The COVID-19 Infodemic: Twitter versus Facebook

Kai-Cheng Yang[1] ⓘ, Francesco Pierri[1,2] ⓘ, Pik-Mai Hui[1],
David Axelrod[1] ⓘ, Christopher Torres-Lugo[1], John Bryden[1] ⓘ
and Filippo Menczer[1] ⓘ

Abstract

The global spread of the novel coronavirus is affected by the spread of related misinformation—the so-called COVID-19 Infodemic—that makes populations more vulnerable to the disease through resistance to mitigation efforts. Here, we analyze the prevalence and diffusion of links to low-credibility content about the pandemic across two major social media platforms, Twitter and Facebook. We characterize cross-platform similarities and differences in popular sources, diffusion patterns, influencers, coordination, and automation. Comparing the two platforms, we find divergence among the prevalence of popular low-credibility sources and suspicious videos. A minority of accounts and pages exert a strong influence on each platform. These misinformation "superspreaders" are often associated with the low-credibility sources and tend to be verified by the platforms. On both platforms, there is evidence of coordinated sharing of Infodemic content. The overt nature of this manipulation points to the need for societal-level solutions in addition to mitigation strategies within the platforms. However, we highlight limits imposed by inconsistent data-access policies on our capability to study harmful manipulations of information ecosystems.

Keywords

Coordination, COVID-19, Facebook, Infodemic, low-credibility, prevalence, Twitter

This article is a part of special theme on Studying the COVID-19 Infodemic at Scale. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/studyinginfodemicatscale

Introduction

The impact of the COVID-19 pandemic has been felt globally, with almost 70 million detected cases and 1.5 million deaths as of December 2020 (coronavirus.jhu.edu/map.html). Epidemiological strategies to combat the virus require collective behavioral changes. To this end, it is important that people receive coherent and accurate information from media sources that they trust. Within this context, the spread of false narratives in our information environment can have acutely negative repercussions on public health and safety. For example, misinformation about masks greatly contributed to low adoption rates and increased disease transmission (Lyu and Wehby, 2020). The problem is not going away any time soon: false vaccine narratives (Loomba et al., 2021) will drive hesitancy, making it difficult to reach herd immunity and prevent future outbreaks.

It is concerning that many people believe, and many more have been exposed to, misinformation about the pandemic (Mitchell and Oliphant, 2020; Nightingale et al., 2020; Roozenbeek et al., 2020; Schaeffer, 2020). The spread of this misinformation has been termed the *Infodemic* (Zarocostas, 2020). Social media play a strong role in propagating misinformation because of peer-to-peer transmission (Vosoughi et al., 2018). There is also evidence that social media are

[1]Observatory on Social Media, Indiana University, Bloomington, IN, USA
[2]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

Corresponding author:
Kai-Cheng Yang, Indiana University, 901 E 10th Street, Bloomington, IN 47408, USA.
Email: yangkc@iu.edu

manipulated (Shao et al., 2018; Stella et al., 2018) and used to spread COVID-19 misinformation (Ferrara et al., 2020). It is therefore important to better understand how users disseminate misinformation across social media networks.

As described in the literature reviewed in the next section, a limitation of most existing studies regarding misinformation spreading on social media is that they focus on a single platform. However, the modern information ecosystem consists of different platforms over which information propagates concurrently and in diverse ways. Each platform can have different vulnerabilities (Allcott et al., 2019). A key goal of the present work is to compare and contrast the extent to which the Infodemic has spread on Twitter and Facebook.

A second gap in our understanding of COVID-19 misinformation is in the patterns of diffusion within social media. It is important to understand how certain user accounts, or groups of accounts, can play a disproportionate role in amplifying the spread of misinformation. Inauthentic social media accounts, known as social bots and trolls, can play an important role in amplifying the spread of COVID-19 misinformation on Twitter (Ferrara, 2020). The picture on Facebook is less clear, as there is little access to data that would enable a determination of social bot activity. It is, however, possible to look for evidence of manipulation in how multiple accounts can be coordinated with one another, potentially controlled by a single entity. For example, accounts may exhibit suspiciously similar sharing behaviors (Pacheco et al., Forthcoming).

We extract website links from social media posts that include COVID-19 related keywords. We identify a link with low-credibility content in one of two ways. First, we follow the convention of classifying misinformation at the source rather than the article level (Lazer et al., 2018). We do this by matching links to an independently generated corpus of low-credibility website *domains* (or *sources*). Second, in the case of links to YouTube, we label videos as suspicious if they have been banned by the site or are otherwise unavailable to the public. This enables us to quantify the prevalence of individual uploads likely to propagate COVID-19 misinformation and the different ways in which they are shared on Twitter and Facebook.

The main contributions of this study stem from exploring three sets of research questions:

1. What is the prevalence of low-credibility content on Twitter and Facebook? Are there similarities in how sources are shared over time? How does this activity compare to that of popular high-credibility sources? Are the same suspicious sources and YouTube videos shared in similar volumes across the two platforms?

2. Is the sharing of misinformation concentrated around a few active accounts? Do a few influential accounts dominate the resharing of popular misinformation? What is the role of verified accounts and those associated with the low-credibility sources on the two platforms?
3. Is there evidence of inauthentic coordinated behavior in sharing low-credibility content? Can we identify clusters of users, pages, or groups with suspiciously similar sharing patterns? Is low-credibility content amplified by Twitter bots more prevalent on Twitter as compared to Facebook?

After systematically reviewing the literature regarding health misinformation on social media in the next section, we describe the methodology and data employed in our analyses. The following three sections present results to answer the above research questions. Finally, we discuss the limitations of our analyses and implications of our findings for mitigation strategies.

Literature review

Concerns regarding online health-related misinformation existed before the advent of online social media. Studies mostly focused on evaluating the quality of information on the web (Eysenbach et al., 2002), and a new research field emerged, namely, "infodemiology," to assess health-related information on the Internet and address the gap between expert knowledge and public perception (Eysenbach, 2002).

With the wide adoption of online social media, the information ecosystem has seen large changes. Peer-to-peer communication can greatly amplify fake or misleading messages by any individual (Vosoughi et al., 2018). Many studies reported on the presence of misinformation on social media during the time of epidemics such as Ebola (Fung et al., 2016; Jin et al., 2014; Pathak et al., 2015; Sell et al., 2020) and Zika (Bora et al., 2018; Seltzer et al., 2017; Sharma et al., 2017; Wood, 2018). Misinformation surrounding vaccines has been particularly persistent and is likely to reoccur whenever the topic comes into public focus (Bahk et al., 2016; DeVerna et al., 2021; Donzelli et al., 2018; Mahoney et al., 2015; Panatto et al., 2018; Schmidt et al., 2018).

These studies focused on specific social media platforms including Twitter (Mahoney et al., 2015; Wood, 2018), Facebook (Schmidt et al., 2018; Sharma et al., 2017), Instagram (Seltzer et al., 2017), and YouTube (Bora et al., 2018; Donzelli et al., 2018). The most common approach was content-based analysis of sampled social media posts, images, and videos to gauge the topics of online discussions and estimate the prevalence of misinformation. Unfortunately, the datasets

analyzed in these studies were usually small (at a scale of hundreds or thousands of items) due to difficulties in accessing and manually annotating large-scale collections.

Unsurprisingly, the COVID-19 pandemic has inspired a new wave of health misinformation studies. In addition to traditional approaches like qualitative analyses of social media content (Al-Rakhami and Al-Amri, 2020; Dutta et al., 2020; Kouzy et al., 2020; Li et al., 2020; Memon and Carley, 2020; Pulido et al., 2020) and survey studies (Mitchell and Oliphant, 2020; Nightingale et al., 2020; Roozenbeek et al., 2020), quantitative studies on the prevalence of links to low-credibility websites at scale have gained popularity in light of the recent development of computational methods (Broniatowski et al., 2020; Cinelli et al., 2020; Gallotti et al., 2020; Guarino et al., 2021; Singh et al., 2020a, 2020b; Yang et al., 2020b).

Many of these studies aimed to assess the prevalence of, and exposure to, COVID-19 misinformation on online social media (Chou et al., 2018). However, different approaches yielded disparate estimates of misinformation prevalence levels ranging from as little as 1% to as much as 70%. These widely varying statistics indicate that different approaches to experimental design, including uneven access to data on different platforms and inconsistent definitions of misinformation, can generate inconclusive or misleading results. In this study, we follow the reasoning from Gallotti et al. (2020) that it is better to clearly define a misinformation metric and then use it in a comparative way to look at how misinformation varies over time or is influenced by other factors.

We identify two gaps in the literature reviewed above. First, it is still unclear how the spreading patterns can differ on different social networks since studies comparing multiple platforms are rare. This might be due to the obstacles in accessing data from different sources simultaneously and the lack of a unified framework to compare very different services. Second, our understanding of the role that different account groups play during the misinformation dissemination is very limited. We address these gaps in this study.

Methods

In this section, we describe in detail the methodology employed in our analyses, allowing other researchers to replicate our approach. The outline is as follows: we collect social media data from Twitter and Facebook using the same keywords list. We then identify low- and high-credibility content from the tweets and posts automatically by tracking the URLs linking to the domains in a pre-defined list. Finally, we identify suspicious YouTube videos by their availability status.

Identification of low-credibility information

We focus on news articles linked in social media posts and identify those pertaining to low-credibility domains by matching the URLs to sources, following a corpus of literature (Bovet and Makse, 2019; Grinberg et al., 2019; Lazer et al., 2018; Pennycook and Rand, 2019; Shao et al., 2018). We define our list of low-credibility domains based on information provided by the Media Bias/Fact Check website (MBFC, mediabiasfactcheck.com), an independent organization that reviews and rates the reliability of news sources. We gather the sources labeled by MBFC as having a "Very Low" or "Low" factual-reporting level. We then add "Questionable" or "Conspiracy-Pseudoscience" sources and we leave out those with factual-reporting levels of "Mostly-Factual," "High," or "Very High." We remark that although many websites exhibit specific political leanings, these do not affect inclusion in the list. The list has 674 low-credibility domains (Yang et al., 2020a).

High-credibility sources

As a benchmark for interpreting the prevalence of low-credibility content, we also curate a list of 20 more credible information sources. We start from the list provided in a recent Pew Research Center report (Mitchell et al., 2014) and used in a few studies on online disinformation (Pierri et al., 2020a, 2020b), and we select popular news outlets that cover the full U.S. political spectrum. These sources have an MBFC factual-reporting level of "Mixed" or higher. In addition, we include the websites of two organizations that acted as authoritative sources of COVID-19 related information, namely, the Centers for Disease Control and Prevention and World Health Organization. For simplicity, we refer to the full list in Table 1 as high-credibility sources.

Table 1. List of high-credibility sources.

huffpost.com	newyorker.com
msnbc.com	newsweek.com
cnn.com	nytimes.com
economist.com	time.com
washingtonpost.com	reuters.com
apnews.com	npr.org
usatoday.com	wsj.com
foxnews.com	marketwatch.com
nypost.com	dailycaller.com
theblaze.com	dailywire.com
cdc.gov	who.int

Data collection

We collect data related to COVID-19 from both Twitter and Facebook. To provide a general and unbiased view of the discussion, we chose the following generic query terms: "coronavirus", "COVID" (to capture keywords like COVID19 and COVID-19), and "sars" (to capture sars-cov-2 and related variations).

Twitter data. Our Twitter data was collected using an application programming interface (API) from the Observatory on Social Media (Davis et al., 2016), which allows to search tweets from the Decahose, a 10% random sample of public tweets. We searched for English tweets containing the keywords between 1 January and 31 October 2020, resulting in over 53M tweets posted by about 12M users. Note that since the Decahose samples tweets and not users, the sample of users in our Twitter dataset is biased toward more active users.

Our collection contains two types of tweets, namely, original tweets and retweets. The content of original tweets is published by users directly, while retweets are generally used to endorse/amplify original tweets by others (no quoted tweets are included). We refer to authors of original tweets as "root" users and to authors of retweets as "leaf" users (see Figure 1).

Facebook data. We used the *posts/search* endpoint of the CrowdTangle API (CrowdTangle Team, 2020) to collect data from Facebook. We filtered the entire set of English posts published by public pages and groups in the period from 1 January to 31 October 2020 using the above list of keywords, resulting in over 37M posts by over 140k public pages/groups.

Our Facebook data collection is limited by the coverage of pages and groups in CrowdTangle, a public tool owned and operated by Facebook. CrowdTangle includes over 6M Facebook pages and groups: all those with at least 100k followers/members, U.S.-based public groups with at least 2k members, and a very small subset of verified profiles that can be followed like public pages. We include these public accounts among pages and groups. In addition, some pages and groups with fewer followers and members are also included by CrowdTangle upon request from users. This might bias the dataset in ways that are hard to gauge. For example, requests from researchers interested in monitoring low-credibility pages and groups might lead to over-representation of such content.

As shown in Figure 1, the collected data contains information about original Facebook posts and the pages/groups that published these posts. For each post, we also have access to *aggregate* statistics such as the number of reshares, comments, and reactions
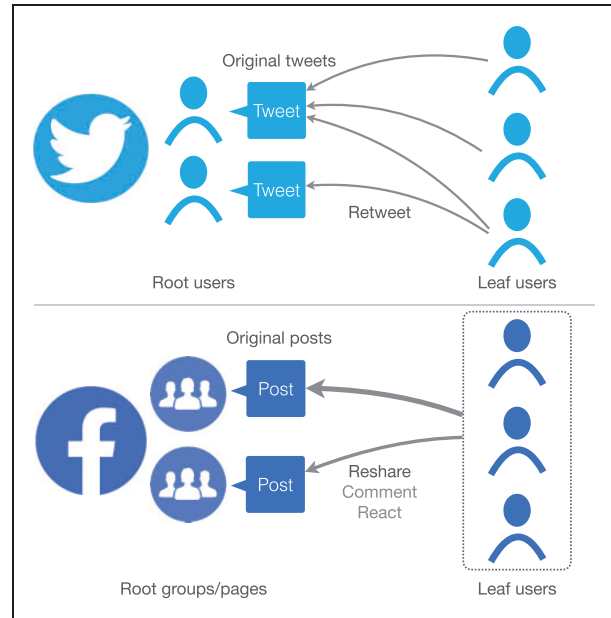

Figure 1. Structure of the data collected from Twitter and Facebook. On Twitter, we have the information about original tweets, retweets, and all the accounts involved. On Facebook, we have information about original posts and public groups/pages that posted them. For each post, we also have aggregate numbers of reshares, comments, and reactions, with no information about the users responsible for those interactions.
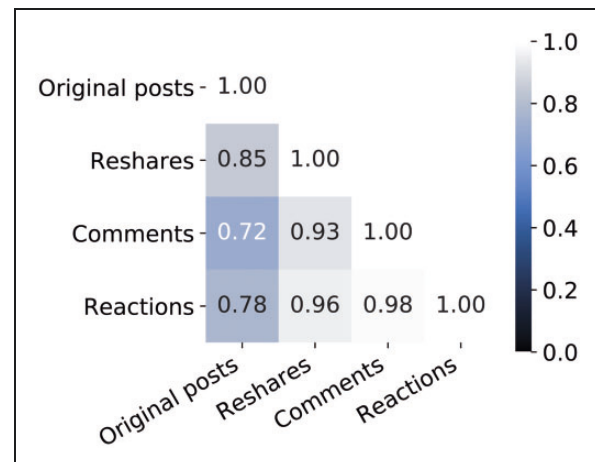

Figure 2. Pearson correlation coefficients between Facebook metrics aggregated at the domain level for low-credibility domains. A reaction can be a "like," "love," "wow," "haha," "sad," "angry," or "care." All correlations are significant ($p < 0.01$).

(e.g., "likes") by Facebook users. The numbers of comments and reactions are highly correlated with reshares (Figure 2), so we focus on reshares in this study.

Similarly to Twitter, Facebook pages and groups that publish posts are referred to as "roots" and users who reshare them are "leaves." However, in contrast to

Twitter, we do not have access to any information about leaf users on Facebook. We refer generically to Twitter users and Facebook pages and groups as "accounts."

To compare Facebook and Twitter in a meaningful way, we compare root users with root pages/groups, original tweets with original posts, and retweet counts with reshare counts. We define *prevalence* as the sum of original tweets and retweets on Twitter, and as the sum of original posts and reshares on Facebook.

YouTube data. We observed a high prevalence of links pointing to youtube.com on both platforms—over 64k videos on Twitter and 204k on Facebook. Therefore, we also provide an analysis of popular videos published on Facebook and Twitter. Specifically, we focus on popular YouTube videos that are likely to contain low-credibility content. An approach analogous to the way we label links to websites would be to identify sources that upload low-credibility videos and then label every video from those sources as misinformation. However, this approach is infeasible because the list of YouTube channels would be huge and fluid. To circumvent this difficulty, we use removal of videos by YouTube as a proxy to label low-credibility content. We additionally consider private videos to be suspicious, since this can be used as a tactic to evade the platform's sanctions when violating terms of service.

To identify the most popular and suspicious YouTube content, we first select the 16,669 videos shared at least once on both platforms. We then query the YouTube API *Videos:list* endpoint to collect their metadata and focus on the 1,828 (11%) videos that had been removed or made private. To validate this approach for identifying low-credibility YouTube content, we follow a two-step manual inspection process for a sample of about 3% of the unavailable videos, comprising a mix of randomly selected and popular ones. We first search for the deleted video IDs in other YouTube videos and web pages. When these references contain the deleted videos' titles, we search for these titles on bitchute.com to find copies of the original videos. This process allows us to identify the narratives of 40 deleted videos, 90% of which contain misinformation. A similar approach was also adopted by Knuutila et al. (2020) in their recent study of COVID-19 misinformation on YouTube.

Ethical considerations. Our studies of public Twitter and Facebook data have been granted exemption from Institutional Review Board review (Indiana University protocols 1102004860 and 10702, respectively). The data collection and analysis are also in compliance with the terms of service of the corresponding social media platforms.

Table 2. List of URL shortening services.

bit.ly	dlvr.it	liicr.nl	tinyurl.com
goo.gl	ift.tt	ow.ly	fxn.ws
buff.ly	back.ly	amzn.to	nyti.ms
nyp.st	dailysign.al	j.mp	wapo.st
reut.rs	drudge.tw	shar.es	sumo.ly
rebrand.ly	covfefe.bz	trib.al	yhoo.it
t.co	shr.lc	po.st	dld.bz
bitly.com	crfrm.us	flip.it	mf.tt
wp.me	voat.co	zurl.co	fw.to
mol.im	read.bi	disq.us	tmsnrt.rs
usat.ly	aje.io	sc.mp	gop.cm
crwd.fr	zpr.io	scq.io	trib.in
owl.li			

Link extraction

Estimating the prevalence of low-credibility information requires matching URLs, extracted from tweets and Facebook metadata, against our lists of low- and high-credibility websites. As shortened links are very common, we also identified 49 link shortening services that appear at least 50 times in our datasets (Table 2) and expanded shortened URLs referring to these services through HTTP requests to obtain the actual domains. We finally match the extracted and expanded links against the lists of low- and high-credibility domains. A breakdown of matched posts/tweets is shown in Table 3. For low-credibility content, the ratio of retweets to tweets is 2.7:1, while the ratio of reshares to posts is 68:1. This large discrepancy is due to various factors: the difference in traffic on the two platforms, the fact that we only have a 10% sample of tweets, and the bias toward popular pages and groups on Facebook.

Infodemic prevalence

In this section, we provide results about the prevalence of links to low-credibility domains on the two platforms. As described in the "Methods" section, we sum tweets and retweets for Twitter, and original posts and reshares for Facebook. Note that deleted content is not included in our data. Therefore, our estimations should be considered as lower bounds for the prevalence of low-credibility information on both platforms.

Prevalence trends

We plot the daily prevalence of links to low-credibility sources on Twitter and Facebook in Figure 3(a). The two time series are strongly correlated (Pearson $r = 0.87$, $p < 0.01$). They both experience a drastic growth during March, when the number of COVID-19 cases was growing worldwide. Towards

Table 3. Breakdown of Facebook and Twitter posts/tweets matched to low- and high-credibility domains.

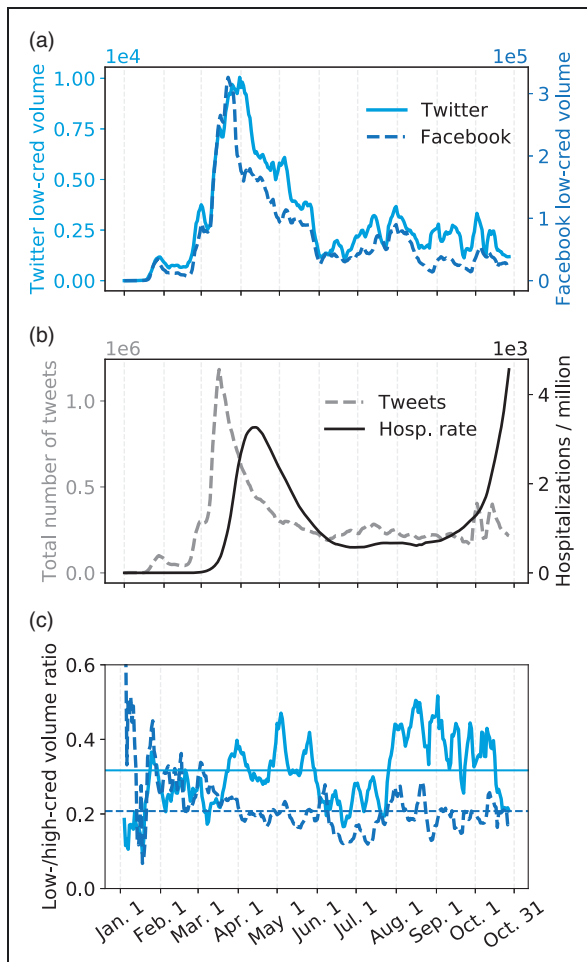|	Low-credibility	High-credibility
Facebook		
Original posts	303,119	1,194,634
Reshares	20,462,035	98,415,973
Twitter		
Original tweets	245,620	734,409
Retweets	653,415	2,184,050

Figure 3. Infodemic content surge on both platforms around the COVID-19 pandemic waves, from 1 January to 31 October 2020. All curves are smoothed via 7-day moving averages. (a) Daily volume of posts/tweets linking to low-credibility domains on Twitter and Facebook. Left and right axes have different scales and correspond to Twitter and Facebook, respectively. (b) Overall daily volume of pandemic-related tweets and worldwide COVID-19 hospitalization rates (data source: Johns Hopkins University). (c) Daily ratio of volume of low-credibility links to volume of high-credibility links on Twitter and Facebook. The noise fluctuations in early January are due to low volume. The horizontal lines indicate averages across the period starting February 1.

summer, the prevalence of low-credibility information decreases to a relatively low level and then becomes more stable.

To analyze the Infodemic surge with respect to the pandemic's development and public awareness, Figure 3(b) shows the worldwide hospitalization rate and the overall volume of tweets in our collection. The Infodemic surge roughly coincides with the general attention given to the pandemic, captured by the overall Twitter volume. The peak in hospitalizations trails by a few weeks. A similar delay was recently reported between peaks of exposure to Infodemic tweets and of COVID-19 cases in different countries (Gallotti et al., 2020). This plot suggests that the delay is related to general attention toward the pandemic rather than specifically toward misinformation.

To further explore whether the decrease in low-credibility information is organic or due to platform interventions, we also compare the prevalence of low-credibility content to that of links to credible sources. As shown in Figure 3(c), the ratios are relatively stable across the observation period. These results suggest that the prevalence of low-credibility content is mostly driven by the public attention to the pandemic in general, which progressively decreases after the initial outbreak. We finally observe that Twitter exhibits a higher ratio of low-credibility information than Facebook (32% vs. 21% on average).

Prevalence of specific domains

We use the high-credibility domains as a benchmark to assess the prevalence of low-credibility domains on each platform. As shown in Figure 4, we notice that the low-credibility sources exhibit disparate levels of prevalence. Low-credibility content as a whole reaches considerable volume on both platforms, with prevalence surpassing every single high-credibility domain considered in this study. On the other hand, low-credibility domains generally exhibit much lower prevalence compared to high-credibility ones (with a few exceptions, notably thegatewaypundit.com and breitbart.com).

Source popularity comparison

As shown in Figure 4, we observe that low-credibility websites may have different prevalence on the two platforms. To further contrast their prevalence levels on Twitter and Facebook, we measure the popularity of websites on each platform by ranking them by prevalence and then compare the resulting ranks in Figure 5. The ranks on the two platforms are not strongly correlated (Spearman $r = 0.57$, $p < 0.01$). A few domains are much more popular or only appear on one of the
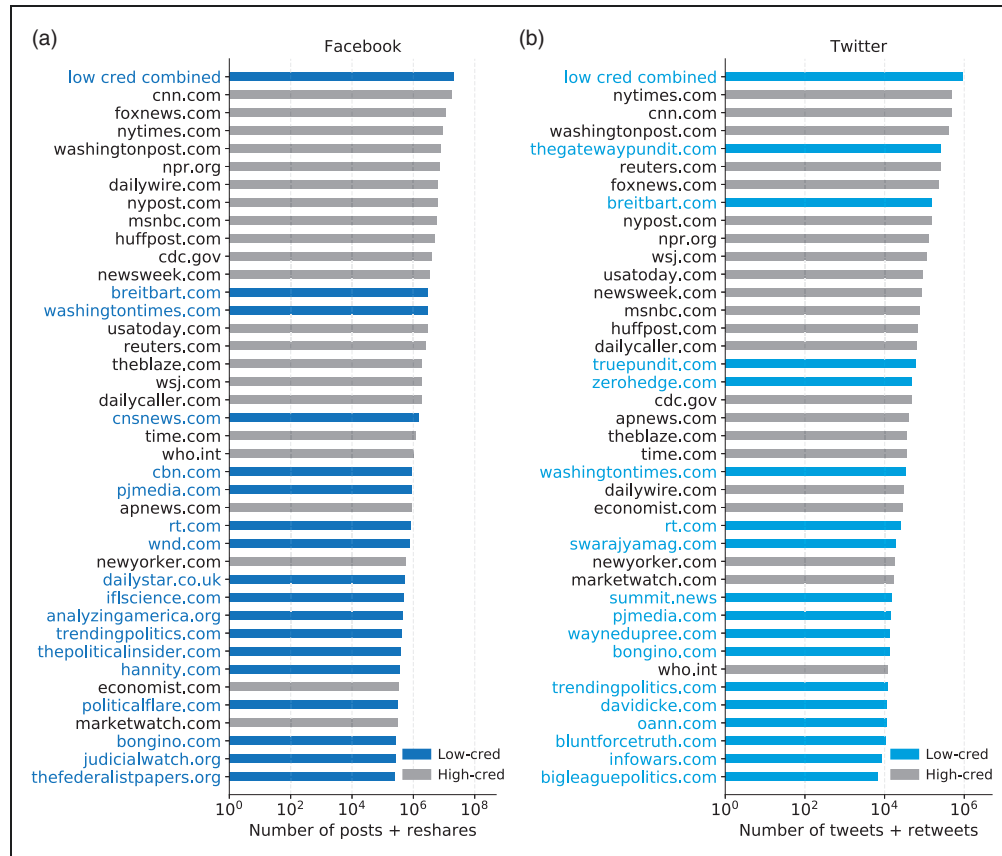
Figure 4. Total prevalence of links to low- and high-credibility domains on both (a) Facebook and (b) Twitter. Due to space limitation, we only show the 40 most frequent domains on the two platforms. The high-credibility domains are all within the top 40. We also show low-credibility information as a whole (cf. "low cred combined").

platforms (see annotations in Figure 5(a)). We also show the domains that are very popular on both platforms in Figure 5(b). They are dominated by right-wing and state sources, such as breitbart.com, washingtontimes.com, thegatewaypundit.com, oann.com, and rt.com.

YouTube Infodemic content

Thus far, we examined the prevalence of links to low-credibility web page sources. However, a significant portion of the links shared on Twitter and Facebook point to YouTube videos, which can also carry COVID-19 misinformation. Previous work has shown that bad actors utilize YouTube in this manner for their campaigns (Wilson and Starbird, 2020). Specifically, anti-scientific narratives on YouTube about vaccines, Idiopathic Pulmonary Fibrosis, and the COVID-19 pandemic have been documented (Donzelli et al., 2018; Dutta et al., 2020; Goobie et al., 2019; Knuutila et al., 2020).

To measure the prevalence of Infodemic content introduced from YouTube, we consider the unavailability (deletion or private status) of videos as an indicator of suspicious content, as explained in the "Methods" section. Figure 6 compares the prevalence rankings on Twitter and Facebook for unavailable videos ranked within the top 500 on both platforms. These videos are linked between 6 and 980 times on Twitter and between 39 and 64,257 times on Facebook. While we cannot apply standard rank correlation measures due to the exclusion of low-prevalence videos, we do not observe a correlation in the cross-platform popularity of suspicious content from a qualitative inspection of the figure. A caveat to this analysis is that the same video content (sometimes re-edited) can be recycled within many video uploads, each having a unique video ID. Some of these videos are promptly removed while others are not. Therefore, the lack of correlation could partly be driven by YouTube's efforts to remove Infodemic content in conjunction with attempts by uploaders to counter those efforts (Knuutila et al., 2020).

Having looked at the prevalence of suspicious content from YouTube, we wish to explore the question from another angle: are videos that are popular on Facebook or Twitter more likely to be flagged as
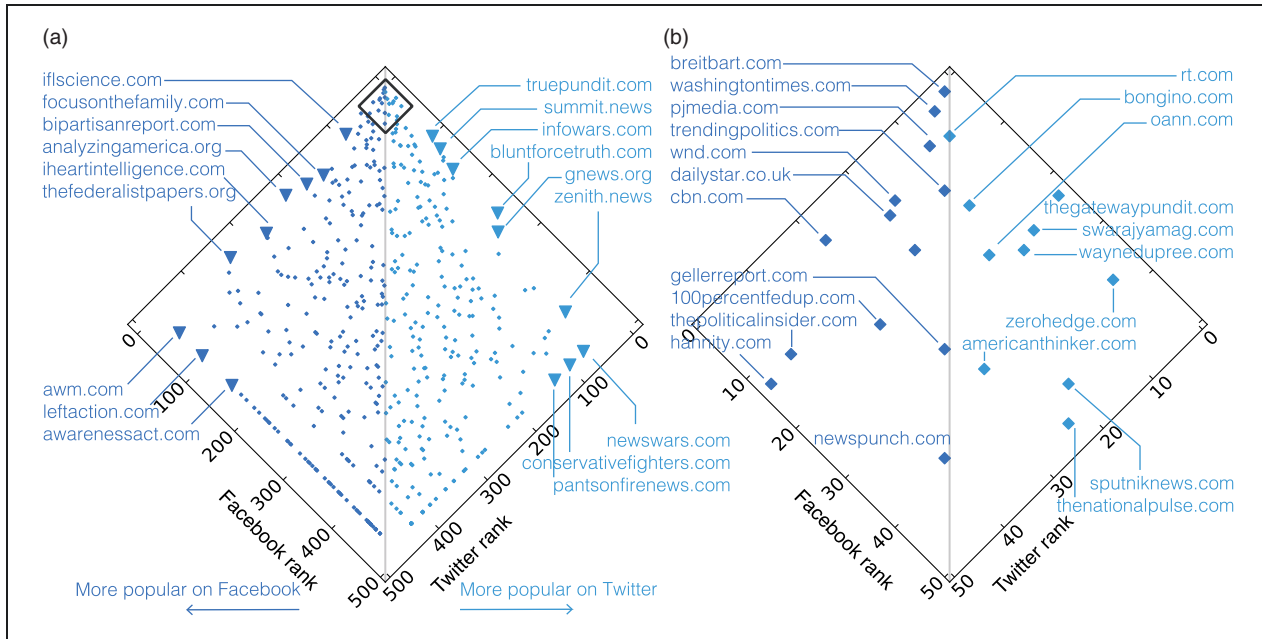
Figure 5. (a) Rank comparison of low-credibility sources on Facebook and Twitter. Each dot in the figure represents a low-credibility domain. The most popular domain ranks first. Domains close to the vertical line have similar ranks on the two platforms. Domains close to the edges are much more popular on one platform or the other. We annotate a few selected domains that exhibit high rank discrepancy. (b) A zoom-in on the sources ranked among the top 50 on both platforms (highlighted square in (a)).
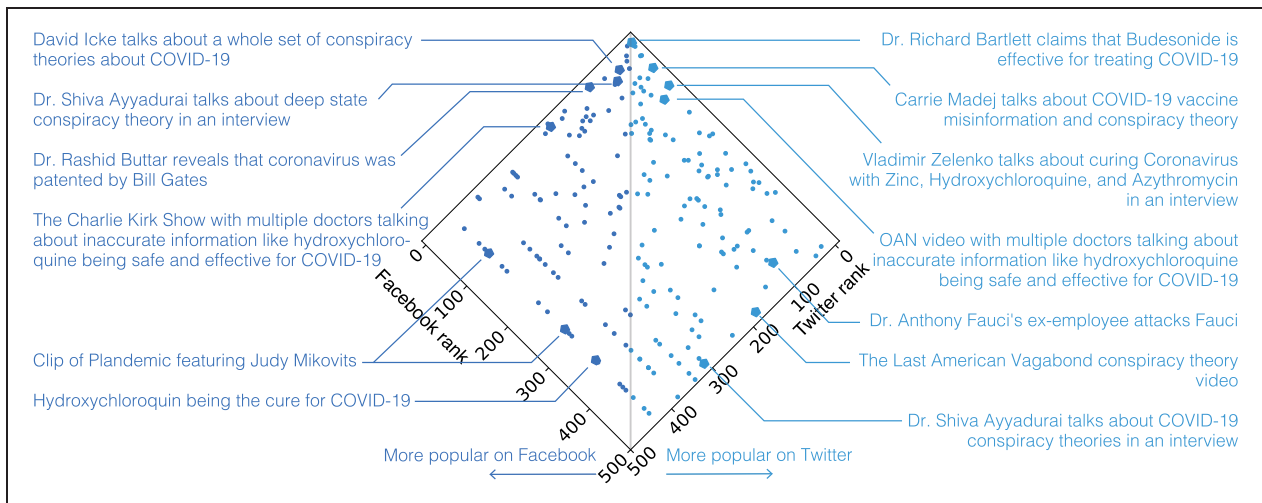

Figure 6. Rank comparison of suspicious YouTube videos within the top 500 on both Facebook and Twitter. The most popular video ranks first. Each dot in the figure represents a suspicious video. Videos close to the vertical line have similar ranks on both platforms. Videos close to the edges are more popular on one platform or the other. We annotated a few selected videos with their narratives extracted from their copies on bitchute.com or other web pages.

suspicious? Figure 7 shows this to be the case on both platforms: a larger portion of videos with higher prevalence are unavailable, but the trend is stronger on Twitter than on Facebook. The overall trend suggests that YouTube may have a bias toward moderating videos that attract more attention. This may be a function of the fact that an Infodemic video that is spreading virally on Twitter/Facebook may receive more abuse reports through YouTube's reporting mechanism. The fact that this trend is greater on Twitter may be explained by the differences between each platform's demographics. Survey data cited in the "Discussion" section shows that Twitter users are younger and more educated; it is therefore plausible that the average Twitter user is more likely to report unreliable content.
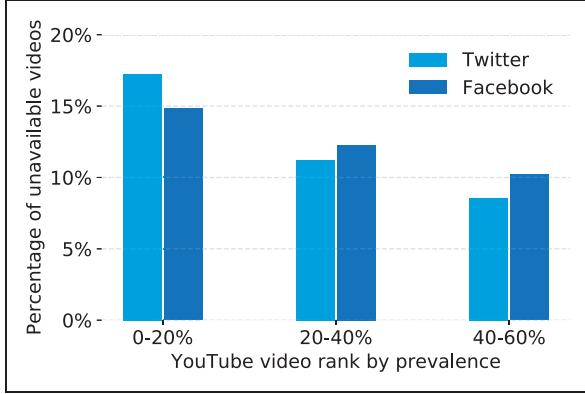
Figure 7. Percentages of suspicious YouTube videos against their percent rank among all videos linked from pandemic-related tweets/posts on both Twitter and Facebook.

Infodemic spreaders

Links to low-credibility sources are published on social media through original tweets and posts first, then retweeted and reshared by leaf users. In this section, we study this dissemination process on Twitter and Facebook with a focus on the top spreaders, or "superspreaders": those disproportionately responsible for the distribution of Infodemic content.

Concentration of influence

We wish to measure whether low-credibility content originates from a wide range of users or can be attributed to a few influential actors. For example, a source published 100 times could owe its popularity to 100 distinct users or to a single root whose post is republished by 99 leaf users. To quantify the concentration of original posts/tweets or reshares/retweets for a source s, we use the inverse normalized entropy (Nikolov et al., 2019), defined as

$$C_s = 1 + \sum_{r=1}^{N_s} \frac{P_r(s)\log P_r(s)}{\log N_s}, \quad (1)$$

where r represents a root user/group/page linking to source s, $P_r(s)$ stands for the fraction of posts/tweets or reshares/retweets linking to s and associated with r, and N_s is the total number of roots linking to s. Entropy measures how evenly quantities of content are distributed across roots; it is normalized to account for the varying numbers of roots in different cases. Its inverse captures concentration and is defined in the unit interval. It is maximized at 1 when the content originates from a single root user/group/page, and minimized at 0 when each root makes an equal contribution. We set $C_s = 1$ when $N_s = 1$.

Let us gauge the *concentration of activity* around root accounts through their numbers of original tweets/posts for each source. Similarly, we calculate the *concentration of popularity* around the root accounts using their numbers of retweets/reshares for each source. We show the distributions of these concentration variables in Figure 8(a). On both platforms, we find that popularity is significantly more concentrated around root accounts compared to their activity ($p < 0.001$ for paired sample t tests). This suggests the existence of superspreaders: despite the diversity of root accounts publishing links to low-credibility content on both platforms, only messages from a small group of influential accounts are shared extensively.

Who are the Infodemic superspreaders?

Both Twitter and Facebook provide verification of accounts and embed such information in the metadata. Although the verification processes differ, we wish to explore the hypothesis that verified accounts on either platform play an important role as top spreaders of low-credibility content. Figure 8(b) compares the popularity of verified accounts to unverified ones on a per-account basis. We find that verified accounts tend to receive a significantly higher number of retweets/reshares on both platforms ($p < 0.001$ for Mann–Whitney U tests).

We further compute the proportion of original tweets/posts and retweets/reshares that correspond to verified accounts on both platforms. Verified accounts are a small minority compared to unverified ones, that is, 1.9% on Twitter and 4.5% on Facebook, among root accounts involved in publishing low-credibility content. Despite this, Figure 8(c) shows that verified accounts yield almost 40% of low-credibility retweets on Twitter and almost 70% of reshares on Facebook.

These results suggest that verified accounts play an outsize role in the spread of Infodemic content. Are superspreaders all verified? To answer this question, let us analyze superspreader accounts separately for each low-credibility source. We extract the top user/page/group (i.e., the account with most retweets/reshares) for each source and find that 19% and 21% of them are verified on Twitter and Facebook, respectively. While these values are much higher than the percentages of verified accounts among all roots, they show that not all superspreaders are verified.

Who are the top spreaders of Infodemic content? Table 4 answers this question for the 23 top low-credibility sources in Figure 5(b). We find that the top spreader for each source tends to be the corresponding official account. For instance, about 20% of the retweets containing links to thegatewaypundit.com pertain to @gatewaypundit, the official handle
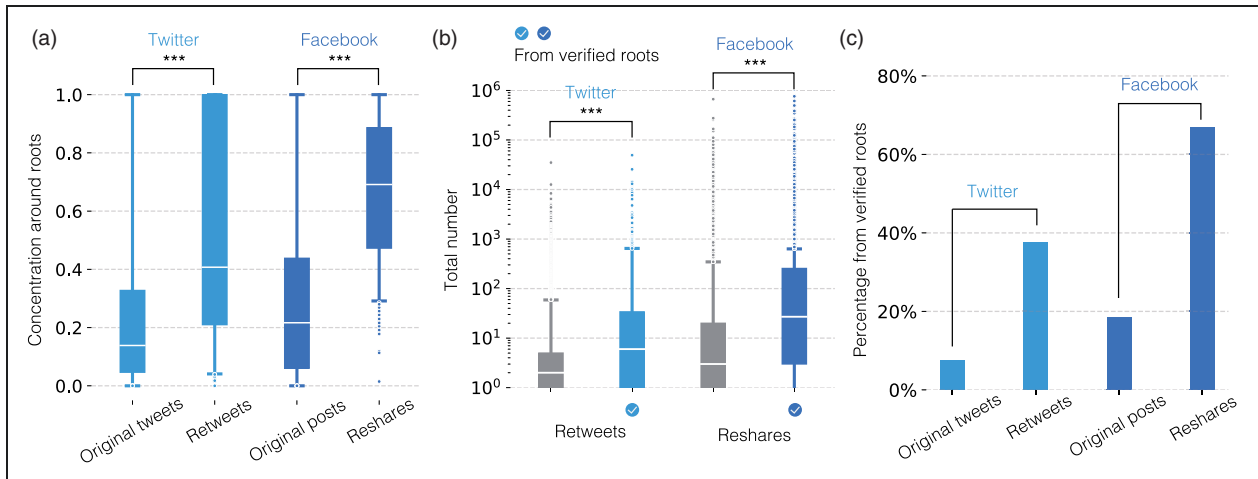
Figure 8. Evidence of Infodemic superspreaders. Boxplots show the median (white line), 25th–75th percentiles (boxes), 5th–95th percentiles (whiskers), and outliers (dots). Significance of statistical tests is indicated by *** ($p < 0.001$). (a) Distributions of the concentration of original tweets, retweets, original posts, and reshares linking to low-credibility domains around root accounts. Each domain corresponds to one observation. (b) Distributions of the total number of retweets and reshares of low-credibility content posted by verified and unverified accounts. Each account corresponds to one observation. (c) Fractions of original tweets, retweets, original posts, and reshares by verified accounts.

associated with *The Gateway Pundit* website, on Twitter. (The @gatewaypundit account was suspended by Twitter in February 2021.) The remaining retweets have 10,410 different root users. Similarly, on Facebook, among all 2,821 pages/groups that post links to thegatewaypundit.com, the official page @gatewaypundit accounts for 68% of the reshares. We observe in Table 4 that most of the top low-credibility sources have official accounts on both Twitter and Facebook, which tend to be verified (71.4% on Twitter and 65.2% on Facebook). They are also the top spreaders of those domains in 16 of 21 cases (76.2%) on Twitter and 18 of 23 (78.3%) on Facebook.

Infodemic manipulation

Here, we consider two types of inauthentic behaviors that can be used to spread and amplify COVID-19 misinformation: coordinated networks and automated accounts.

Coordinated amplification of low-credibility content

Social media accounts can act in a coordinated fashion (possibly controlled by a single entity) to increase influence and evade detection (Nizzoli et al., 2020; Pacheco et al., Forthcoming; Sharma et al., 2020). We apply the framework proposed by Pacheco et al. (Forthcoming) to identify coordinated efforts in promoting low-credibility information, both on Twitter and Facebook.

The idea is to build a network of accounts where the weights of edges represent how often two accounts link

to the same domains. A high weight on an edge means that there is an unusually high number of domains shared by the two accounts. We first construct a bipartite graph between accounts and low-credibility domains linked in tweets/posts. The edges of the bipartite graph are weighted using term frequency-inverse document frequency (TF-IDF) (Spärck Jones, 1972) to discount the contributions of popular sources. Each account is therefore represented as a TF-IDF vector of domains. A projected co-domain network is finally constructed, with edges weighted by the cosine similarity between the account vectors.

We apply two filters to focus on active accounts and highly similar pairs. On Twitter, the users must have at least 10 tweets containing low-credibility links, and we retain edges with similarity above 0.99. On Facebook, the pages/groups must have at least five posts containing links, and we retain edges with similarity above 0.95. These thresholds are selected by manually inspecting the outputs.

Figure 9 shows densely connected components in the co-domain networks for Twitter and Facebook. These clusters of accounts share suspiciously similar sets of sources. They likely act in a coordinated fashion to amplify Infodemic messages and are possibly controlled by the same entity or organization. We highlight the fact that using a more stringent threshold on the Twitter dataset yields a higher number of clusters than a more lax threshold on the Facebook dataset. However, this does not necessarily imply a higher level of abuse on Twitter; it could be due to the difference in the units of analysis. On Facebook, we only

Table 4. Official social media handles for the 23 top low-credibility sources from Figure 5(b).

Domain	Twitter handle	Facebook page/group
thegatewaypundit.com	@gatewaypundit ✓* †	@gatewaypundit ✓*
breitbart.com	@BreitbartNews ✓*	@Breitbart ✓*
zerohedge.com	@zerohedge *	@ZeroHedge *
washingtontimes.com	@WashTimes ✓*	@TheWashingtonTimes ✓
rt.com	@RT_com ✓*	@RTnews ✓*
swarajyamag.com	@SwarajyaMag ✓*	@swarajyamag ✓*
pjmedia.com	@PJMedia_com *	@PJMedia ✓*
waynedupree.com	@WayneDupreeShow ✓*	@WayneDupreeShow ✓*
bongino.com	@dbongino ✓*	@dan.bongino ✓*
trendingpolitics.com	–	@trendingpoliticsdotcom
oann.com	@OANN ✓*	@OneAmericaNewsNetwork ✓*
wnd.com	@worldnetdaily ✓*	@WNDNews *
sputniknews.com	@SputnikInt ✓*	@SputnikNews ✓*
dailystar.co.uk	@dailystar ✓*	@thedailystar ✓*
politicalflare.com	@nicolejames	@nicolejameswriter
thenationalpulse.com	@RaheemKassam ✓*	@thenationalpulse *
americanthinker.com	@AmericanThinker	@AmericanThinker *
gellerreport.com	@PamelaGeller ✓†	@pamelageller ✓
cbn.com	@CBNOnline ✓	@cbnonline ✓
100percentfedup.com	@100PercFEDUP *	@100PercentFEDUp *
newspunch.com	–	@thepeoplesvoicetv *
thepoliticalinsider.com	@TPInsidr	@ThePoliticalInsider ✓*
hannity.com	@seanhannity ✓*	@SeanHannity ✓*

Note: Accounts with a checkmark (✓) are verified. Accounts with an asterisk (*) are the top spreaders for the corresponding domains. Accounts with a dagger (†) were suspended as of February 2021.

have access to public groups and pages with a bias toward high popularity and not to all accounts as on Twitter.

An examination of the sources shared by the suspicious clusters on both platforms shows that they are predominantly right-leaning and mostly U.S.-centric. The list of domains shared by likely coordinated accounts on Twitter is mostly concentrated on the leading low-credibility sources, such as breitbart.com and thegatewaypundit.com, while likely coordinated groups and pages on Facebook link to a more varied list of sources. Some of the amplified websites feature polarized rhetoric, such as defending against "attack by enemies" (see www.frontpagemag.com/about) or accusations of "liberal bias" (cnsnews.com/about-us), among others. Additionally, there are clusters on both platforms that share Russian state-affiliated media such as rt.com and an Indian right-wing magazine (swarajyamag.com).

In terms of the composition of the clusters, they mostly consist of users, pages, and groups that mention President Trump or his campaign slogans. Some of the Facebook clusters are notable because they consist of groups or pages that are owned by organizations with a wider reach beyond the platform, or that are given an appearance of credibility by being verified. Examples of the former are the pages associated with *The Answer* radio stations (highlighted in Figure 9). These are among 100 stations owned by the publicly traded Salem Media Group, which also airs content on 3,100 affiliate stations. Examples of verified pages engaged in likely coordinated behavior are those affiliated with the non-profit Media Research Center, some of which have millions of followers. On Twitter, some of the clusters include accounts with tens of thousands of followers. Many of the suspicious accounts in Figure 9 no longer exist.

Role of social bots

We are interested in revealing the role of inauthentic actors in spreading low-credibility information on social media. One type of inauthentic behavior stems from accounts controlled in part by algorithms, known as *social bots* (Ferrara et al., 2016). Malicious bots are known to spread low-credibility information (Shao et al., 2018) and in particular create confusion in the online debate around health-related topics like vaccination (Broniatowski et al., 2018).

We adopt BotometerLite (rapidapi.com/OSoMe/api/botometer-pro), a publicly available tool that allows efficient bot detection on Twitter (Yang et al., 2020c). BotometerLite generates a bot score between 0
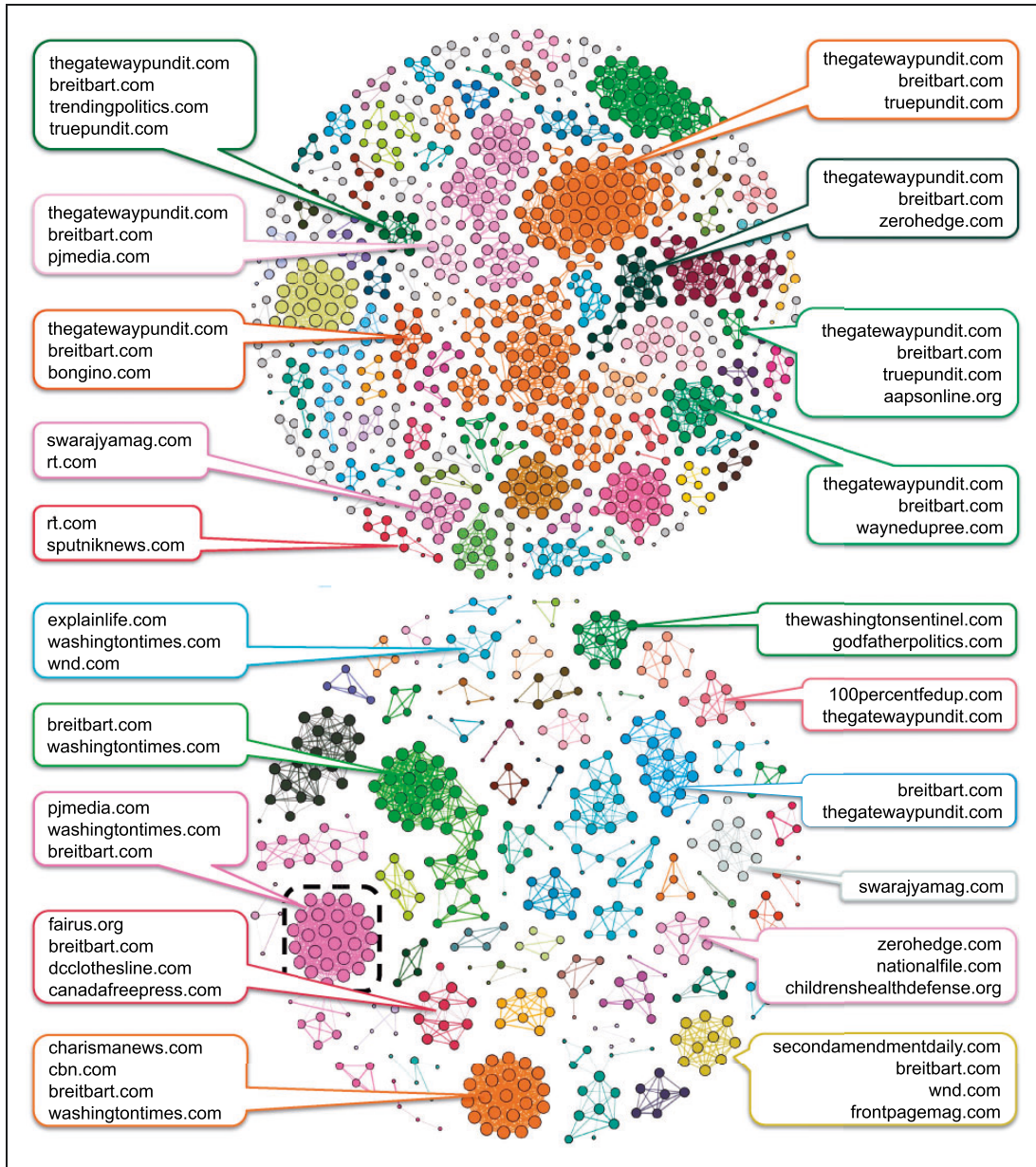
Figure 9. Networks showing clusters that share suspiciously similar sets of sources on (top) Twitter and (bottom) Facebook. Nodes represent Twitter users or Facebook pages/groups. The size of the each node is proportional to its degree. Edges are drawn between pairs of nodes that share an unlikely high number of the same low-credibility domains. The edge weight represents the number of co-shared domains. The most shared sources are annotated for some of the clusters. Facebook pages associated with Salem Media Group radio stations are highlighted by a dashed box.

and 1 for each Twitter account; higher scores indicate bot-like profiles. To the best of our knowledge, there are no similar techniques designed for Facebook because insufficient training data is available. Therefore, we limit this analysis to Twitter.

When applying BotometerLite to our Twitter dataset, we use 0.5 as the threshold to categorize accounts as likely humans or likely bots. For each domain, we calculate the total number of original tweets plus retweets authored by likely humans (n_h) and bots (n_b). We plot the relationship between the two in Figure 10. The linear trend on the log–log plot signifies a power law $n_h \sim n_b^\gamma$ with exponent $\gamma \approx 1.04$, suggesting a weak level of bot amplification (4%) (Shao et al., 2018).

While we are unable to perform automation detection on Facebook groups and pages, the ranks of the low-credibility sources on both platforms allow us to investigate whether sources with more Twitter bot
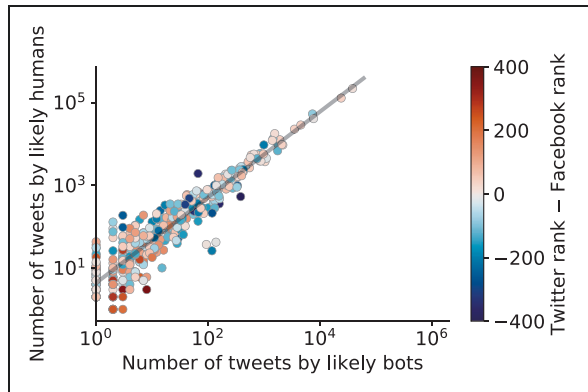
Figure 10. Total number of tweets with links posted by likely humans versus likely bots for each low-credibility source. The slope of the fitted line is 1.04. The color of each source represents the difference between its popularity rank on the two platforms. Red means more popular on Facebook, blue more popular on Twitter.

activity are more prevalent on Twitter or Facebook. For each domain, we calculate the difference of its ranks on Twitter and Facebook and use the value of the difference to color the dots in Figure 10. The results show that sources with more bot activity on Twitter are equally shared on both platforms.

Discussion

In this study, we provide the first comparison between the prevalence of low-credibility content related to the COVID-19 pandemic on two major social media platforms, namely, Twitter and Facebook. Our results indicate that the primary drivers of low-credibility information tend to be high-profile, official, and verified accounts. We also find evidence of coordination among accounts spreading Infodemic content on both platforms, including many controlled by influential organizations. Since automated accounts do not appear to play a strong role in amplifying content, these results indicate that the COVID-19 Infodemic is an overt, rather than a covert, phenomenon.

We find that low-credibility content, as a whole, has higher prevalence than content from any single high-credibility source. However, there is evidence of differences in the misinformation ecosystems of the two platforms, with many low-credibility websites and suspicious YouTube videos at higher prevalence on one platform when compared to the other. Such a discrepancy might be due to a combination of the supply and demand factors. On the supply side, the official accounts associated with specific low-credibility websites are not symmetrically present on both platforms. On the demand side, the two platforms have very different user demographics. According to recent surveys,

69% of adults in the United States say they use Facebook, but only 22% of adults are on Twitter. Furthermore, while Facebook usage is relatively common across a range of demographic groups, Twitter users tend to be younger, more educated, and have higher than average income. Finally, Facebook is a pathway to consuming online news for around 43% of U.S. adults, while the same number for Twitter is 12% (Perrin and Anderson, 2019; Wojcik and Adam, 2019).

During the first months of the pandemic, we observe similar surges of low-credibility content on both platforms. The strong correlation between the timelines of low- and high-credibility content volume reveals that these peaks were likely driven by public attention to the crisis rather than by bursts of malicious content.

Our results provide us with a way to assess how effective the two platforms have been at combating the Infodemic. The ratio of low- to high-credibility information on Facebook is lower than on Twitter, suggesting that Facebook may be more effective. On the other hand, we also find that verified accounts played a stronger role on Facebook than Twitter in spreading low-credibility content. However, the accuracy of these comparisons is subject to the different data collection biases. Suspicious YouTube uploads also exhibit an asymmetric prevalence between Facebook and Twitter. As stated previously, this may be partly a result of uploaders recycling sections of videos and uploading the content with a new video ID. Having such duplicates can mean that one version becomes popular on Facebook and another on Twitter, each potentially shared by a different demographic. This asymmetry might also be driven by Twitter users being more likely to flag videos. YouTube may then quickly remove reported videos before Facebook users have a chance to share them.

There are a number of limitations to our work. As we have remarked throughout the paper, differences between platform data availability and biases in sampling and selection make direct and fair comparisons impossible in many cases. The content collected from the Twitter Decahose is biased toward active users due to being sampled on a per-tweet basis. The Facebook accounts provided by CrowdTangle are biased toward popular pages and public groups, and data availability is also based upon requests made by other researchers. The small set of keywords driving our data collection pipeline may have introduced additional biases in the analyses. This is an inevitable limitation of any collection system, including the Twitter COVID-19 stream (developer.twitter.com/en/docs/labs/covid19-stream/filtering-rules). The use of source-level rather than article-level labels for selecting low-credibility content is necessary (Lazer et al., 2018)

but not ideal; some links from low-credibility sources may point to credible information. In addition, the list of low-credibility sources was not specifically tailored to our subject of inquiry. Finally, we do not have access to many deleted Twitter and Facebook posts, which may lead to an underestimation of the Infodemic's prevalence. All of these limitations highlight the need for cross-platform, privacy-sensitive protocols for sharing data with researchers (Pasquetto et al., 2020).

Low-credibility information on the pandemic is an ongoing concern for society. Our study raises a number of questions. For example, user demographics might strongly affect the consumption of low-credibility information on social media: how do users in distinct demographic groups interact with different information sources? The answer to this question can lead to a better understanding of the Infodemic and more effective moderation strategies but will require methods that scale with the nature of Big Data from social media. Another critical question is how social media platforms are handling the flow of information and allowing dangerous content to spread. Regrettably, since we find that high-status accounts play an important role, addressing this problem will prove difficult. As Twitter and Facebook have increased their moderation of COVID-19 misinformation, they have been accused of political bias. While there are many legal and ethical considerations around free speech and censorship, our work suggests that these questions cannot be avoided and are an important part of the debate around how we can improve our information ecosystem.

Declaration of conflicting interests

Funding

ORCID iDs

Kai-Cheng Yang (iD) https://orcid.org/0000-0003-4627-9273
Francesco Pierri (iD) https://orcid.org/0000-0002-9339-7566
David Axelrod (iD) https://orcid.org/0000-0002-9840-8711
John Bryden (iD) https://orcid.org/0000-0002-5301-5367
Filippo Menczer (iD) https://orcid.org/0000-0003-4384-2876

References

Allcott H, Gentzkow M and Yu C (2019) Trends in the diffusion of misinformation on social media. *Research & Politics* 6(2): 2053168019848554.

Al-Rakhami MS and Al-Amri AM (2020) Lies kill, facts save: Detecting COVID-19 misinformation in Twitter. *IEEE Access* 8: 155961–155970.

Bahk CY, Cumming M, Paushter L, et al. (2016) Publicly available online tool facilitates real-time monitoring of vaccine conversations and sentiments. *Health Affairs* 35(2): 341–347.

Bora K, Das D, Barman B, et al. (2018) Are internet videos useful sources of information during global public health emergencies? A case study of YouTube videos during the 2015–16 Zika virus pandemic. *Pathogens and Global Health* 112(6): 320–328.

Bovet A and Makse HA (2019) Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications* 10(1): 1–14.

Broniatowski DA, Jamison AM, Qi S, et al. (2018) Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health* 108(10): 1378–1384.

Broniatowski DA, Kerchner D, Farooq F, et al. (2020) The COVID-19 social media Infodemic reflects uncertainty and state-sponsored propaganda. *arXiv:2007.09682*.

Chou WYS, Oh A and Klein WM (2018) Addressing health-related misinformation on social media. *Journal of the American Medical Association* 320(23): 2417–2418.

Cinelli M, Quattrociocchi W, Galeazzi A, et al. (2020) The COVID-19 social media Infodemic. *Scientific Reports* 10(1): 16598.

CrowdTangle Team (2020) *CrowdTangle*. Menlo Park: Facebook.

Davis CA, Ciampaglia GL, Aiello LM, et al. (2016) OSoMe: the IUNI observatory on social media. *PeerJ Computer Science* 2: e87.

DeVerna MR, Pierri F, Truong BT, et al. (2021) CoVaxxy: A collection of English-language Twitter posts about COVID-19 vaccines. In: *Proceedings of the AAAI international conference on web and social media (ICWSM)*.

Donzelli G, Palomba G, Federigi I, et al. (2018) Misinformation on vaccination: A quantitative analysis of youtube videos. *Human Vaccines & Immunotherapeutics* 14(7): 1654–1659.

Dutta A, Beriwal N, Van Breugel LM, et al. (2020) YouTube as a source of medical and epidemiological information during COVID-19 pandemic: A cross-sectional study of content across six languages around the globe. *Cureus* 12(6): e8622.

Eysenbach G (2002) Infodemiology: The epidemiology of (mis) information. *The American Journal of Medicine* 113(9): 763–765.

Eysenbach G, Powell J, Kuss O, et al. (2002) Empirical studies assessing the quality of health information for consumers on the world wide web: A systematic review. *Journal of the American Medical Association* 287(20): 2691–2700.

Ferrara E (2020) What types of COVID-19 conspiracies are populated by Twitter bots? *First Monday* 25(6).

Ferrara E, Cresci S and Luceri L (2020) Misinformation, manipulation, and abuse on social media in the era of COVID-19. *Journal of Computational Social Science* 3(2): 271–277.

Ferrara E, Varol O, Davis C, et al. (2016) The rise of social bots. *Communications of the ACM* 59(7): 96–104.

Fung ICH, Fu KW, Chan CH, et al. (2016) Social media's initial reaction to information and misinformation on Ebola, August 2014: Facts and rumors. *Public Health Reports* 131(3): 461–473.

Gallotti R, Valle F, Castaldo N, et al. (2020) Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nature Human Behaviour* 4: 1285–1293.

Goobie GC, Gulera SA, Johannson KA, et al. (2019) YouTube videos as a source of misinformation on idiopathic pulmonary fibrosis. *Annals of the American Thoracic Society* 16(5): 572–579.

Grinberg N, Joseph K, Friedland L, et al. (2019) Fake news on Twitter during the 2016 US presidential election. *Science* 363(6425): 374–378.

Guarino S, Pierri F, Di Giovanni M, et al. (2021) Information disorders during the COVID-19 Infodemic: The case of Italian Facebook. *Online Social Networks and Media* 22: 100124.

Jin F, Wang W, Zhao L, et al. (2014) Misinformation propagation in the age of Twitter. *IEEE Annals of the History of Computing* 47(12): 90–94.

Knuutila A, Herasimenka A, Au H, et al. (2020) *COVID-Related Misinformation on YouTube: The Spread of Misinformation Videos on Social Media and the Effectiveness of Platform Policies*. Oxford: Project on Computational Propaganda.

Kouzy R, Abi Jaoude J, Kraitem A, et al. (2020) Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on Twitter. *Cureus* 12(3): e7255.

Lazer D, Baum M, Benkler Y, et al. (2018) The science of fake news. *Science* 359(6380): 1094–1096.

Li H, Bailey A, Huynh D, et al. (2020) YouTube as a source of information on COVID-19: A pandemic of misinformation? *BMJ Global Health* 5(5): e002604.

Loomba S, de Figueiredo A, Piatek SJ, et al. (2021) Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour* 5(3): 337–348.

Lyu W and Wehby GL (2020) Community use of face masks and COVID-19: Evidence from a natural experiment of state mandates in the US. *Health Affairs* 39(8): 1419–1425.

Mahoney LM, Tang T, Ji K, et al. (2015) The digital distribution of public health news surrounding the human papillomavirus vaccination: A longitudinal infodemiology study. *JMIR Public Health and Surveillance* 1(1): e2.

Memon SA and Carley KM (2020) Characterizing COVID-19 misinformation communities using a novel Twitter dataset. *CEUR Workshop Proceedings* (Vol. 2699).

Mitchell A, Gottfried J, Kiley J, et al. (2014) *Political polarization & media habits*. Available at: http://pewrsr.ch/1vZ9MnM (accessed November 2020).

Mitchell A and Oliphant JB (2020) *Americans immersed in coronavirus news; most think media are doing fairly well covering it*. Available at: https://pewrsr.ch/3dbTpxs (accessed November 2020).

Nightingale S, Faddoul M and Farid H (2020) Quantifying the reach and belief in COVID-19 misinformation. *arXiv:2006.08830*.

Nikolov D, Lalmas M, Flammini A, et al. (2019) Quantifying biases in online information exposure. *Journal of the Association for Information Science and Technology* 70(3): 218–229.

Nizzoli L, Tardelli S, Avvenuti M, et al. (2020) Coordinated behavior on social media in 2019 UK general election. *arXiv:2008.08370*.

Pacheco D, Hui PM, Torres-Lugo C, et al. (2021) Uncovering coordinated networks on social media: Methods and case studies. In: Proceedings of the AAAI international conference on web and social media (ICWSM).

Panatto D, Amicizia D, Arata L, et al. (2018) A comprehensive analysis of Italian web pages mentioning squalene-based influenza vaccine adjuvants reveals a high prevalence of misinformation. *Human Vaccines & Immunotherapeutics* 14(4): 969–977.

Pasquetto IV, Swire-Thompson B, Amazeen B, et al. (2020) Tackling misinformation: What researchers could do with social media data. *Harvard Kennedy School Misinformation Review* 1(8).

Pathak R, Poudel DR, Karmacharya P, et al. (2015) YouTube as a source of information on Ebola virus disease. *North American Journal of Medical Sciences* 7(7): 306.

Pennycook G and Rand DG (2019) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116(7): 2521–2526.

Perrin A and Anderson M (2019) *Share of US adults using social media, including Facebook, is mostly unchanged since 2018*. Available at: https://pewrsr.ch/2VxJuJ3 (accessed February 2021).

Pierri F, Piccardi C and Ceri S (2020a) A multi-layer approach to disinformation detection in US and Italian news spreading on twitter. *EPJ Data Science* 9(35).

Pierri F, Piccardi C and Ceri S (2020b) Topology comparison of Twitter diffusion networks effectively reveals misleading news. *Scientific Reports* 10: 1372.

Pulido CM, Villarejo-Carballido B, Redondo-Sama G, et al. (2020) COVID-19 Infodemic: More retweets for science-based information on coronavirus than for false information. *International Sociology* 35(4): 377–392.

Roozenbeek J, Schneider CR, Dryhurst S, et al. (2020) Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science* 7(10): 201199.

Schaeffer K (2020) *Nearly three-in-ten Americans believe COVID-19 was made in a lab*. Available at: https://pewrsr.ch/2XlJqAa (Accessed November 2020).

Schmidt AL, Zollo F, Scala A, et al. (2018) Polarization of the vaccination debate on Facebook. *Vaccine* 36(25): 3606–3612.

Sell TK, Hosangadi D and Trotochaud M (2020) Misinformation and the us ebola communication crisis: Analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. *BMC Public Health* 20: 1–10.

Seltzer E, Horst-Martz E, Lu M, et al. (2017) Public sentiment and discourse about Zika virus on Instagram. *Public Health* 150: 170–175.

Shao C, Ciampaglia GL, Varol O, et al. (2018) The spread of low-credibility content by social bots. *Nature Communications* 9: 4787.

Sharma K, Ferrara E and Liu Y (2020) Identifying coordinated accounts in disinformation campaigns. *arXiv:2008.11308*.

Sharma M, Yadav K, Yadav N, et al. (2017) Zika virus pandemic–analysis of Facebook as a social media health information platform. *American Journal of Infection Control* 45(3): 301–302.

Singh L, Bansal S, Bode L, et al. (2020a) A first look at COVID-19 information and misinformation sharing on Twitter. *arXiv:2003.13907*.

Singh L, Bode L, Budak C, et al. (2020b) Understanding high-and low-quality URL Sharing on COVID-19 Twitter streams. *Journal of Computational Social Science* 3: 343–366.

Spärck Jones K (1972) A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation* 28(1): 11–21.

Stella M, Ferrara E and De Domenico M (2018) Bots increase exposure to negative and inflammatory content in online social systems. *Proceedings of the National Academy of Sciences* 115(49): 12435–12440.

Vosoughi S, Roy D and Aral S (2018) The spread of true and false news online. *Science* 359(6380): 1146–1151.

Wilson T and Starbird K (2020) Cross-platform disinformation campaigns: Lessons learned and next steps. *Harvard Kennedy School Misinformation Review* 1(1).

Wojcik S and Adam H (2019) *Sizing up Twitter users*. Available at: https://pewrsr.ch/2VUkzj4 (accessed February 2021).

Wood MJ (2018) Propagating and debunking conspiracy theories on Twitter during the 2015–2016 Zika virus outbreak. *Cyberpsychology, Behavior, and Social Networking* 21(8): 485–490.

Yang KC, Pierri F, Hui PM, et al. (2020a) Dataset for paper: The COVID-19 Infodemic: Twitter versus Facebook. *Zenodo*. https://doi.org/10.5281/zenodo.4313903

Yang KC, Torres-Lugo C and Menczer F (2020b) Prevalence of low-credibility information on twitter during the COVID-19 outbreak. In: Proceedings of the ICWSM international workshop on cyber social threats.

Yang KC, Varol O, Hui PM, et al. (2020c) Scalable and generalizable social bot detection through data selection. *Proceedings of the AAAI Conference on Artificial Intelligence* 34(1): 1096–1103.

Zarocostas J (2020) How to fight an infodemic. *The Lancet* 395(10225): 676.