

Bad robot? The benevolent use of automated software and social bots by influencers in the #antivaxx discourse on Twitter

Antonia Egli and Theo Lynn

DCU Business School, Dublin City University, Dublin, Ireland

Pierangelo Rosati

J.E. Cairnes School of Business and Economics, University of Galway, Galway, Ireland, and

Gary Sinclair

DCU Business School, Dublin City University, Dublin, Ireland

Abstract

Purpose – Automated social media messaging tactics can undermine trust in health institutions and public health advice. As such, we examine automated software programs (ASPs) and social bots in the Twitter anti-vaccine discourse before and after the release of COVID-19 vaccines.

Design/methodology/approach – We compare two Twitter datasets comprising user accounts and associated English-language tweets featuring the keywords “#antivaxx” or “anti-vaxx.” The first dataset, from 2018 (pre-COVID vaccine), includes 3,154 user accounts and 6,380 tweets. The second comprises 327,067 accounts and 545,268 tweets published during the 12 months following December 1, 2020 (post-COVID vaccine). Using Information Laundering Theory (ILT), the datasets were examined manually and through user analytics and machine learning to identify activity, visibility, verification status, vaccine position, and ASP or bot technology use.

Findings – The post-COVID vaccine dataset showed an increase in highly probable bot accounts (31.09%) and anti-vaccine accounts. However, both datasets were dominated by pro-vaccine accounts; most highly active (59%) and highly visible (50%) accounts classified as probable bots were pro-vaccine.

Originality/value – This research is the first to compare bot behaviors in the “#antivaxx” discourse before and after the release of COVID-19 vaccines. The prevalence of mostly benevolent probable bot accounts suggests a potential overstatement of the threat posed by anti-vaccine accounts using ASPs or bot technologies. By highlighting bots as intermediaries that disseminate both pro- and anti-vaccine content, we extend ILT by identifying a benevolent variant and offering insights into bots as “pathways” to generating mainstream information.

Keywords Twitter, Social bots, COVID-19, Social media, Anti-Vaccine, Information laundering theory

Paper type Research paper

Introduction

Global vaccination rates have declined since the COVID-19 pandemic despite substantial evidence that vaccines are safe and contribute to disease prevention and control (WHO and

© Antonia Egli, Theo Lynn, Pierangelo Rosati and Gary Sinclair. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Research funding: This study was partially funded by the Irish Institute of Digital Business in Dublin, Ireland. There is no other funding to report.

Data Availability: The datasets generated and analyzed for this study were done so using a paid Twitter license and therefore cannot be published.

Competing Interests: The authors declare no competing interests.



UNICEF, 2022). Indeed, the World Health Organization (WHO) has reported the largest regression in childhood vaccinations in 30 years (WHO and UNICEF, 2022). Over the same 30-year period, the Internet and social media have become powerful alternative sources of health information to the public (Lynn *et al.*, 2020). Online social platforms facilitate the implementation of public health interventions and provide valuable support in promoting health advice, advancing rapid responses to health crises, and tracking disease outbreaks (Moorhead *et al.*, 2013). During the COVID-19 pandemic, social media was used widely to educate and influence public opinion (Kothari *et al.*, 2022). However, public health monitoring and advocacy on social media are simultaneously endangered by a lack of accuracy in and regulation of shared health information (Sinapuelas and Ho, 2019). Manipulative and deceptive communication practices on social media, propagated, for example, by automated software programs (ASPs) and social bots, can interfere with public health communication by creating a false sense of uniformity and validity (Lynn *et al.*, 2020). Combined with inadequate regulatory structures surrounding health-related content, social bots, which aim to influence discussions online, and other automated messaging tactics stand to endanger public consensus by obscuring sound health advice and legitimizing unverified, false, or misleading information (Ferrara *et al.*, 2016; Broniatowski *et al.*, 2018).

To date, only few studies (e.g. Yuan *et al.*, 2019) have addressed the role of ASPs and social bots in the anti-vaccine (*anti-vaxx*) discourse and none have compared the prevalence of specific bot-like behaviors before and after the release of COVID-19 vaccines. Studies that examine social bots in the overall vaccine discourse focus on the content published rather than the actors themselves (Suarez-Lledo and Alvarez-Galvez, 2022; Duan *et al.*, 2022; Zhang *et al.*, 2023). Furthermore, little attention has been paid to the social media user accounts contributing to the #antivaxx discourse to discern whether these are truly anti-vaxx or, instead, represent pro-vaccine (*pro-vaxx*) actors countering anti-vaxx positions or attempting to sway the vaccine hesitant (for a pre-COVID-19 analysis, see Dunn *et al.*, 2020). This is deserving of deeper investigation, as a better understanding of probable bot accounts and the bot-like behaviors they exhibit informs both public health and platform policy and supports the development of countermeasures to information originating from potentially malicious accounts.

Against this backdrop and expanding on existing research, this study investigates the extent to which user accounts involved in the #antivaxx discourse on Twitter presented potentially manipulative communication behaviors before and after the release of COVID-19 vaccines. This study builds on existing findings by (1) analyzing a global sample of English-language tweets and (2) comparing two distinct periods that include a time in which the topic of vaccines became notably more prominent. It also applies a novel theoretical lens to the study of social bots within the discourse by means of Information Laundering Theory (Klein, 2012), which describes how information is made available and gains credibility in mainstream online spaces by progressively passing through digital intermediaries.

We examine increases in social bot volume across two distinct periods, detect specific bot-like behavior within the discourse, and determine the vaccine positions of the datasets' most active and most influential actors. Analyses suggest that probable bot accounts participating in the Twitter #antivaxx discourse were largely benevolent and that numerous pro-vaxx accounts exhibited bot-like behavior.

Literature review

The impact of social media misinformation on COVID-19 vaccine hesitancy

COVID-19 vaccines present an opportunity to control a highly transmissible disease that, in addition to unprecedented social and economic crises, resulted in approximately 700 million infections and over six million deaths worldwide (WHO, 2023). However, false, misleading, and inaccurate information on social media has led to increased vaccine hesitancy, thereby reducing the effectiveness of COVID-19 vaccination programs. *Vaccine hesitancy*, otherwise

described as delayed vaccination despite the availability of and access to vaccines, significantly contributes to avoidable deaths and the adverse social, financial, and political effects of diseases (WHO, 2019). Although the dangers of vaccine-related mis- and disinformation on social media were observed long before the spread of COVID-19 (Steffens *et al.*, 2019), research has since shown a substantial uptake in both vaccine-related content and false claims surrounding the COVID-19 vaccine (Prasad, 2022). In contrast, social media platforms have also been identified as spaces that may cater to “disadvantaged social groups” (Laor, 2022), a descriptor that, within a health context, also applies to the vaccine hesitant (Mendonça and Hilário, 2023). Vaccine hesitant individuals have been found to build both in-person and online networks to manage stigma surrounding their positions towards vaccines and exchange information regarding vaccines and vaccination (Mendonça and Hilário, 2023). Research nonetheless finds a correlation between media exposure and heightened risk perceptions of COVID-19 and highlight the negative effects of consuming polarizing content or mis- and disinformation relating to COVID-19 vaccines on users’ intent to vaccinate (Loomba *et al.*, 2021; Mendonça and Hilário, 2023; Sitar-Taut and Mican, 2024). As such, exposure to *social bots* can also influence opinions towards vaccines (Zhang *et al.*, 2022).

Social bots and automated software programs

Social bots are part of a wider category of digital tools defined as *automated software programs* (ASPs). ASPs are not limited to social environments and include any type of agent that automates repetitive tasks, like web crawlers (Grimme *et al.*, 2017). Nonetheless, they may be used to publish and manage content on different digital channels (Lynn *et al.*, 2020). Examples of ASPs include automated posts from activity trackers that are connected to user Twitter accounts, such as updates generated by Fitbits, or enterprise marketing technologies like Hootsuite (Lynn *et al.*, 2020). Increasingly, ASPs are also used in combination with machine learning to support marketing decision-making processes (Järvinen and Taiminen, 2016). In general, ASPs are designed for instrumental or communicative purposes to amplify reach or efficiently communicate with target users (Lynn *et al.*, 2020).

Social bots, however, are distinctly programmed to mimic human behavior and fulfill functions ranging from simple replications of online human communication patterns to autonomous social actions (Assenmacher *et al.*, 2020; Lynn *et al.*, 2020). The objectives of social bots are typically to amplify content through likes or shares, build networks by connecting with other user accounts, or generate original content (Assenmacher *et al.*, 2020). “Social” in this sense refers to their interactions with humans and their presence on social networking sites (Ferrara *et al.*, 2016). Research finds that social bots retweet more frequently but generate fewer replies, retweets, and mentions (“@”) from humans and are more disconnected from their networks than human users (Ferrara *et al.*, 2016; Alhayan *et al.*, 2023). Social bots may be used individually or in so-called *social botnets* or *spreading groups* consisting of large groups of bots functioning under a single coordinator, or *botmaster*, who manages interactions to generate tweets independently or within a *retweet chain* (Cook *et al.*, 2013; Stieglitz *et al.*, 2017; Sela *et al.*, 2020). Typically, these networks spread specific messages by creating information cascades through ping-pong-like exchanges (Sela *et al.*, 2020).

Social bots within the social media health discourse

Within a health context, social bots can both spread verified health content or negatively impact valid health-related messaging by disseminating false or misleading information (Subrahmanian *et al.*, 2016; Zhang *et al.*, 2023). Evidence of the malicious use of bots to spread rumors, spam, disinformation, slander, or noise is substantial (Ferrara *et al.*, 2016). This poses a threat to human social media users, particularly those who are uncertain in their decisions to vaccinate. Extant research suggests that the anti-vaxx movement has employed social bots to negatively influence vaccine-related narratives and decisions to vaccinate among

hesitant individuals (Subrahmanian *et al.*, 2016; Zhang *et al.*, 2022). Examining an online network of approximately 100 million user accounts expressing positions about vaccination, Johnson *et al.* (2020) found that anti-vaxx clusters, while smaller in size and primarily consisting of ideologically fringe groups, were most often intertwined with larger, thematically mainstream vaccine hesitant communities. Repeated exposure to such content has further been shown to result in increased hesitancy towards vaccines, as well as a preference for finding information online rather than from accredited healthcare organizations (Jones *et al.*, 2012; Jolley and Douglas, 2014).

Benevolent social bots have also been observed within the health and vaccine discourse on Twitter. For example, during the initial phases of the vaccine roll-out, benign pro-vaxx social bots emphasized the safety and convenience of vaccination, cited authoritative sources, and clarified rumors surrounding the volume of available vaccines (Zhang *et al.*, 2023). Even before COVID-19, Yuan *et al.* (2019) found that although only a small percentage of user accounts (1.45%) discussing their decision to vaccinate against measles, mumps, and rubella on Twitter were bots, these were almost equally active in disseminating pro-vaxx and anti-vaxx messages. Similarly, Ruiz-Núñez *et al.* (2022) detected bot usage levels ranging from 14.31% in pro-vaxx networks to 16.19% in anti-vaxx networks. Abrahams and Aljizawi (2020) in their analysis of COVID-19-related social media manipulation by Twitter bots in the Middle East found the level of plausible bot accounts comparable to pre-COVID-19 levels, suggesting that fears of widespread bot use on Twitter may, at least in the Middle East, have been overstated (Abrahams and Aljizawi, 2020). Nonetheless, the use of bots in spreading health-related information warrants further analysis due to their overall role as intermediaries within the information dissemination process.

Theoretical framework: the Model of Information Laundering

Information Laundering Theory (ILT; Klein, 2012) provides a structure to position social bots and ASPs within the #antivaxx discourse on Twitter. ILT posits that mis- or disinformation is legitimized and established as mainstream public knowledge through repeated exposure in online environments, or *cyberspace* (Klein, 2012). Like money laundering, information is “cleaned” and gains credibility by being disseminated across the “interconnected information superhighway of web directories, research engines, news outlets, and social networks” (Klein, 2012, p. 433). Klein’s (2012) Model of Information Laundering incorporates both ILT and the theoretical concept of *techno-ethos*, which ascribes credibility or authority to an online source based on technological signals and its ability to mimic the “aesthetic Internet standard” (Borrowman, 1999, p. 238). Within the context of websites, for example, Borrowman (1999) names visual appeal, technical sophistication, and user traffic as instances of *techno-ethos*.

To date, literature has explored laundering networks (Klinger *et al.*, 2023), the content normalized through laundering processes (Stolze, 2022), and the “digital social objects” (i.e. URLs, hashtags, photos, mentions) native to social media platforms that were employed to launder information (Puschmann *et al.*, 2016). The overwhelming majority of studies does not account for the presence of social bots and only Klinger *et al.* (2023) examine automation levels within the German-language far-right news and information sites, albeit descriptively. Klein’s (2012) model and existing literature furthermore assume that propagated information is malicious *ab initio* and dismiss the role of benevolent actors within the information laundering process. This infers malicious intent, with actors intentionally propagating harmful content to “sanitize” information, and ignores the benevolent actors who may aim to amplify or counter information without spreading mis- or disinformation. Lastly, ILT has not been applied to the context of vaccines, leaving a gap in understanding how both vaccine-related misinformation and credible content are disseminated and legitimized by means of information laundering.

Although Klein (2012) focuses exclusively on hate speech, we propose that social bots and ASPs can be viewed as general information laundering intermediaries within the #antivaxx

discourse. Specifically, these actors function as integral entities within the structure of social networks to include, generate, amplify, or distort vaccine-related content. This allows us to better approach their role within the discourse and examine the complex relationship between the technology, the vaccine-related information that is disseminated using the technology, and the types of users that employ the technology. Applying ILT also allows for a closer examination of highly active and highly influential information intermediaries and how these contribute to the spread and perceived legitimacy of vaccine-related positions. To gain a better understanding of concrete communication dynamics within the laundering process, this research further asks which bot-like behaviors are employed as mechanisms to legitimize vaccine-related information. Lastly, examining these changes before and after the introduction of the COVID-19 vaccine generates insights into the adaptation of communication tactics of social bots over time.

Research questions

This study leverages two Twitter datasets that comprise English-language tweets posted across two time periods. The level of detail in these datasets and their longitudinal nature provide us with a unique opportunity to compare the characteristics, communication behaviors, and vaccine positions of user accounts involved in the #antivaxx discourse on Twitter during these periods. To contribute to the discussed literature surrounding the role of social media on health information communication in a broader sense, this study seeks to answer the following research questions.

- (1) To what extent do highly active and highly influential user accounts use ASPs and social bots in the Twitter #antivaxx discourse before and after the introduction of COVID-19 vaccines?
- (2) Which social bot behaviors are exhibited by highly active and highly influential user accounts in the Twitter #antivaxx discourse before and after the introduction of COVID-19 vaccines?

Data and methods

This study leverages and compares two datasets comprising Twitter accounts that published English-language tweets featuring the hashtag “#antivaxx” or variations of the keyword “anti-vaxx.” The first dataset (hereafter “pre-COVID vaccine dataset”) comprises 6,380 tweets generated by 3,154 user accounts between January and December 2018. The second dataset (hereafter “post-COVID vaccine dataset”) comprises 545,258 tweets generated by 327,067 user accounts between December 1, 2020 (when COVID-19 vaccines were first approved for public use in the United States) and November 30, 2021. [Table 1](#) provides a high-level overview of the two datasets. All data was obtained using Twitter’s enterprise Application Programming Interface (API). 2018 was chosen as a focal period to compare a year during which the vaccine discourse was relatively standard with a highly topical, high-volume online conversation such as that brought about by the COVID-19 pandemic. While 2018 saw various virus outbreaks, these were geographically contained and comparably small. 2019 was excluded due to several major multi-country virus outbreaks, including measles ([CDC, 2021](#)) and dengue fever ([Hosangadi, 2019](#)).

This study uses Twitter as its empirical context because it is a largely open network that facilitates the spread of content between connected user accounts and strangers, largely through hashtags (“#”; [Lynn et al., 2020](#)). Twitter is particularly relevant as it has been widely used for health surveillance and related research ([Zhang et al., 2022, 2023](#); [Suarez-Lledo and Alvarez-Galvez, 2022](#); [Duan et al., 2022](#)) and is a popular channel among pro-vaxx, anti-vaxx, and vaccine hesitant individuals alike ([Broniatowski et al., 2018](#); [Ruiz et al., 2021](#)).

Table 1. Descriptive tweet and user account analysis

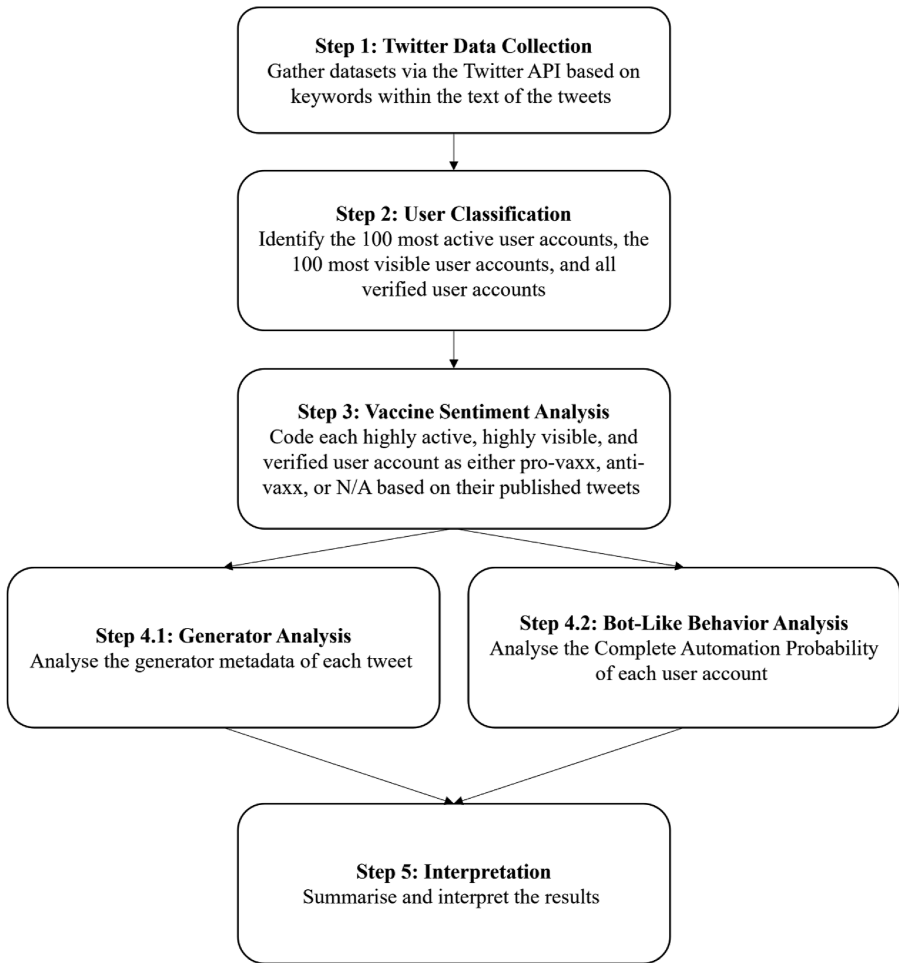
	Pre-COVID vaccine dataset	Post-COVID vaccine dataset
<i>Panel 1: Tweet Characteristics</i>		
No. of Tweets	6,380	545,258
No. of Original Tweets ^a	2,454	114,669
No. of Retweets	372	366,298
No. of Replies	3,554	64,291
<i>Panel 2: User Account Characteristics</i>		
No. of Distinct Users	3,154	327,067
No. of Verified Users	57	3,674
Avg. Tweets per User	2.02	1.67
Avg. User Activity	2.02	1.49
Avg. User Visibility	1.02	1.14
Note(s): Tweets are neither replies (“@”) nor retweets		
Source(s): Authors’ own work		

This study adopted a multi-step methodology that is graphically summarized in [Figure 1](#). Once the datasets were generated (Step 1), all user accounts were classified based on their activity, visibility, and verification status (Step 2). User activity is measured as the sum of tweets, retweets, and replies *posted* by a user account. Active user accounts publish the largest amount of Twitter content and primarily *drive* discussions on a given topic due to the volume of tweets they publish ([Park and Macy, 2015](#)). Visibility is measured as the number of retweets and replies *received* by a user account and is an indicator of an account’s ability to reach and *attract* other accounts ([Chae, 2014](#)). Highly visible user accounts are therefore highly influential user accounts. Lastly, verified user accounts are those that were authenticated or determined to be of public interest by Twitter. At the time of data collection, verified user accounts were of public interest and deemed as “active, notable, and authentic” by Twitter ([X, 2023](#)). They are therefore also viewed as highly influential [1]. More specifically, we identified the 100 most active (hereafter “highly active user accounts”), the 100 most visible (hereafter “highly visible user accounts”) and verified user accounts (57 in the pre-COVID vaccine dataset and 3,674 in the post-COVID vaccine dataset) due to their roles as opinion leaders on the platform. This initial classification allowed us to conduct a series of analyses at various levels of granularity (i.e. the full dataset, highly active, highly visible, and verified user accounts) and compare the results.

Step 3 of the methodology involved manual coding and focused on highly active, highly visible, and verified user accounts only. Due to the substantial number of user accounts, particularly in the post-COVID vaccine dataset, the elbow method was used to reduce sample size ([Al-Rawi et al., 2019](#)) and identify the data’s most active and most influential user accounts. The elbow method produces the optimal number of clusters within a given dataset based on an “elbow,” or a stark drop-off within a graph. While the drop-offs varied, examining the top 100 user accounts per category presented the most feasible solution across all datasets. Based on this method, we examined all tweets posted across the three cohorts to determine their vaccine positions, i.e. whether they were pro-vaxx, anti-vaxx, or neutral.

Step 4 consisted of two separate yet interconnected analyses. The first (Step 4.1) leveraged the *generator metadata* available in our datasets to reveal the type of software used to post each tweet. More specifically, we extracted the full list of generators and classified them across the following five categories:

- (1) Twitter clients: the official applications released by Twitter (Twitter Web Client, Twitter for iPhone/Android, etc.).



Source(s): Authors' own work

Figure 1. Research methodology

- (2) Other social networks: social networking sites that allow user accounts to cross-post to Twitter.
- (3) Third-party clients: applications built on the Twitter platform by external developers that are not owned or operated by Twitter. Marketing professionals often use these to automate social media posting and manage user engagement and social media marketing campaigns (Hootsuite, TweetDeck, Buffer, etc.).
- (4) Social bots: generators whose name contains “bot,” therefore suggesting that they are designed to automatically post content or engage with content posted by other user accounts (Lynn *et al.*, 2020).
- (5) Other: all generators that could not be included in any of the prior categories (e.g. generators that were no longer in use, untraceable generators, or mobile applications).

The second analysis (Step 4.2) focused on identifying user accounts' bot-like behavior and the specific technique(s) they may have implemented. More specifically, each discrete account across the two datasets was analyzed using the Botometer (Davis *et al.*, 2016). The Botometer is a supervised machine learning classifier designed to detect social bots on Twitter. It leverages approximately 1,000 features from a Twitter user account and its activity to evaluate similarity with known features of social bots. In doing so, it reports a social bot detection accuracy over 95%, as well as the *Complete Automation Probability* (CAP), or the probability that a user account presenting said CAP or greater is a social bot (Davis *et al.*, 2016; Ferrara *et al.*, 2016). Extant research has deemed Botometer CAP results as useful indicators of potentially deceptive accounts that require more in-depth analysis (Shao *et al.*, 2018). The Botometer also estimates a score ranging from 0 to 5 to measure the extent to which each user account engaged in the following bot-like behaviors.

- (1) *Astroturfing*: evidence of mimicking grassroots campaigns.
- (2) *Fake followers*: the degree to which a user account has purchased bot accounts to increase its followership.
- (3) *Financial*: evidence of posting cashtags (“\$”), an indexing feature that allows Twitter user accounts to browse stock symbols and related content.
- (4) *Self-declared*: verified online bots registered on the documentation site <https://botwiki.org>.
- (5) *Spammer*: user accounts previously identified as spambots.
- (6) *Other*: bots manually identified through user feedback.

The results of these analyses were aggregated in Step 5 to provide a final interpretation of the findings.

Results

Descriptive analytics

Analyzing the activity, visibility, and verification status of user accounts within the #antivaxx discourse is key to creating a foundational sense of the highly influential actors that propagate information by means of bot technology and ASPs. These actors are crucial elements of the information laundering process. High user account activity, visibility, and a verified status are indicators of techno-ethos, as they deliver signals of credibility to other user accounts. In describing the characteristics of highly influential actors, we can examine who amplified the dissemination and furthered the legitimization of vaccine-related information within the discourse, and how.

A comparative analysis across the two datasets (see Table 1) shows that the average level of activity in the post-COVID vaccine dataset (1.49 published posts per account) is lower than in the pre-COVID vaccine dataset (2.02 published posts per account), although both values suggest a generally low average level of activity. Visibility across both datasets is comparable, with 1.02 mentions/retweets/replies received on average by each user account in the pre-COVID vaccine dataset and 1.14 in the post-COVID vaccine dataset. These findings suggest that the #antivaxx discourse on Twitter comprised a disperse network of only a few highly connected accounts in both datasets.

As outlined in Step 3, highly active, highly visible, and verified user accounts were examined for their vaccine positions. This analysis (Table 2) shows that anti-vaxx positions decreased solely among highly active user accounts – namely from 41.18% (70 user accounts) to 33.77% (51 user accounts) – in the post-COVID vaccine dataset. Nonetheless, most highly active and highly visible user accounts remained pro-vaxx, making up 66.23% (100 accounts) and 50% (100 accounts) of the post-COVID sample respectively.

Table 2. Vaccine position analysis

User group	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
	Pro-vaxx	Anti-vaxx	Pro-vaxx	Anti-vaxx
Highly Active Users	100	70	100	51
Highly Visible Users	100	16	100	100
Verified Users	55	1	2,999	65

Source(s): Authors' own work

Among verified user accounts, 96.49% (55 accounts) were coded as pro-vaxx in the pre-COVID vaccine dataset. In the post-COVID vaccine dataset, 79.51% (2,999 accounts) of verified user accounts were identified as pro-vaxx, 1.72% (65 accounts) as anti-vaxx, and 18.02% (680 accounts) as neutral. Overall, highly active anti-vaxx user accounts decreased from 41.18% to 33.77%, while pro-vaxx verified user accounts decreased from 96.49% to 79.51% across the two datasets.

Generator analysis

The use of bot-like generators or ASPs contributes to the information laundering process by introducing information to the discourse and significantly amplifying it through automated or semi-automated tactics. Importantly, this applies to both legitimate and false vaccine-related information. Comparing both datasets shows that the overall volume of tweets and the number of discrete generators used to publish them increased significantly post-COVID vaccine. In fact, only 25 unique generators could be detected in the pre-COVID vaccine dataset, while 649 generators featured in the post-COVID vaccine dataset.

Panel 1 in [Table 3](#) shows that official Twitter clients were the most used generators in terms of the number of tweets and associated user accounts across both datasets. In contrast, only 1.25% of user accounts in the pre-COVID vaccine dataset and 0.46% of user accounts in the post-COVID vaccine dataset make use of generators that are classified as social bots. A closer examination, however, found that 3 (12%) of the 25 generators used within the pre-COVID vaccine dataset and 211 (32.51%) of all 649 generators within the post-COVID vaccine dataset self-identified as bots. This indicates a rise in the overall number of bot generators created to publish content within the #antivaxx discourse in the post-COVID vaccine dataset, but with minor impact on the overall conversation as they were used by only a fraction of user accounts. Through the lens of ILT, this suggests a deliberate strategy to automate the dissemination and amplification of vaccine-related information and offers first insights into the infrastructure of the actors within the discourse.

The top 100 most active and most visible pro- and anti-vaxx user accounts from both datasets were analyzed to provide a more granular view of generator software. Panels 2 and 3 in [Table 3](#) indicate that bot generators became more prevalent among both user groups over the two focal periods, rising from 0% to 1% among highly active pro-vaxx user accounts and from 1% to 2% among highly visible pro-vaxx user accounts. 0% of highly active anti-vaxx user accounts made use of bot generators in the pre-COVID vaccine dataset, while 3.29% of these accounts used bot generators in the post-COVID vaccine dataset. Highly visible anti-vaxx user accounts using bot generators rose from 0% to 1%. Nonetheless, official Twitter clients remained the most popular generators on Twitter throughout.

Lastly, verified user accounts were examined in [Table 3](#), Panel 4. The percentage of accounts using bot generators increased from 0% (0 accounts) to 1.07% (32 accounts) in the post-COVID vaccine dataset, as did the percentage of accounts using official Twitter generators (from 90.9% [50 accounts] to 92.43% [2,832 accounts]). Closer examination of verified user accounts employing bot generators presented solely pro-vaxx positions (32 user accounts representing 1.07% of accounts in the post-COVID vaccine dataset). This indicates

Table 3. Generator analysis

Panel 1: All Users	Pre-COVID Vaccine Dataset			Post-COVID Vaccine Dataset		
	No. of Generators	No. of Tweets	No. of Accounts	No. of Generators	No. of Tweets	No. of Accounts
Twitter Clients	5 (20%)	371 (87.5%)	522 (93.55%)	37 (5.70%)	536,222 (98.34%)	335,525 (98.78%)
3 rd Party Clients	11 (44%)	36 (8.49%)	21 (3.76%)	225 (34.67%)	4,666 (0.86%)	1,996 (0.59%)
Other Social Network	2 (8%)	5 (1.18%)	4 (0.72%)	4 (0.62%)	423 (0.08%)	306 (0.09%)
Social Bots	3 (12%)	5 (1.18%)	7 (1.25%)	211 (32.51%)	3,300 (0.61%)	1,575 (0.46%)
Other	4 (16%)	7 (1.65%)	4 (0.72%)	172 (26.50%)	64 (0.12%)	261 (0.08%)
Panel 2: Highly Active Users	Pre-COVID Vaccine Dataset		Post-COVID Vaccine Dataset			
	Pro (n = 100)	Anti (n = 70)	Pro (n = 100)	Anti (n = 51)		
Twitter Clients	91 (91%)	66 (94.29%)	90 (90%)	48 (94.12%)		
3 rd Party Clients	7 (7%)	2 (2.85%)	7 (7%)	1 (1.96%)		
Other Social Network	1 (1%)	2 (2.85%)	1 (1%)	0 (0%)		
Social Bots	0 (0%)	0 (0%)	1 (1%)	2 (3.29%)		
Other	1 (1%)	0 (0%)	1 (1%)	0 (0%)		
Panel 3: Highly Visible Users	Pre-COVID Vaccine Dataset		Post-COVID Vaccine Dataset			
	Pro (n = 100)	Anti (n = 16)	Pro (n = 100)	Anti (n = 100)		
Twitter Clients	93 (93%)	14 (87.5%)	92 (92%)	95 (95%)		
3 rd Party Clients	5 (5%)	2 (12.5%)	6 (6%)	4 (4%)		
Other Social Network	0 (0%)	0 (0%)	0 (0%)	0 (0%)		
Social Bots	1 (1%)	0 (0%)	2 (2%)	1 (1%)		
Other	1 (1%)	0 (0%)	0 (0%)	0 (0%)		
Panel 4: Verified Users	Pre-COVID Vaccine Dataset		Post-COVID Vaccine Dataset			
	Pro (n = 55)	Anti (n = 1)	Pro (n = 2,999)	Anti (n = 65)		
Twitter Clients	50 (90.9%)	0 (0%)	2,769 (92.3%)	63 (96.92%)		
3 rd Party Clients	5 (9.09%)	1 (100%)	192 (6.4%)	2 (3.08%)		
Other Social Network	0 (0%)	0 (0%)	1 (0.03%)	0 (0%)		
Social Bots	0 (0%)	0 (0%)	32 (1.07%)	0 (0%)		
Other	0 (0%)	0 (0%)	5 (0.17%)	0 (0%)		

Source(s): Authors' own work

that although verified user accounts remained predominantly pro-vaxx, a growing percentage of these accounts functioned as information intermediaries by promoting their positions using bot generators.

Bot-like behavior analysis

In the first step of the study's final analysis, all accounts for which the Botometer returned an "error" message were removed. These were not available for analysis due to account deactivation, deletion, or suspension. As a result, 2,632 user accounts were included in the following analyses of the pre-COVID vaccine dataset and 286,428 user accounts in the post-COVID vaccine dataset.

The average CAP scores across the datasets are insightful. User accounts in the pre-COVID vaccine dataset presented an average CAP of 51%, while the post-COVID vaccine dataset average increased slightly to 52%. Notwithstanding this relatively minor increase, developments in the prevalence of accounts presenting *high* levels of bot-like behavior – a

CAP higher than 70% – were significant. Extant research uses CAP scores ranging from 50% (Zhang *et al.*, 2022) to 80% (Suarez-Lledo and Alvarez-Galvez, 2022) to determine probable bot user accounts. For this study, we consider social bots as those accounts with a CAP score of 70% or higher in line with Duan *et al.* (2022). Table 4, Panel 1 shows that 29.45% of user accounts active in the pre-COVID vaccine dataset presented a CAP of 70% or higher and that this figure increased to 31.09% in the post-COVID vaccine dataset.

The average CAP of all highly active and highly visible pro- and anti-vaxx accounts increased over time; in particular, the average CAP of all highly active pro-vaxx user accounts rose from 56% to 72% across the two datasets (Table 4, Panel 2). Looking at the number of user accounts with a CAP of 70% or higher also shows more highly active and highly visible pro-vaxx user accounts in the post-COVID vaccine dataset compared to the pre-COVID vaccine dataset (Table 4, Panel 3). These results indicate that more highly active and highly visible pro-vaxx user accounts exhibited bot-like behavior, with highly active pro-vaxx accounts rising from 34% to 59% and highly visible pro-vaxx accounts increasing from 45% to 50%.

In contrast, the average CAP presented by pro-vaxx verified user accounts decreased from 57% in the pre-COVID vaccine dataset to 49% in the post-COVID vaccine dataset (Table 4, Panel 2). Anti-vaxx verified users decreased from 80% in the pre-COVID vaccine dataset to 43% in the post-COVID vaccine dataset. A closer examination of verified user accounts classified as social bots showed only a relatively small percentage of anti-vaxx positions in both datasets. In fact, only one account (1.75% of verified accounts) in the pre-COVID vaccine dataset and 9 accounts (13.84% of verified accounts) in the post-COVID vaccine dataset were labeled as anti-vaxx.

Examining bot-like behaviors uncovers the specific mechanisms used to launder information within the discourse. These behaviors indicate how bots not only spread, but also enhance the legitimacy of both factual and false information through signals of techno-ethos as it moves from fringe discourses into the mainstream. In terms of specific bot-like behaviors, Panel 1 in Table 5 shows that user accounts in the post-COVID vaccine dataset presented, on average, more *fake followers* and *financial bot* techniques compared to the pre-COVID vaccine dataset, while other techniques such as *astroturfing*, *self-declared*, and *spammer* are less visible. Table 5, however, suggests that *astroturfing* and *fake followers* became more prominent in the post-COVID vaccine dataset among highly active user accounts (Table 5, Panel 2) and highly visible user accounts (Table 5, Panel 3). All bot-like behaviors (apart from *astroturfing*) became less prevalent among verified user accounts in the post-COVID vaccine dataset. This aligns with a

Table 4. Distribution of users with a CAP of 70% or higher

	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
	No. of tweets	No. of accounts	No. of tweets	No. of accounts
<i>Panel 1: All users</i>				
CAP ≥70%	1,694 (26.55%)	775 (29.45%)	189,789 (34.81%)	89,037 (31.09%)
	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
	Pro	Anti	Pro	Anti
<i>Panel 2: Average CAP User Breakdown</i>				
Highly Active Users	56%	64%	72%	65%
Highly Visible Users	63%	70%	65%	71%
Verified Users	57%	80%	49%	43%
<i>Panel 3: CAP ≥70% User Breakdown</i>				
Highly Active Users	34 (34%)	25 (35.7%)	59 (59%)	20 (39.22%)
Highly Visible Users	45 (45%)	3 (18.75%)	50 (50%)	41 (41%)
Verified Users	22 (40%)	1 (100%)	7 (0.23%)	9 (13.65%)
Source(s): Authors' own work				

Table 5. Bot-like behavior overview and user breakdown

Panel 1: All users	Pre-COVID vaccine dataset Average score		Post-COVID vaccine dataset Average score	
Astroturfing	1.35		1.22	
Fake Follower	0.69		0.72	
Financial	0.23		0.29	
Self-Declared	1.32		0.17	
Spammer	0.16		0.14	
Other	1.32		1.22	

Panel 2: Highly active users	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
	Pro (n = 100)	Anti (n = 70)	Pro (n = 100)	Anti (n = 51)
Astroturfing	1.26	0.74	2.59	1.33
Fake Follower	0.49	0.81	0.93	0.93
Financial	0.13	0.16	0.14	0.21
Self-Declared	0.07	0.47	0.22	0.47
Spammer	0.08	0.33	0.17	0.25
Other	1.18	1.23	1.90	1.97
Suspended/Deleted Accounts	0	0	10	19

Panel 3: Highly visible users	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
	Pro (n = 100)	Anti (n = 16)	Pro (n = 100)	Anti (n = 100)
Astroturfing	1.3	0.95	1.74	1.69
Fake Follower	0.53	0.58	0.57	0.8
Financial	0.09	0.1	0.02	0.03
Self-Declared	0.13	0.22	0.09	0.25
Spammer	0.09	0.07	0.02	0.11
Other	1.29	1.15	1.19	1.47
Suspended/Deleted Accounts	1	5	5	37

Panel 4: Verified users	Pre-COVID vaccine dataset		Post-COVID vaccine dataset	
Astroturfing	1.29		1.35	
Fake Follower	0.61		0.46	
Financial	0.17		0.04	
Self-Declared	0.19		0.13	
Spammer	0.16		0.03	
Other	1.49		1.26	
Suspended/Deleted Accounts	0		98	

Source(s): Authors' own work

final observation that the number of suspended or deleted user accounts increased among all highly active and highly influential user accounts in the post-COVID vaccine dataset.

Discussion

Use of automated software programs and social bots by highly active and highly influential user accounts in the Twitter #antivaxx discourse

This research delivers new insights into the degree to which bot-like communication behaviors and ASPs were adopted within the #antivaxx discourse on Twitter and by whom. They show that (1) the prevalence of user accounts classified as social bots increased to 31.09% in the

post-COVID vaccine dataset and (2) the most active and most visible accounts exhibiting bot-like communication behaviors were primarily pro-vaxx in nature (59 and 50% respectively). While, for example, Ruiz-Núñez *et al.* (2022) show comparable results in the volume of probable bot accounts within pro- and anti-vaxx networks on Twitter, this study also examines bot-like behaviors and content generators. It demonstrates that bot-like behavior and the use of bot generators increased dramatically between the pre- and post-COVID vaccine periods, but that these were primarily used as communication techniques in pro-vaxx narratives and comprised only a relatively small fraction of the overall discourse.

Social bot behaviors of highly active and highly influential user accounts in the Twitter #antivaxx discourse

This is the first granular analysis of the specific bot-like behavior exhibited within the #antivaxx discourse before and after the introduction of COVID-19 vaccines. It shows that *astroturfing*, *fake followers*, and *financial* bot behaviors were more prevalent in the post-COVID vaccine dataset particularly among highly active and highly visible user accounts. Specifically, adopting *astroturfing* techniques and purchasing *fake followers* may have served to amplify the credibility of pro-vaccine user accounts by means of heightened techno-ethos, i.e. creating an impression of authenticity or increasing follower numbers to ensure prioritization by Twitter's algorithm and enhance visibility in a wider and potentially more active followership. This is a key step in the information laundering process and demonstrates that laundering tactics are employed not only by anti-vaxx cohorts, but also by actors who have legitimate interest in improving the accuracy and truthfulness of content within the vaccine discourse. *Financial* bot techniques, like using cashtags, may have otherwise supported the spread of information surrounding investment in pharmaceutical companies or other health and vaccine-related organizations. It may also have served to reach targeted audiences active on Twitter.

Scholarly contribution

This study of probable bot-like behaviors and generators exhibited by highly active and highly influential user accounts within the #antivaxx discourse on Twitter addresses gaps in extant literature by expanding on previously examined time periods and detailing specific bot-like behaviors. It also supports methodological contributions to the field of health information communication research. This study's manually coded highly active, highly visible, and verified user accounts, annotated for vaccine position, and CAP scores can feed into future research surrounding the manual and computational classification and analysis of vaccine narratives on social media and their effects on user engagement and influence (Peng and Wang, 2024).

This study contributes to existing literature by solidifying behavioral insights into bot use on social media and underlining that perceptions of malignant bot use within the online vaccine discourse may have been overstated. Analyses show that probable bot accounts within the #antivaxx discourse on Twitter were primarily benevolent and that largely pro-vaxx accounts exhibited bot-like behaviors. These findings echo research conducted before the emergence of the COVID-19 pandemic by Dunn *et al.* (2020) demonstrating that most of the vaccine-related content encountered by typical user accounts in the United States was generated by humans rather than bots. In fact, the median percentage of exposures originating from bots was 0.0% (interquartile range = 0.0–0.5%), indicating that bots had minimal to no influence on the Twitter vaccine discourse in the United States between 2017 and 2019 (Dunn *et al.*, 2020). This study expands on these results by (1) examining a corpus of English-language tweets not geographically restricted to the United States and (2) comparing two distinct periods during which the topic of vaccines became significantly more salient.

In describing the Model of Information Laundering, Klein (2012) states that, "in cyberspace, the pathways to false knowledge and propaganda are the same as those that lead to legitimate and credible resources" (p. 435). A similar thought is echoed in discussions surrounding the ethics of bot use. Due to their abilities to mislead public perception and manufacture consensus, sway public

sentiment, and spread false health claims, social bots have been reported as threats within the online vaccine discourse (Woolley and Guilbeault, 2017). However, *bot civility* calls to view social bots not exclusively as “unethical apparatuses” (Coleman, 2018, p. 120), but rather as a new method of joining the online public discourse similar to other technologies human social media users generally find permissible (Haeg, 2017). This research provides an initial theoretical input and a first step in understanding the *intent polarity* of social bots within the context of information laundering. By highlighting bots as intermediaries that facilitate the flow of both false and credible information, this work offers early insights into viewing bots as “pathways” that contribute to generating mainstream information. Within this context, we posit that social bots and ASPs can facilitate information laundering by amplifying messages at an accelerated rate – whether pro-social or harmful – as they integrate with other platforms, like social networks. For example, while malicious actors may use bots to amplify falsehoods, ethical actors work to combat such narratives by amplifying truthful and credible information. Contrary to Klein’s (2012) theory, this assumes that *benevolent information laundering* is just as feasible as the malicious use of automated communication tactics. This perspective presents a foundation for theory development surrounding the online information landscape in that it explores the intersection between bot use, information dissemination, and information laundering within contested discourses.

Societal impact

Countering vaccine related mis- and disinformation online is a significant multi-stakeholder challenge that can be addressed by interventions from individual user accounts, public health agencies, policymakers, pharmaceutical companies, and social media platforms (Luo *et al.*, 2024). The work conducted by these stakeholders is at risk if deceptive practices hinder public health messages from reaching social media’s most vulnerable human user accounts. Understanding the use of ASPs and social bots in the Twitter #antivaxx discourse is crucial in identifying the sources and spreaders of mis- and disinformation, which can significantly negatively impact public health (Steffens *et al.*, 2019; Sitar-Taut and Micán, 2024). The rising use of social bots and ASPs post-COVID vaccine highlights how such activities might increase as discourses become more topical. The predominance of pro-vaxx bot-like user accounts within this study provides important insight into potentially effective measures in counteracting anti-vaxx mis- and disinformation, promoting vaccine uptake, and creating awareness surrounding the benefits of vaccination. Here, amplification by means of automated communication tactics serves as a reactive strategy to combat misinformation. However, although most examined bots were pro-vaxx, the risk remains that even a small percentage of anti-vaxx bots can amplify misinformation, contributing to vaccine hesitancy and undermining public health efforts (Loomba *et al.*, 2021). The prevalence of bot-like behavior among influential accounts, even for pro-social causes like vaccine advocacy, equally raises ethical concerns surrounding content manipulation and the authenticity of user accounts active in the discourse (Haeg, 2017). Specifically, the use of bots to artificially inflate support (*astroturfing*) or follower counts (*fake followers*) distorts public opinion and misleads users, potentially polarizing debates and hindering human users’ ability to discern credible information. This can negatively affect public understanding of critical health issues like vaccination. In addition to delivering an improved understanding of ASP strategies, observations of benevolent bot-like activity and bot generator use may (1) guide policy decisions and regulation efforts on social media platforms, (2) support the development of pro-social communication and counter-misinformation strategies, and (3) inform discussions surrounding social media platform accountability in social bot detection and account moderation.

Limitations and future research

This study is not without limitations, which in themselves represent opportunities for future research. Firstly, this research is limited to one social network, Twitter, and the English language. There is urgent need for further research focusing on other languages, other countries, and other

platforms with varying content moderation policies including other social networks, Website forums, and encrypted messaging apps such as WhatsApp and Telegram, where anti-vaxx discourses are widespread. Secondly, this study does not measure the impact of bot usage exhibited within the pre-COVID vaccine and post-COVID vaccine discourses. Measuring word of mouth (WOM), or user opinions, experiences, and emotions surrounding products or services shared within their social networks (Chetioui *et al.*, 2023), by focusing both on CAP percentages and available metrics provided by Twitter (such as likes, replies, or retweets) will enable a better understanding of the impact of bot-like behaviors and bot generators. A final limitation is the dearth of tools available for social bot detection and analysis, many of which were limited to Twitter (now known as X). These tools are dependent on API access and the data made available by social network platforms. Changes in the breadth and availability of data, along with changing terms of access to such APIs, may be detrimental to the development and evolution of bot detection and analysis software, as well as future research and replication studies critical for scientific validation. For example, Twitter has introduced a fee for accessing data for academic purposes that is prohibitive to most researchers. Even when available, software such as Botometer could not analyze suspended or deleted user accounts and therefore ignored a sample of accounts that may have been active during both focal periods but were since removed from the platform (Bovet and Makse, 2019). Lastly, such bot detection and analysis software are trained on posts that may no longer truly reflect new generations of ASPs or behaviors exhibited by genuine human users during the COVID-19 pandemic. For example, it may be that levels of retweeting by human users increased significantly during pandemic lockdowns.

This study finds that the most active and visible user accounts within the #antivaxx discourse on Twitter were, in fact, pro-vaccine and made use of ASPs and social bots to promote their vaccine position. However, further research is required to better understand how COVID-19 impacted the publication of vaccine-related content. This creates ample opportunity for future research avenues focusing on (1) user migration to less restrictive social media channels as refuges for human users who were banned from other platforms, (2) changes in the general prevalence of suspended and deleted user accounts during these focal periods, (3) comparative examinations of bot-like behavior across social media platforms, and (4) theorizing in regard to ILT. From a theoretical perspective, research avenues should include broadening the scope of ILT to explore ethical information laundering and further investigate the intent polarity of bots in both propagating misinformation and disseminating legitimate health content within the vaccine discourse. Future research should also attempt to better understand the behavior of automated actors, the structure and impact of their networks, and the validity of information shared before, during, and after the COVID-19 pandemic within the #antivaxx discourse. Such efforts overall will deliver supporting findings towards mitigating the adverse effects of mis- and disinformation and restoring faith in vaccines and vaccination programs.

Notes

1. Twitter's verification program was modified towards a payment-based verification model in November 2022.

References

- Abrahams, A. and Aljizawi, N. (2020), "Middle East Twitter bots and the Covid-19 infodemic", *Workshop Proceedings of the 14th International AAI Conference on Web and Social Media*.
- Al-Rawi, A., Groshek, J. and Zhang, L. (2019), "What the fake? Assessing the extent of networked political spamming and bots in the propagation of #fakenews on Twitter", *Online Information Review*, Vol. 43 No. 1, pp. 53-71, doi: [10.1108/oir-02-2018-0065](https://doi.org/10.1108/oir-02-2018-0065).
- Alhayan, F., Pennington, D. and Ayouni, S. (2023), "Twitter use by the dementia community during COVID-19: a user classification and social network analysis", *Online Information Review*, Vol. 47 No. 1, pp. 41-58, doi: [10.1108/oir-04-2021-0208](https://doi.org/10.1108/oir-04-2021-0208).

- Assenmacher, D., Clever, L., Frischlich, L., Quandt, T., Trautmann, H. and Grimme, C. (2020), "Demystifying social bots: on the intelligence of automated social media actors", *Social Media+Society*, Vol. 6 No. 3, 2056305120939264, doi: [10.1177/2056305120939264](https://doi.org/10.1177/2056305120939264).
- Borrowman, S. (1999), "Critical surfing: holocaust deniability and credibility on the web", *College Teaching*, Vol. 47 No. 2, pp. 44-54, doi: [10.1080/87567559909595782](https://doi.org/10.1080/87567559909595782).
- Bovet, A. and Makse, H.A. (2019), "Influence of fake news in Twitter during the 2016 US presidential election", *Nature Communications*, Vol. 10 No. 1, p. 7, doi: [10.1038/s41467-018-07761-2](https://doi.org/10.1038/s41467-018-07761-2).
- Broniatowski, D.A., Jamison, A.M., Qu, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S.C. and Dredze, M. (2018), "Weaponized health communication: twitter bots and Russian trolls amplify the vaccine debate", *American Journal of Public Health*, Vol. 108 No. 10, pp. 1378-1380.
- CDC (2021), "Measles cases and outbreaks", available at: <https://www.cdc.gov/measles/cases-outbreaks.html> (accessed December 2023).
- Chae, B.K. (2014), "A complexity theory approach to IT-enabled services (IESs) and service innovation: Business analytics as an illustration of IES", *Decision Support Systems*, Vol. 57, pp. 1-10, doi: [10.1016/j.dss.2013.07.005](https://doi.org/10.1016/j.dss.2013.07.005).
- Chetioui, Y., Butt, I., Fathani, A. and Lebdaoui, H. (2023), "Organic food and instagram health and wellbeing influencers: an emerging country's perspective with gender as a moderator", *British Food Journal*, Vol. 125 No. 4, pp. 1181-1205.
- Coleman, M.C. (2018), "Bots, social capital, and the need for civility", *Journal of Media Ethics*, Vol. 33 No. 3, pp. 120-132, doi: [10.1080/23736992.2018.1476149](https://doi.org/10.1080/23736992.2018.1476149).
- Cook, D.M., Waugh, B., Abdipannah, M., Hashemi, O. and Rahman, S.A. (2013), "Twitter deception and influence: issues of identity, slacktivism, and puppetry", *Journal of Information Warfare*, Vol. 13 No. 1, pp. 58-71.
- Davis, C.A., Varol, O., Ferrara, O.E., Flammini, A. and Menczer, F. (2016), "Botornot: a system to evaluate social bots", *Proceedings of the 25th International Conference Companion on World Wide Web*, pp. 273-274, doi: [10.1145/2872518.2889302](https://doi.org/10.1145/2872518.2889302).
- Duan, Z., Li, J., Lukito, J., Yang, K.C., Chen, F., Shah, D.V. and Yang, S. (2022), "Algorithmic agents in the hybrid media system: social bots, selective amplification, and partisan news about COVID-19", *Human Communication Research*, Vol. 48 No. 3, pp. 1-27, doi: [10.1093/hcr/hqac012](https://doi.org/10.1093/hcr/hqac012).
- Dunn, A.G., Surian, D., Dalmazzo, J., Rezazadegan, D., Steffens, M., Dyda, A., Leask, J., Coiera, E., Dey, A. and Mandl, K.D. (2020), "Limited role of bots in spreading vaccine-critical information among active twitter users in the United States: 2017-2019", *American Journal of Public Health*, Vol. 110 No. S3, pp. S319-S325, doi: [10.2105/ajph.2020.305902](https://doi.org/10.2105/ajph.2020.305902).
- Ferrara, E., Varol, O., Davis, C., Menczer, F. and Flammini, A. (2016), "The rise of social bots", *Communications of the ACM*, Vol. 59 No. 7, pp. 96-104, doi: [10.1145/2818717](https://doi.org/10.1145/2818717).
- Grimme, C., Preuss, M., Adam, L. and Trautmann, H. (2017), "Social bots: human-like by means of human control?", *Big Data*, Vol. 5 No. 4, pp. 279-293, doi: [10.1089/big.2017.0044](https://doi.org/10.1089/big.2017.0044).
- Haeg, J. (2017), "The ethics of political bots: should we allow them for personal use?", *Journal of Practical Ethics*, Vol. 5 No. 2, pp. 85-104.
- Hosangadi, D. (2019), "The global rise of dengue infections", available at: <https://www.outbreakobservatory.org/outbreakthursday-1/3/21/2019/theglobal-rise-of-dengue-infections> (accessed December 2023).
- Järvinen, J. and Taiminen, H. (2016), "Harnessing marketing automation for B2B content marketing", *Industrial Marketing Management*, Vol. 54, pp. 164-175, doi: [10.1016/j.indmarman.2015.07.002](https://doi.org/10.1016/j.indmarman.2015.07.002).
- Johnson, N.F., Velasquez, N., Restrepo, N.J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S. and Lupu, Y. (2020), "The online competition between pro-and anti-vaccination views", *Nature*, Vol. 582 No. 7811, pp. 230-233, doi: [10.1038/s41586-020-2281-1](https://doi.org/10.1038/s41586-020-2281-1).
- Jolley, D. and Douglas, K.M. (2014), "The effects of anti-vaccine conspiracy theories on vaccination intentions", *PLoS One*, Vol. 9 No. 2, e89177, doi: [10.1371/journal.pone.0089177](https://doi.org/10.1371/journal.pone.0089177).

- Jones, A.M., Omer, S.B., Bednarczyk, R.A., Halsey, N.A., Moulton, L.H. and Salmon, D.A. (2012), "Parents' source of vaccine information and impact on vaccine attitudes, beliefs, and nonmedical exemptions", *Advances in Preventive Medicine*, Vol. 2012, pp. 1-8, doi: [10.1155/2012/932741](https://doi.org/10.1155/2012/932741).
- Klein, A. (2012), "Slipping racism into the mainstream: a theory of information laundering", *Communication Theory*, Vol. 22 No. 4, pp. 427-448, doi: [10.1111/j.1468-2885.2012.01415.x](https://doi.org/10.1111/j.1468-2885.2012.01415.x).
- Klinger, U., Lance Bennett, W., Knüpfer, C.B., Martini, F. and Zhang, X. (2023), "From the fringes to mainstream politics: intermediary networks and movement-party coordination of a global anti-immigration campaign in Germany. Information", *Communications Society*, Vol. 26 No. 9, pp. 1890-1907, doi: [10.1080/1369118x.2022.2050415](https://doi.org/10.1080/1369118x.2022.2050415).
- Kothari, A., Walker, K. and Burns, K. (2022), "#CoronaVirus and public health: the role of social media in sharing health information", *Online Information Review*, Vol. 46 No. 7, pp. 1293-1312, doi: [10.1108/oir-03-2021-0143](https://doi.org/10.1108/oir-03-2021-0143).
- Laor, T. (2022), "My social network: group differences in frequency of use, active use, and interactive use on Facebook, Instagram and Twitter", *Technology in Society*, Vol. 68, 101922, doi: [10.1016/j.techsoc.2022.101922](https://doi.org/10.1016/j.techsoc.2022.101922).
- Loomba, S., de Figueiredo, A., Piatek, S.J., de Graaf, K. and Larson, H.J. (2021), "Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA", *Nature Human Behaviour*, Vol. 5 No. 3, pp. 337-348, doi: [10.1038/s41562-021-01056-1](https://doi.org/10.1038/s41562-021-01056-1).
- Luo, C., Zhu, Y. and Chen, A. (2024), "What motivates people to counter misinformation on social media? Unpacking the roles of perceived consequences, third-person perception and social media use", *Online Information Review*, Vol. 48 No. 1, pp. 105-122, doi: [10.1108/oir-09-2022-0507](https://doi.org/10.1108/oir-09-2022-0507).
- Lynn, T., Rosati, P., Santos, G.L. and Endo, P.T. (2020), "Sorting the healthy diet signal from the social media expert noise: preliminary evidence from the healthy diet discourse on twitter", *International Journal of Environmental Research and Public Health*, Vol. 17 No. 22, p. 8557, doi: [10.3390/ijerph17228557](https://doi.org/10.3390/ijerph17228557).
- Mendonça, J. and Hilário, A.P. (2023), "Touching the cornerstone: an illustrative example of the effects of stigma and discrimination on vaccine-hesitant parents", *Public Health in Practice*, Vol. 6, 100438, doi: [10.1016/j.puhip.2023.100438](https://doi.org/10.1016/j.puhip.2023.100438).
- Moorhead, S.A., Hazlett, D.E., Harrison, L., Carroll, J.K., Irwin, A. and Hoving, C. (2013), "A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication", *Journal of Medical Internet Research*, Vol. 15 No. 4, e85, doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933).
- Park, P. and Macy, M. (2015), "The paradox of active users", *Big Data and Society*, Vol. 2 No. 2, 2053951715606164, doi: [10.1177/2053951715606164](https://doi.org/10.1177/2053951715606164).
- Peng, R.X. and Wang, R.Y. (2024), "The infinity vaccine war: linguistic regularities and audience engagement of vaccine debate on Twitter", *Online Information Review*, Vol. 48 No. 1, pp. 84-104, doi: [10.1108/oir-03-2022-0186](https://doi.org/10.1108/oir-03-2022-0186).
- Prasad, A. (2022), "Anti-science misinformation and conspiracies: COVID-19, post-truth, and science and technology studies (STS)", *Science Technology and Society*, Vol. 27 No. 1, pp. 88-112, doi: [10.1177/09717218211003413](https://doi.org/10.1177/09717218211003413).
- Puschmann, C., Ausserhofer, J., Maan, N. and Hametner, M. (2016), "Information laundering and counter-publics: the news sources of islamophobic groups on Twitter", *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 10, pp. 143-150, doi: [10.1609/icwsm.v10i2.14847.2](https://doi.org/10.1609/icwsm.v10i2.14847.2).
- Ruiz, J., Featherstone, J.D. and Barnett, G.A. (2021), "Identifying vaccine hesitant communities on twitter and their geolocations: a network approach", *Proceedings of the 54th Hawaii International Conference on System Sciences*, pp. 3964-3969.
- Ruiz-Núñez, C., Segado-Fernández, S., Jiménez-Gómez, B., Hidalgo, P.J.J., Magdalena, C.S.R., Pollo, M.D.C.Á. and Herrera-Peco, I. (2022), "Bots' activity on COVID-19 Pro and anti-vaccination networks: analysis of Spanish-written messages on Twitter", *Vaccines*, Vol. 10 No. 8, pp. 1-10, doi: [10.3390/vaccines10081240](https://doi.org/10.3390/vaccines10081240).

- Sela, A., Milo, O., Kagan, E. and Ben-Gal, I. (2020), "Improving information spread by spreading groups", *Online Information Review*, Vol. 44 No. 1, pp. 24-42, doi: [10.1108/oir-08-2018-0245](https://doi.org/10.1108/oir-08-2018-0245).
- Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A. and Menczer, F. (2018), "The spread of low-credibility content by social bots", *Nature Communications*, Vol. 9 No. 1, pp. 1-9, doi: [10.1038/s41467-018-06930-7](https://doi.org/10.1038/s41467-018-06930-7).
- Sinapuelas, I.C. and Ho, F.N. (2019), "Information exchange in social networks for health care", *Journal of Consumer Marketing*, Vol. 36 No. 5, pp. 692-702, doi: [10.1108/jcm-12-2017-2470](https://doi.org/10.1108/jcm-12-2017-2470).
- Sitar-Taut, D.A. and Mican, D. (2024), "Social media exposure assessment: influence on attitudes toward generic vaccination during the COVID-19 pandemic", *Online Information Review*, Vol. 47 No. 1, pp. 138-161, doi: [10.1108/oir-11-2021-0621](https://doi.org/10.1108/oir-11-2021-0621).
- Steffens, M.S., Dunn, A.G., Wiley, K.E. and Leask, J. (2019), "How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation", *BMC Public Health*, Vol. 19 No. 1, pp. 1-2, doi: [10.1186/s12889-019-7659-3](https://doi.org/10.1186/s12889-019-7659-3).
- Stieglitz, S., Brachten, F., Ross, B. and Jung, A.K. (2017), "Do social bots dream of electric sheep? A categorization of 147 social media bot accounts", *Proceedings of Australasian Conference on Information Systems*, pp. 1-11.
- Stolze, M. (2022), *Information Laundering via Baltnews on Telegram: How Russian State-Sponsored Media Evade Sanctions and Narrate the War*, NATO Strategic Communications Centre of Excellence, Riga.
- Suarez-Lledo, V. and Alvarez-Galvez, J. (2022), "Assessing the role of social bots during the COVID-19 pandemic: infodemic, disagreement, and criticism", *Journal of Medical Internet Research*, Vol. 24 No. 8, e36085, doi: [10.2196/36085](https://doi.org/10.2196/36085).
- Subrahmanian, V.S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A. and Menczer, F. (2016), "The DARPA Twitter bot challenge", *Computer*, Vol. 49 No. 6, pp. 38-46, doi: [10.1109/mc.2016.183](https://doi.org/10.1109/mc.2016.183).
- WHO (2019), "Ten threats to global health in 2019", available at: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019> (accessed December 2023).
- WHO (2023), "WHO coronavirus (COVID-19) dashboard", available at: <https://covid19.who.int/?mapFilter=deaths> (accessed December 2023).
- WHO and UNICEF (2022), "COVID-19 pandemic fuels largest continued backslide in vaccinations in three decades", available at: <https://www.who.int/news/item/15-07-2022-covid-19-pandemic-fuels-largest-continued-backslide-in-vaccinations-in-three-decades> (accessed December 2023).
- Woolley, S. and Guilbeault, D. (2017), *Computational Propaganda in the United States of America: Manufacturing Consensus Online*, Computational Propaganda Worldwide, in Samuel, W. and Howard, P.N. (Eds), Working Paper 2017.5, Oxford, p. 27, available at: <https://demtech.oi.ox.ac.uk/>
- X (2023), "How to get the blue checkmark on X", available at: <https://help.twitter.com/en/managing-your-account/about-x-verified-accounts> (accessed January 2024).
- Yuan, X.Y., Schuchard, R.J. and Crooks, A.T. (2019), "Examining emergent communities and social bots within the polarized online vaccination debate in Twitter", *Social Media+ Society*, Vol. 5 No. 3, 2056305119865465, doi: [10.1177/2056305119865465](https://doi.org/10.1177/2056305119865465).
- Zhang, M., Chen, Z., Qi, X. and Liu, J. (2022), "Could social bots' sentiment engagement shape humans' sentiment on COVID-19 vaccine discussion on twitter?", *Sustainability*, Vol. 14 No. 9, p. 5566, doi: [10.3390/su14095566](https://doi.org/10.3390/su14095566).
- Zhang, Y., Song, W., Shao, J., Abbas, M., Zhang, J., Koura, Y.H. and Su, Y. (2023), "Social bots' role in the COVID-19 pandemic discussion on twitter", *International Journal of Environmental Research and Public Health*, Vol. 20 No. 4, p. 3284, doi: [10.3390/ijerph20043284](https://doi.org/10.3390/ijerph20043284).

Corresponding author

Antonia Egli can be contacted at: antonia.egli@dcu.ie

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com