publicly available.[1]

Figure 1 shows spikes in Twitter activity on September 3rd, 9th and 15th, 2020. On September 3rd, the most popular URLs point to a Twitter event on Dwayne Johnson, an American actor, and his family testing positive for COVID-19 and the URL is shared 1756 times.[2] On September 9th, the most popular URLs point to mainstream news articles, with an article published on `statnews.com` receiving the highest number (1223) of shares.[3] This article reports the halt in AstraZeneca vaccine trials in response to a potentially harmful reaction of a trial participant. We observe another spike on September 15th, where the majority of tweets (574) cite an article related to COVID re-infection.[4] There are some other tweets (354) citing an article written about a whistle-blower named Dr. Yan who released a report suggesting coronavirus was 'Lab Modified'.[5]

We consider two periods in our analysis, before (September 1–8) and after (September 9–15) the temporary interruption of the vaccine development on September 9, 2020. While the majority of the tweets mention AstraZeneca as expected (Fig. 2),

---

[1] https://github.com/pnnl/socialsim

[2] https://twitter.com/i/events/1301328910227963907.

[3] https://www.statnews.com/2020/09/08/astrazeneca-covid-19-vaccine-study-put-on-hold-due-to-suspected-adverse-reaction-in-participant-in-the-u-k/.

[4] https://dranganathans.blogspot.com/2020/09/covid-reinfection-sign-of-super.html.

[5] https://thenationalpulse.com/2020/09/14/whistleblower-dr-yan-releases-report-suggesting-coronavirus-was-lab-modified.

Table 1 Top-10 topics identified from Tweet texts (September 1–8)

Topic label	Top 10 words	# Tweets
Vaccines	Vaccin, antivax, polio, smallpox, hpv, pox, vacc, vax, chickenpox, typhoid	41,634
Covid Safety	Aye, awhil, suscept, exagger, blindli, isnt, huh, arent, diplomaci, ahem	27,520
Politics	Trump, djt, trumpster, drumpf, maga, trumpism, donald, trumper, libtard, honestli	25,641
Covid Deaths	Deaths, death, die, mortem, dead, lethal, rip, tue, morbid, fatal	22,972
Covid Related News	Amp, vox, rrb, rha, interf, wire, slash, tc, pedal, tele	22,896
Facemasks	Mask, facemask, unmask, helmet, helm, wig, visor, bandana, hide, conceal	22,837
Donald Trump	Vaccin, antivax, trump, djt, drumpf, trumpster, trumpism, barron, kushner, hpv	20,994
Child Safety	School, schoolchildren, preschool, classroom, homeschool, teacher, uncov, kindergarten, isd, teach	13,173
Covid Tests	Test, tester, retest, assay, exam, swab, trial, checkup, appt, dmv	12,914
Flu Virus	Flu, influenza, flue, sars, sicker, swine, cdc, vaccin, ill, ebola	11,325

we also detect tweets that mention Pfizer, SputnikV, and Moderna vaccines. When comparing the two periods, there is a significant increase in AstraZeneca mentioned tweets (363.7% increase) in the second period. Additionally, we observed that Pfizer mentioned tweets also increased significantly by 149.1%. We observed that SputnikV and Moderna vaccines mentioned tweets decreased by 50.2% and 57.5% respectively. There is also a significant increase (149.36%) of tweets mentioning both Pfizer and AstraZeneca vaccines.

We used a pretrained embedding model *universal-sentence-encoder*[6] provided in the Top2Vec package (Angelov 2020) to identify topics in the tweets dataset from the above mentioned two time frames. Using the default parameters, we obtained 2,527 and 2,225 topics before and after the vaccine development halt, respectively. We selected the top 10 topics from most occurring keywords from both time periods and performed a qualitative analysis (as shown in Tables 1 and 2). We noticed that face mask-related topics are prevalent in the discussions after the vaccine development halt. Twitter users have discussed about the usability and efficacy of the face-masks and the importance of wearing one (see Appendix Tables 7 and 8 for sample tweets). Other topics include COVID-19 related deaths, safety, and testing. In addition, we identified tweets discussing unemployment due to COVID-19 or due to shutdown of businesses (see Appendix Table 9 for sample tweets).

---

[6] https://tfhub.dev/google/universal-sentence-encoder/4.

Table 2 Top-10 topics identified from Tweet texts (September 9–15)

Topic label	Top 10 words	# Tweets
Facemasks	Mask, facemask, unmask, helmet, helm, safeguard, conceal, hide, visor, impun	29,824
Vaccines	Vaccin, antivax, polio, smallpox, hpv, narrow, vax, pox, cdc, bcg	25,316
Covid Safety	aye, awhil, arent, bhi, diplomaci, isnt, exagger, webpag, meh, blindli	22,205
Covid Related News	Amp, vox, rha, interf, wire, circuit, ion, tele, pedal, slash	20,691
Donald Trump	Trump, djt, trumpster, drumpf, maga, trumper, donald, trumpism, lyin, bluster	20,172
Covid Deaths	Deaths, death, mortem, die, lethal, dead, fatal, morbid, rip, tue	19,759
Vaccines	Vaccin, antivax, polio, smallpox, pox, hpv, chickenpox, vax, typhoid, cdc	18,073
Unemployment due to Covid	Work, jobless, quit, iam, job, awhil, ive, dole, newsengin, workload	14,881
Covid Safety	Aye, exagger, weve, plz, blindli, let, suscept, breweri, section, about	13,000
Politics	Biden, vp, joe, djt, trumpster, trump, obama, msnbc, sotu, mccain	12,537

4 URL analysis

We analyze the URLs cited in the tweets to identify the extent of sharing poor quality information that were originated from outside of the platform. To this end, we group the URLs into low credibility URLs (Sect. 4.1) and malicious URLs (Sect. 4.2).

4.1 Low credibility information sources

We aim to investigate how Twitter users react to low credibility information sources across the two different periods. We grouped the web domains according to the classification made by two sources (i.e., Media Bias/Fact Check (MBFC 2020) and the Factual (Factual 2020). We considered as low credibility those web domains that were in at least one of the following categories: questionable sources, conspiracy-pseudoscience, or (very)low credibility rating. We identified 14,215 (2%) URLs from 377 low credibility information sources that are shared in 42,271 (2.7%) messages posted before the halt of trials, and 12,768 URLs from 363 low credibility sources shared in 36,906 (2.7%) messages posted after the halt of trials. Table 3 shows the top 10 most popular domains by number of tweets from each time frame of interest.

thegatewaypundit.com and rt.com are the most popular low credibility web domains by number of mentions (4326 and 3626) and number of engaged users (1947 and 1971) despite publishing fewer URLs/articles than other domains in the period corresponding to before the halt. These two web sources are known for spreading propaganda and promoting conspiracy theories in their articles. After the halting of the AstraZeneca trials, the most shared low credibility web domains were rt.com and zerohedge.com. Both domains are classified with low to very low

Table 3 Twitter sharing characteristics for low credibility domains as identified by MBFC

Period	Domain	# Tweets	# Users	# URLs	Overall rank
September 1–8	thegatewaypundit.com	4326	1947	358	27
	rt.com	3626	1971	687	31
	zerohedge.com	2948	1815	237	36
	dailymail.co.uk	2548	1392	1122	45
	foxnews.com	2281	1196	622	52
	granma.cu	1536	486	192	86
	westernjournal.com	1459	876	277	93
	fr24news.com	1414	21	1407	96
	sputniknews.com	1363	602	649	99
	news18.com	1070	273	660	129
September 9–15	rt.com	2972	1478	646	36
	zerohedge.com	2930	1692	237	39
	dailymail.co.uk	2485	1411	1039	49
	foxnews.com	2273	1390	616	56
	granma.cu	1864	518	238	72
	thegatewaypundit.com	1589	902	208	83
	sputniknews.com	1509	648	698	91
	fr24news.com	1188	15	1184	108
	news18.com	879	270	541	137
	westernjournal.com	727	493	177	169

We also include the overall ranking of the low credibility domains out of all domains in the dataset

credibility due to the promotion of pseudoscience misinformation. During both periods, we also observed that `sputniknews.com` made it to the top-10 of low credibility domains shared in Twitter. Similar to other Russian state-sponsored outlets in the list, `sputniknews.com` exploited the halt of the AstraZeneca trials to push narratives discrediting the vaccine and boost the promotion of SputnikV instead.

The list of most popular low credibility web domains did not changed drastically across the two distinct time periods, but their rankings did (as shown in Fig. 3a and b).

We also noted some cases in which tweets citing the same URL often share the same article heading. These users promoted certain topics through massive repetition of messages via injecting URLs. For example, an article published in `zerohedge.com` was in the Top-10 most popular URLs on the day when the AstraZeneza vaccine development halted.[7] However, this article tried to build an alternative frame highlighting a statement by the US House Speaker Nancy Pelosi about the issue instead of reporting the details of the main event.

---

[7] https://www.zerohedge.com/markets/ft-confirms-astrazeneca-covid-19-vaccine-caused-serious-spinal-issues-test-patient

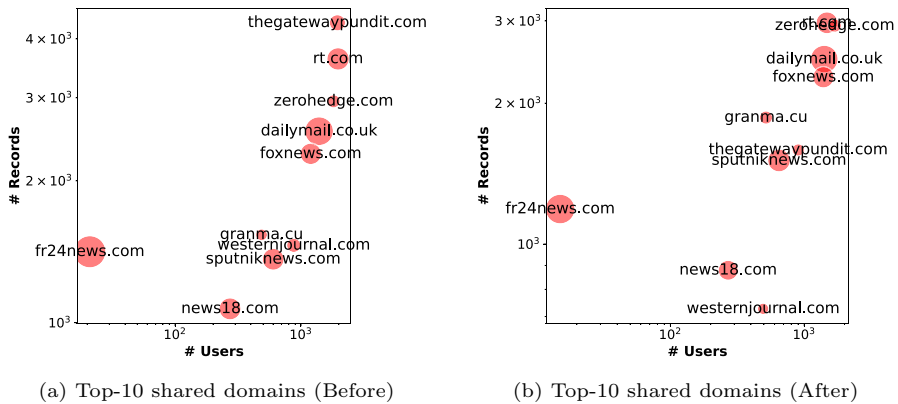(a) Top-10 shared domains (Before)   (b) Top-10 shared domains (After)

Fig. 3 Twitter sharing characteristics of most popular domains. **a** and **b** Show the Top-10 domains by the number of distinct URLs in each time frame. The size of the markers in this plot are proportional to the number of URLs associated with the domain

4.2 Malicious URLs

We used VirusTotal (VT) (VirusTotal 2021) to extract the *maliciousness* of URLs. VT provides the state-of-the-art aggregated intelligence for domains and URLs, and relies on more than 70 third-party updated antivirus (AV) engines. For all distinct URLs in our collection, we extracted VT scan reports via querying the publicly available API. Each VT scan report contains of the verdict from every AV engine, information related to the URL such as first and last seen dates of the URL in the VT system, hosting IP address, final redirected URL (if applicable), content length, etc. Each AV engine in a VT report detects if the URL is malicious or not. In order to indicate the maliciousness of a URL, we looked at how many engines flag it as such.

In this study, we labelled a URL as malicious if at least one AV engine detects it as malicious. Such malicious URLs, in general, are either phishing websites that steal user credentials and/or personally identifiable information from victims or malware hosting websites that attempt to install malware on victims' devices. Before suspending the trial, we observed that 35.9% of the malicious URLs utilize URL shortening services with top 4 services being `bit.ly`, `tinyurl.com`, `ow.ly` and `goo.su` whereas as only 20.2% of benign URLs utilize such services. We noticed rather different proportions after the trail. 25.9% of the malicious URLs utilize URL shortening services with top 4 services being `bit.ly`, `ow.ly`, `hubs.ly`, and `rb.gy` whereas only 7.2% of benign URLs utilize such services. This observation is consistent with the trend that malicious actors are increasingly using URL shortening services to camouflage malicious URLs to present non-suspicious-looking URLs to users (FAS 2020). We found that 30.80% and 40.66% of the domains related to malicious URLs are ranked below 100K by

Alexa Amazon (2021) before and after the halt respectively (the lower the rank value, the higher the popularity). This indicates the alarming reality that malicious actors are able to reach a large user base reaping a high return on investment for their attacks.

We further analyzed the malicious URLs to identify related malicious URLs. To this end, based on the lexical features in the literature Silva et al. (2021) and the hosting features mentioned in Table 4, we clustered the malicious URLs using PCA/OPTICS algorithm. While lexical features identify characteristics related to URLs themselves, hosting features, extracted from Farsight Passive DNS (PDNS) data (Farsight Security 2021), capture the characteristics of underlying hosting infrastructure. As shown in Fig. 4, these features collectively identify 4 distinct malicious URL clusters. We manually verified the accuracy of the top 2 clusters by checking the web page content, registration information and domain certificate information. The clusters observed in the two time frame are quite similar in their properties such as short-lived newly created domains, self-signed certificates, content management system (CMS) technology and hosting geographical location. We further analyzed the clusters based on the maliciousness of URLs. The maliciousness of a URL can loosely be measured by #VT, the number of VT positives. An interesting observation is that URLs belonging to different maliciousness levels share similar lexical and hosting features. We further analyzed these malicious URLs in terms of where they are hosted. Unexpectedly, we found that 80.04% of these malicious URLs are hosted in content delivery networks (CDNs) such as Cloudflare and Akamai. While CDNs provide fast delivery of content across the globe through their distributed computing infrastructure, we believe a key reason why malicious actors utilize such services is to improve attack agility and stay below the radar of malicious domain detection mechanisms in place. This observation is further reinforced with the increased utilization of public cloud computing infrastructure (33.5% of all malicious URLs) sharing hosting IPs with tens of thousands of unrelated domains, which are mostly benign. Such shared IPs are usually not blocked in practice due to the collateral damage.

5　Coordinated URL sharing behavior

Previous work on detecting coordination has focused on accounts who consistently amplify/boost sources of information (e.g., users who co-retweet the same tweets) in a social media platform (Keller et al. 2020; Pacheco et al. 2021; Weber and Neumann 2020). In this work, we are interested in those accounts who introduce new information into the platform, specifically in the form of URLs. One of our objectives is to characterize the URL sharing activities in the COVID-19 vaccine related discussions. To this end, we describe the presence of two groups of URLs (i.e., low-credibility and malicious URLs) in Twitter tweets. According to previous research (Pacheco et al. 2021; Kin Wai et al. 2021), low-credibility URLs are often promoted by coordinated groups of users in diverse contexts (e.g., U.S. elections, Hong Kong protests, and the Syrian civil war).

Table 4 Details of the URL hosting and lexical features

Feature	Description	Type
VT_Dur	URL duration in VirusTotal (VT)	Hosting
PDNS_Dur	Domain duration in Passive Domain Name Resolution (PDNS)	Hosting
#IPs	# hosting IPs	Hosting
#Queries	# times the domain is accessed	Hosting
#NSes	# Name servers	Hosting
Is_NS	Do the apexes of the domain name and NS domain name match?	Hosting
#SOAs	# administrative domains	Hosting
Is_SOA	Do the apexes of the domain name and admin domain match?	Hosting
#Domains	# domains hosted on the IP	Hosting
#Queries_IP	# times the IP is accessed	Hosting
ASN	Autonomous System Number	Hosting
Org	Organization owning the Autonomous System Numbers (ASN)	Hosting
Geo	Geographic location of the ASN of particular IP	Hosting
Server	Web server used for hosting	Hosting
Minus	The number of dashes appear in the fully qualified domain name (FQDN)	Lexical
Suspicious_TLD	Does the domain name include a suspicious country code top-level domain (ccTLD)	Lexical
Fake_TLD	Does the domain name include a fake gTLD (com, edu, net, org, gov)?	Lexical
Brand	Does it impersonate a popular Alexa top 1000 brand?	Lexical
Pop_Keywords	Does the domain name include popular keywords	Lexical
URL length	The length of the URL	Lexical
Is_IDN	Is internationalised domain name?	Lexical
NS_Domain	Is Name Server Domain?	Lexical

(a) September 1–8   (b) September 9–15
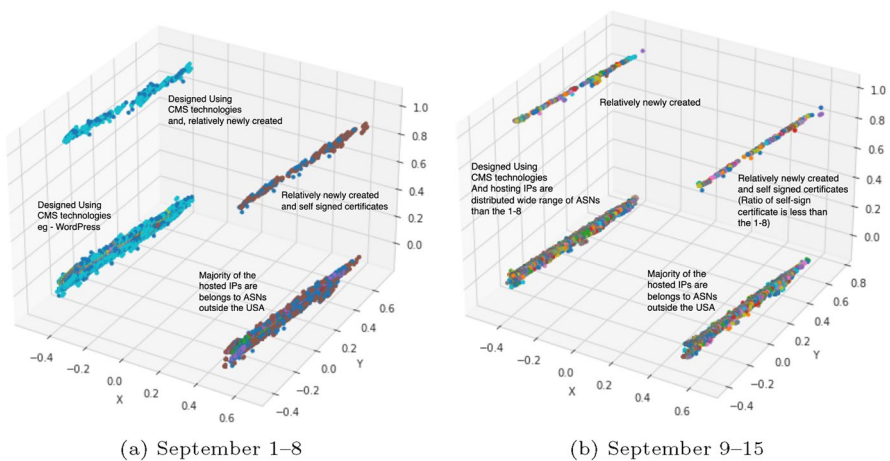
Fig. 4 Malicious URL clusters based on the lexical and hosting features. Each point is a URL, and it is colored according to the cluster it belongs

To detect coordinated link-sharing behavior, we employed the methodology proposed in Giglietto et al. (2020), where coordination is defined as "different users who repeatedly share the same URLs in an unusually short period of time." The computation of this time threshold is based on the analysis of inter-arrival times between tweets of the same URL for each period of interest. Particularly, the timing behavior of the top quickest shared URLs, as identified based on the time differences between consecutive tweets, is further explored. The desired threshold is computed by calculating the median time that it takes these top URLs to reach a certain proportion of their total shares. As can be seen, the inference of this threshold relies on three parameters: (1) the percentage of top URLs for the analysis, (2) the number of consecutive tweets to consider in the inter-arrival analysis (e.g., time difference between first and second post, or first and last post), and (3) the proportion of tweets each URL needs to reach in order to compute the median time.

We experimented with different configurations of these parameters to identify a reasonable coordination time threshold, in which simultaneous postings of the same URL would be considered unusual compared to the activity patterns of the entire period. The time threshold is mostly sensitive to both the percentage of quickest URLs and the number of consecutive tweets considered in the time difference analysis. For example, when considering high percentage of URLs (e.g., 25%, 50% or 75%) or a large number of consecutive tweets, the coordination interval would be too long and not sufficiently strict. Stringent time thresholds were found when considering smaller values for these parameters. Specifically, we considered the median time that it takes the 10% quickest URLs, ranked by their time differences between the first and second posts, to reach 50% of their shares. The threshold was 5 s for both periods, before and after the halt of trials.

We compared the coordinated networks that spanned across the two different time periods: September 1–8 and September 9–15, which correspond to Twitter activities from before and after the halt of the AZ trials, respectively. We only considered URLs that were shared at least twice and by different user accounts. From September 1 to September 8, 143,782 URLs were posted in 705,917 tweets. From September 9 to September 10, 128,426 URLs were posted in 643,667 tweets.

For each time period, we constructed the network of timely-coordinated accounts by considering only those pairs of users who post the same URL within the corresponding time interval threshold. The edge weights between users correspond to the number of URLs posted simultaneously within the threshold. In an attempt to reduce the chances of false positives (i.e., simultaneous postings happening by chance), we remove from these networks connections with an edge weight of 1. We acknowledged that additional filtering strategies could reduce even further the amount of false positives in the network. For example, inspecting the time difference between the publication date of an article and the timestamps of simultaneous tweets could reveal instances of organic behavior (e.g., crowds sharing synchronously due to a share-button functionality on websites). Unfortunately, we failed to scrape accurate publication dates for many articles due to web pages being inactive or changes to the original date caused by updates.

Table 5 shows a comparison between the coordinated networks corresponding to each time period we investigate across various network properties. We found that

Table 5 Basic network properties for the two coordinated networks from before and after the halt of trials	Statistics	Before (September 1–8)	After (September 9–15)
	Number of Nodes	2794	2528
	Number of Edges	2998	2745
	Number of URLs	8831	7420
	Number of Domains	1136	1030
	Connected Components	1039	919
	Number of Dyads	774	699
	Number of Triads	148	124

there are no drastic differences in terms of the size, density, and number of unique URLs shared between the two networks across different periods. Most connected components in both periods consist of dyads and triads (89% in the before network and 90% in the after network).

To investigate the level of trustworthiness of the news sources present in each coordinated network, we grouped the web domains according to their credibility ratings as mentioned in Sect. 4.1. We identified 385 URLs from 34 low credibility web domains, 2860 URLs from 77 credibility web domains, and 12,895 URLs from 1387 web domains with unknown credibility score. Figure 5a and b show the coordinated networks, from before and after the halt of trials, induced on only those network components that consist of at least one coordinated connection sharing credibility or low credibility domains. That is, we ignore those components in the coordinated network that exclusively shared domains with unknown credibility. The induced network for the period before the halt consists of 508 nodes and 820 edges (out of which 56 are to low credibility sources, which is 2% of the total edges in the original coordinated network). The induced network for the period after the halt consists of 444 nodes and 515 edges (out of which 58 are to low credibility sources, which also represents 2% of the total edges in the non-induced network).

Similarly, we also investigated the level of maliciousness of URLs in the coordinated networks by inducing on network components with at least one connection to a URL classified as malicious by VT. We identified 623 malicious URLs from 52 different web domains. The induced network for the period before the halt consists of 276 nodes and 632 edges (out of which 188 are to malicious sources, which is 6% of the total edges in the original coordinated network). The induced network for the period after the halt consists of 222 nodes and 273 edges (out of which 130 are to malicious sources, which represents 4.7% of the total edges in the non-induced network).

These observations suggest that the extent to which both low credibility and malicious sources are promoted by coordinated groups of users is relatively similar across the two time frames. The number of coordinated nodes that overlap between the two different periods is 314 for the induced networks by credibility and 147 for the induced networks by maliciousness, which is more than half of the users in both cases. This highlights that groups of coordinated users, who promote either
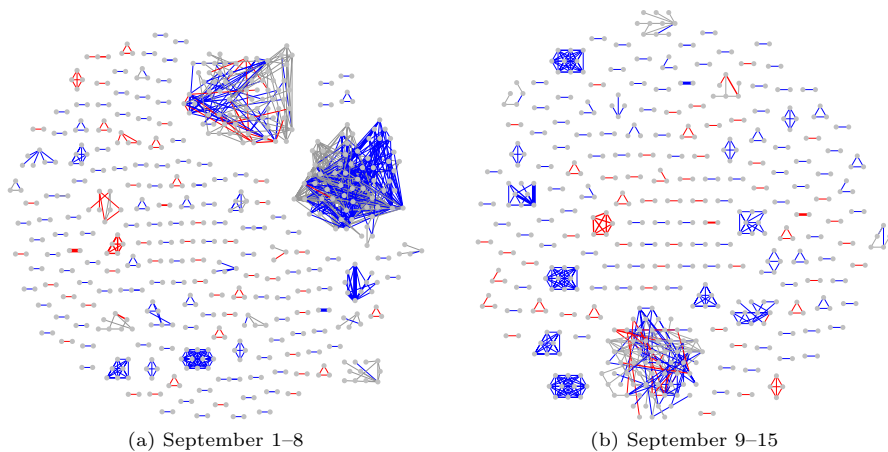
(a) September 1–8 (b) September 9–15

Fig. 5 Network of coordinated users before (left) and after (right) the halt of AstraZeneca trials. The edge weight represents the number of URLs posted simultaneously between two users. The red edges occur between users who co-shared at least one non-credibility web domain, the blue edges are between users who co-shared credibility web domains, and the gray edges are between users where the credibility of web domains is unknown. Pair of users with edge weight of 1 are removed

low credibility or malicious URLs, are persistent across the two different periods we study.

Table 6 shows the top-10 most popular low credibility sources shared by coordinated users across the two periods of interest. We found that, in both time frames, the most popular sources were `zerohedge.com`, `news18.com`, and `sputniknews.com`. The first is a website in the conspiracy-pseudoscience category, and the last two are classified as questionable sources with a poor fact-checking record. We observed a high presence of Russian web sources among the low credibility domains in the coordinated networks. Some are well-known domains such as `sputniknews.com` and `rt.com`, and others less popular such as `cnnn.ru` and `inosminews.ru`, which we found are news aggregators that often re-publish articles from the first two. We noticed that web domains shared by coordinated groups remained relatively similar in terms of their popularity across the two different periods of interest. Finally, we found that only a small number of coordinately shared URLs (26) from two web domains are classified as both low credibility and malicious: `rt.com` and `tmz.com`.

6 Conclusions

In times of crisis, whether political or health-related, online disinformation is amplified by social media promotion of alternative media outlets (Horawalavithana et al. 2020; 2021). This study adds to the growing body of work (Ferrara et al. 2020) that investigates the misinformation activity during the COVID-19 crisis by analyzing a Twitter dataset collected between September 1 and 15, 2020. This period covers

Table 6 Twitter sharing characteristics for low credibility URLs posted by users found in the coordinated networks from before and after the halt of AstraZeneca trials

Period	Domain	# Tweets	# Users	# URLs
September 1–8	zerohedge.com	152	12	47
	news18.com	89	6	24
	sputniknews.com	44	6	21
	westernjournal.com	33	3	16
	dailymail.co.uk	31	13	15
	yc.news	28	4	7
	theepochtimes.com	26	3	13
	rt.com	21	8	8
	breitbart.com	18	3	9
	foxnews.com	17	4	7
September 9–15	zerohedge.com	198	11	19
	news18.com	68	6	19
	sputniknews.com	60	6	30
	foxnews.com	37	6	14
	theepochtimes.com	34	5	16
	rt.com	28	9	9
	westernjournal.com	23	5	11
	dailymail.co.uk	22	9	11
	breitbart.com	22	6	9
	yc.news	20	4	5

events related to the AstraZeneca vaccine development phase trials. Our contributions complement previous observations (Horawalavithana et al. 2021; Singh et al. 2020) in multiple ways.

First, we found a significant increase of AstraZeneca vaccine mentioned tweets in the period following the halt of AstraZeneca vaccine trials. These messages also contain topics related to the usability and efficacy of the facemasks, vaccine safety and COVID related deaths. We also detect tweets that mention SputnikV, Moderna and Pfizer vaccines with the stories that aimed at casting doubts and fear towards the safety and efficacy of COVID-19 vaccines. This vaccination trial event was also linked to multiple other COVID-19 stories. For example, politicized discussions concerning the decisions by government officials to roll-out prematurely the vaccines were also prominent. On the other hand, topics related to unemployment due to COVID-19 or due to shutdown of businesses were popular after the vaccination trial event.

Second, we discover a strong presence of malicious and low-credibility information sources shared on Twitter messages. Not only URLs from low-credibility sources, as classified by independent fact-checking organizations, were present in the dataset, but many of them pointed to pages with malicious code. We found that a significant portion of these low-credibility and malicious URLs (36%) used URL shortening services to a greater extent than non-malicious URLs (<1%). In addition,

they were usually hosted on well-established and reputable content delivery networks in an attempt, we believe, to avoid detection.

Third, we found similar properties of the malicious URLs shared before and after the vaccine development halt event. For example, most popular malicious URLs in two time periods are quite similar in their properties such as short-lived newly created domains, self-signed certificates, content management system and hosting geographical location. Moreover, URLs belonging to different maliciousness levels share similar lexical and hosting features.

Finally, we discovered potential signs of coordination to promote malicious and low-credibility URLs on Twitter. Specifically, we identified groups of users who potentially engage in the rapid propagation and artificial inflation of a particular URL through simultaneous tweeting. Our observations suggest that the extent to which both low credibility and malicious sources are promoted by coordinated groups of users is relatively similar across the two time periods before and after the vaccine development halt event. The code and the dataset used in this work are publicly available (De Silva et al. 2022).

Our analysis is useful for multiple stakeholders ranging from individuals, educators, health professionals, journalists, researchers and governments. According to the Surgeon General of the United States, understanding the malign objectives around COVID-19 discussions would help to reduce confusion and mistrust around vaccines and promote public health efforts (Vivek 2021). In this work, we press the needs to extend the definition of COVID-19 related malign content into both low-credibility and malicious content types. This would open new directions for journalists and researchers to broaden health misinformation research done with social media datasets. In addition, they can identify high quality information sources to avoid amplifying malign content. As these malign content can change the perception of general public towards participation in future vaccine trials, we urge health professionals to better inform vulnerable communities. We also show how the malign content spreads across Twitter with potential signs of coordination. This might be helpful for individuals to identify potential coordination campaigns and avoid amplifying malign content unwittingly. Educators can use our findings to share common tactics used by bad actors and use them as evidence to improve educational programs.

Further work is needed to fully comprehend the dubious objectives of bad actors active in times of crisis. For example, bad actors might have chosen this event strategically to maximize the spread of low quality information. These actors can deploy the same strategy in future conversations, thus having content moderation techniques to limit what they can share is important. However, we can only speculate on the motivation behind the use of malware shared along with vaccine-related stories. Bad actors can use this strategy to target specific user communities with the new stories promoting certain vaccines favorably, while questioning others (FAS 2020). On the other hand, the low-credibility news sources might have reported this event opportunistically in an attempt to promote vaccine hesitancy. People might have engaged with these low quality sources to watch out the information space around COVID-19 vaccines. According to (Smith et al. 2020), there is a deficit of high quality information sources to seek vaccine information. Bad actors use this information

deficit as an advantage to push low quality information. We believe this analysis can be extended in understanding the role of bad actors during similar emotionally-charged conversations in the future. Another direction is to analyze the change of public perception around COVID-19 vaccines before and after the issues reported in the vaccine trials or the exposure to the news of side effects.

Appendix A Example tweets in Twitter discussions

See Tables 7, 8 and 9.

Table 7 An excerpt from tweets discussing face masks (September 1–8)

Tweets
Wear a mask, covid is https://t.co/fCKTDlvDNF
Think twice before wearing a face shield to protect against COVID-19 instead of a cloth face mask—here's why https://t.co/4DYERlEbd9 https://t.co/fJeOtgT7zA
@NBCNews Protect yourself From Covid 19 Use Mask & Facial Shield (see photo).Remember that on the street & in public Transportation People R wearing the mask Incorrectly. U can Get Coronavirus. It's Ur Health to Protect Yourself. https://t.co/VS3noQKhq5 We R 501(C)(3) Donate Now https://t.co/WqfpP4q8vx
@NTVNewsNL @DonBradshawNTV So if someone gets sick with COVID-19, do they not have to self quarantine now because we are all wearing NON-MEDICAL masks???
Wearing home-made reusable face-cover/mask is essential to stay protected from COVID-19.
@TODAYshow @TodayParents Please share my face mask shop so we can all look good masked up & get back to living MadMaskr https://t.co/f8rz6kRtC3 via @Etsy #facemasks #MaskMandate #TuesdayTips #fashionstyle #fashionblogger #COVID #Corona #MaskUpNOLA #MaskUpAZ #MaskUpMN #Masks #MaskMyAss #TuesdayThoughts
BarackObama Protect yourself From Covid 19 Use Mask & Facial Shield (see photo).Remember that on the street & in public Transportation People R wearing the mask Incorrectly. U can Get Coronavirus. It's Ur Health to Protect Yourself. https://t.co/VS3noQKhq5 We R 501(C)(3) https://t.co/5qWZZ46MJv
WHO #WearAMask challenge! By wearing a mask, you are sending a message of solidarity & protecting other people, especially those most vulnerable to COVID_19 Take a photo or a video of yourself wearing a mask, share it & nominate friends to do the same https://t.co/b0OunAFcpx 02
Your brain's powers of facial recognition are going to need some time to get used to the face masks we're wearing to keep each other healthy https://t.co/GAWbElS0Ec
This Face Mask Doesn't Stop COVID After All https://t.co/kre8POd3fF

Table 8 An excerpt from tweets discussing facemasks (9–15 September)

Tweets

Do you want to take of your mask, meet your friends, hug your family, go where you want, and take your chances with the covid flu. If you think living in fear wearing a mask is no life for anybody. Support an open and uncensored debate. Sign And Share x https://t.co/kjCiuNvI1E https://t.co/DHtAg9nnDI

Guys remember to wear a mask. covid is https://t.co/obnxQUGnMN

@ewsunionmp @Indersinghsjp @JPNadda @OfficeOfDrNM @OfficeofSSC @BJP4MP @INCMP @narendramodi @OfficeOfKNath @vdsharmabjp @TCGEHLOT Im staying in tamilnadu near thiruporur & i went out for groceries i saw ppl gathering without masks spitting coughing smoking holding hands i feel horible n scary about community spreading of covid shopkeeprs r nt even wearing mask or gloves y dey r behaving like illiterates??

@ElijahSchaffer Please share my face mask shop so we can all look good masked up & get back to living MadMaskr https://t.co/f8rz6kRtC3 via @Etsy #facemasks #MaskMandate #TuesdayTips #fashionstyle #fashionblogger #COVID #Corona #MaskUpNOLA #MaskUpAZ #MaskUpMN #Masks #MaskMyAss #TuesdayThoughts

Wearing a mask could protect you from COVID-19 in more ways than you think https://t.co/JNMzggXx6n

@enews Protect yourself From Covid 19 Use Mask & Facial Shield (see photo).Remember that on the street & in public Transportation People R wearing the mask Incorrectly. U can Get Coronavirus. It's Ur Health to Protect Yourself. https://t.co/VS3noQKhq5 We R 501(C)(3) Donate Now https://t.co/QKh6sN1zl8

@realDonaldTrump @POTUS Have you observed that just about all the people sitting behind Trump were wearing masks during his rally in Henderson, NV? I guess they are trying to protect him from possibly deadly Covid exhales. Notice a significant space between him and the front row. No masks.

@OfficialSidFC @sardesairajdeep Im staying in tamilnadu & i went out for groceries i saw ppl gathering without masks spitting on road coughing holding hands i feel horible n scary about community spreading of covid shopkeeprs r nt even wearing mask or gloves y dey r behaving like illiterates.r nt they unaware?

@JThakers Im staying in tamilnadu & i went out for groceries i saw ppl gathering without masks spitting on road coughing holding hands i feel horible n scary about community spreading of covid shopkeeprs r nt even wearing mask or gloves y dey r behaving like illiterates.r nt they aware?

@CNBC Test the Mask. The Covid-19 is captured in the masks. YouTube Link: https://t.co/9ju0Pc3gMF

Table 9 An excerpt from tweets discussing unemployment (9–15 September)

Tweets

@SenThomTillis Please do not leave me behind. I'm a 56 year old woman that lost my job because of COVID-19. I've worked since I was 19. This pandemic is not over! #SaveThe600 #ExtendPUA #DoYourJob #Extend600 #ExtendUI

@SenSchumer Hello Senator Schumer Im hoping you can help me.I filed for UI June 22 To this day I heard nothing.Over 100 calls and new stories every time.I spoke level 3 reps. Nothing. I worked in a public school,shut down due to COVID. I thought we would be taken care of but I was wrong HELP https://t.co/GHjba95vUv

@LoisWeiss So wait I lost my job during COVID 19 and I only make 200 dollars in UI...am I qualified for 300 dollars?

@halsey @KarriKuzma PLEASE PLEASE HELP. We're really struggling right now. My uncle just got let go from his job due to covid. I'm trying really hard to better my health so I can get a job. Please anything helps CashApp: $Jenn052192

@CashpersCraving @TriciaHuff14 @piccmeeprizes @J4CKMULL I lost my job due to Covid,I had to sell most of our stuff just to be able to eat these last few months we can't afford food..my family always comes first for me & if anyone can help us we would appreciate it more than you know. .Godbless and I'm sorry I even have to ask.ty

@pulte Please I would be sooooo thankful I've been unemployed since March from COVID and have been able to land a job bills are piling up so please $stefany12

@Trump_owo @LindseyGrahamSC @actblue I've donated A LOT and I am unemployed due to your abusive inaction re: COVID 19. And I will continue to donate. SOOOO STFU! https://t.co/g6vqc gcs8z

@tapairportugal Can you take care of your customers instead of screwing them?? Please and thank you. Covid has messed a lot up–be willing to work with your customers. Why are you refusing to extend my voucher when it was impossible to use it this summer and cannot choose dates for next year yet

@AlwayzInTrouble Hope you will check out my doc. If you like it please spread the word. Due to Covid, no work for me. So, I'm living on my marketing budget. So, word of mouth is now my marketing. https://t.co/o2AkKx8iNB

@pulte I've been laid off work because I have COVID symptoms i don't know how I'm going to pay my rent and all other bills. £abbielengthorn

Data Availability This work uses a publicly available dataset focused on the Twitterdiscussions around this event and released as part of the 2021 GrandChallenge of the North American Social Network Conference.

References

Amazon (2021) Alexa Top 1M. https://www.alexa.com/topsites. Accessed Dec 2021

Angelov D (2020) Top2vec: distributed representations of topics. CoRR abs/2008.09470. https://arxiv.org/abs/2008.09470

Bagherpour A (2020) Covid misinformation is killing people. https://www.scientificamerican.com/article/covid-misinformation-is-killing-people1/. Accessed Dec 2021

Chadwick A, Kaiser J, Vaccari C et al (2021) Online social endorsement and covid-19 vaccine hesitancy in the United Kingdom. Soc Media + Soc. https://doi.org/10.1177/20563051211008817

Cyranoski D, Mallapaty S (2020) Scientists relieved as coronavirus vaccine trial restarts but question lack of transparency. Nature 585(7825):331–332. https://doi.org/10.1038/d41586-020-02633-6

De Silva R, Weerasekara N, Kin Wai NG et al (2022) COVID-19 vaccine disinformation. https://github.com/qcri/Vaccine_Disinfo

DeVerna M, Pierri F, Truong BT et al (2021) Covaxxy: a global collection of english Twitter posts about COVID-19 vaccines. CoRR abs/2101.07694. https://arxiv.org/abs/2101.07694

Factual T (2020) The best and worst sources for unbiased news. https://www.thefactual.com/blog/biased-factual-reliable-new-sources/. Accessed Dec 2021

Farsight Security (2021) DNS database. https://www.dnsdb.info/. Accessed Dec 2021

FAS (2020) Spanish-language vaccine news stories hosting malware disseminated via URL shorteners. https://fas.org/disinfoblog/spanish-language-vaccine-news-stories-hosting-malware-disseminated-via-url-shorteners/. Accessed Dec 2021

Ferrara E, Cresci S, Luceri L (2020) Misinformation, manipulation, and abuse on social media in the era of covid-19. J Comput Soc Sci 3(2):271–277. https://doi.org/10.1007/s42001-020-00094-5

Forbes (2020) Why suspension of the AstraZeneca vaccine trial matters. https://www.forbes.com/sites/williamhaseltine/2020/09/09/why-astrazenecas-move-to-pause-their-vaccine-trial-matters/. Accessed Dec 2021

Giglietto F, Righetti N, Rossi L et al (2020) It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections. Inf Commun Soc 23(6):867–891. https://doi.org/10.1080/1369118X.2020.1739732

Horawalavithana S, Ng KW, Iamnitchi A (2020) Twitter is the megaphone of cross-platform messaging on the White Helmets. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. SBP-BRiMS 2020, Springer, pp 235–244. https://doi.org/10.1007/978-3-030-61255-9_23

Horawalavithana S, De Silva R, Nabeel M et al (2021) Malicious and low credibility URLs on Twitter during the AstraZeneca covid-19 vaccine development. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. SBP-BRiMS 2021, Springer, pp 3–12. https://doi.org/10.1007/978-3-030-80387-2_1

Huang B, Carley KM (2020) Disinformation and misinformation on Twitter during the novel coronavirus outbreak. CoRR abs/2006.04278. https://arxiv.org/abs/2006.04278

Jemielniak D, Krempovych Y (2021) An analysis of AstraZeneca covid-19 vaccine misinformation and fear mongering on Twitter. Public Health 200:4–6. https://doi.org/10.1016/j.puhe.2021.08.019

Karami A, Zhu M, Goldschmidt B et al (2021) Covid-19 vaccine and social media in the U.S.: exploring emotions and discussions on Twitter. Vaccines. https://doi.org/10.3390/vaccines9101059

Keller FB, Schoch D, Stier S et al (2020) Political astroturfing on Twitter: how to coordinate a disinformation campaign. Polit Commun 37(2):256–280. https://doi.org/10.1080/10584609.2019.1661888

Kin Wai NG, Horawalavithana S, Iamnitchi A (2021) Multi-platform information operations: Twitter, Facebook and YouTube against the White Helmets. In: Workshop proceedings of the 15th international AAAI conference on web and social media. https://doi.org/10.36190/2021.36

MBFC (2020) Media bias/fact check—search and learn the bias of news media. https://mediabiasfactcheck.com/. Accessed Dec 2021

NASN (2021) Nasn 2021: grand challenge. https://www.insna.org/nasn-2021-grand-challenge

Nuzhath T, Tasnim S, Sanjwal RK et al (2020) Covid-19 vaccination hesitancy, misinformation and conspiracy theories on social media: a content analysis of Twitter data. SocArXiv https://doi.org/10.31235/osf.io/vc9jb

Pacheco D, Hui PM, Torres-Lugo C et al (2021) Uncovering coordinated networks on social media: methods and case studies. In: Proceedings of the international AAAI conference on web and social media, vol 15, pp 455–466. https://ojs.aaai.org/index.php/ICWSM/article/view/18075

PBS (2020) FDA weighs whether to resume AstraZeneca vaccine trials after patient's serious side effect. https://www.pbs.org/newshour/health/fda-weighs-whether-to-resume-astrazeneca-vaccine-trials-after-patients-serious-side-effect. Accessed Dec 2021

Pranesh R, Farokhenajd M, Shekhar A et al (2021) CMTA: COVID-19 misinformation multilingual analysis on Twitter. In: Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing: student research workshop. Association for Computational Linguistics, pp 270–283. https://doi.org/10.18653/v1/2021.acl-srw.28

Reuters (2021) Covid-19 vaccine AstraZeneca roller-coaster ride. https://www.reuters.com/business/healthcare-pharmaceuticals/covid-19-vaccine-astrazeneca-roller-coaster-ride-2021-03-24/. Accessed Dec 2021

Robbins R, Feuerstein A, Branswell H (2020) AstraZeneca covid-19 vaccine study is put on hold. https://www.statnews.com/2020/09/08/astrazeneca-covid-19-vaccine-study-put-on-hold-due-to-suspected-adverse-reaction-in-participant-in-the-u-k/

Rosenberg H, Syed S, Rezaie S (2020) The Twitter pandemic: the critical role of Twitter in the dissemination of medical information and misinformation during the covid-19 pandemic. Can J Emerg Med 22(4):418–421. https://doi.org/10.1017/cem.2020.361

Sharma K, Zhang Y, Liu Y (2022) COVID-19 vaccines: characterizing misinformation campaigns and vaccine hesitancy on Twitter. In: Proceedings of the international AAAI conference on web and social media, vol 16, pp 920–931. https://ojs.aaai.org/index.php/ICWSM/article/view/19346

Silva M, Ceschin F, Shrestha P et al (2020) People still care about facts: Twitter users engage more with factual discourse than misinformation: a comparison between covid and general narratives on Twitter. CoRR abs/2012.02164. https://arxiv.org/abs/2012.02164

Silva RD, Nabeel M, Elvitigala C et al (2021) Compromised or attacker-owned: a large scale classification and study of hosting domains of malicious urls. In: 30th USENIX security symposium (USENIX Security 21). USENIX Association

Singh L, Bode L, Budak C et al (2020) Understanding high and low-quality url sharing on covid-19 Twitter streams. J Comput Soc Sci 3(2):343–366. https://doi.org/10.1007/s42001-020-00093-6

Smith R, Cubbon S, Wardle C (2020) Under the surface: Covid-19 vaccine narratives, misinformation and data deficits on social media. https://firstdraftnews.org/long-form-article/under-the-surface-covid-19-vaccine-narratives-misinformation-and-data-deficits-on-social-media/

Teng X, Lin YR, Chung WT et al (2022) Characterizing user susceptibility to covid-19 misinformation on Twitter. In: Proceedings of the international AAAI conference on web and social media, vol 16, pp 1005–1016

VirusTotal (2021) VirusTotal free online virus, Malware and URL scanner. https://www.virustotal.com/. Accessed Dec 2021

Vivek M (2021) Confronting health misinformation: the U.S. surgeon general's advisory on building a healthy information environment. https://www.ncbi.nlm.nih.gov/books/NBK572169/. Accessed Dec 2021

Weber D, Neumann F (2020) Who's in the gang? Revealing coordinating communities in social media. In: 2020 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), IEEE, pp 89–93

Wilson SL, Wiysonge C (2020) Social media and vaccine hesitancy. BMJ Glob Health. https://doi.org/10.1136/bmjgh-2020-004206

Yang KC, Torres-Lugo C, Menczer F (2020) Prevalence of low-credibility information on twitter during the COVID-19 outbreak. In: Proc. ICWSM Intl. Workshop on Cyber Social Threats (CySoc). https://doi.org/10.36190/2020.16

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Sameera Horawalavithana[1] [iD] **· Ravindu De Silva[2] · Nipuna Weerasekara[2] · N G Kin Wai[1] · Mohamed Nabeel[3] · Buddhini Abayaratna[2] · Charitha Elvitigala[2] · Primal Wijesekera[4] · Adriana Iamnitchi[5]**

Ravindu De Silva
ravindud@scorelab.org

Nipuna Weerasekara
w.nipuna@scorelab.org

N G Kin Wai
kinwaing@usf.edu

Mohamed Nabeel
mnabeel@hbku.edu.qa

Buddhini Abayaratna
buddhinia@scorelab.org

Charitha Elvitigala
charitha@scorelab.org

Primal Wijesekera
primal@berkeley.edu

[1]   CSE, University of South Florida, Tampa, USA

[2]   SCoRe Lab, Colombo, Sri Lanka

[3]   Qatar Computing Research Institute, Ar Rayyān, Qatar

[4]   ICSI, University of California, Berkeley, Berkeley, USA

[5]   Institute of Data Science, Maastricht University, Maastricht, Netherlands