INFORMATION, COMMUNICATION & SOCIETY

Routledge
Taylor & Francis Group

Information, Communication & Society

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rics20

# It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections

Fabio Giglietto, Nicola Righetti, Luca Rossi & Giada Marino

**To cite this article:** Fabio Giglietto, Nicola Righetti, Luca Rossi & Giada Marino (2020) It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections, Information, Communication & Society, 23:6, 867-891, DOI: 10.1080/1369118X.2020.1739732

**To link to this article:** https://doi.org/10.1080/1369118X.2020.1739732

INFORMATION, COMMUNICATION & SOCIETY
2020, VOL. 23, NO. 6, 867–891
https://doi.org/10.1080/1369118X.2020.1739732

# It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections

Fabio Giglietto [a], Nicola Righetti [a], Luca Rossi [b] and Giada Marino [a]

[a]Department of Communication Sciences, Humanities and International Studies, University of Urbino Carlo Bo, Urbino, Italy; [b]Networks Data & Society research group, IT University of Copenhagen, Copenhagen, Denmark

**ABSTRACT**
Over the last few years, a proliferation of attempts to define, understand and fight the spread of problematic information in contemporary media ecosystems emerged. Most of these attempts focus on false content and/or bad actors detection. In this paper, we argue for a wider ecological focus. Using the frame of *media manipulation* and a revised version of the 'coordinated inauthentic behavior' original definition, the paper presents a study based on an unprecedented combination of Facebook data, accessed through the CrowdTangle API, and two datasets of Italian political news stories published in the run-up to the 2018 Italian general election and 2019 European election. By focusing on actors' collective behavior, we identified several networks of pages, groups, and verified public profiles ('entities'), that shared the same political news articles on Facebook within a very short period of time. Some entities in our networks were openly political, while others, despite sharing political content too, deceptively presented themselves as entertainment venues. The proportion of inauthentic entities in a network affects the wideness of the range of news media sources they shared, thus pointing to different strategies and possible motivations. The paper has both theoretical and empirical implications: it frames the concept of 'coordinated inauthentic behavior' in existing literature, introduces a method to detect coordinated link sharing behavior and points out different strategies and methods employed by networks of actors willing to manipulate the media and public opinion.

**ARTICLE HISTORY**
Received 25 October 2019
Accepted 2 March 2020

**KEYWORDS**
Media manipulation; coordinated inauthentic behavior; problematic information; disinformation; Facebook; elections

## Introduction

Citizens' exposure to online disinformation has become a major concern all over the world for a while now. The fear that malicious actors could sow the seeds of discord and distrust among digitally connected citizens, feeding polarization and stirring up insurmountable divisions so as to undermine the democratic process, has filled the unceasing flow of reports and news articles that have been published on the topic (Benkler, 2019). In the

**CONTACT** Fabio Giglietto fabio.giglietto@uniurb.it
This article was originally published with errors, which have now been corrected in the online version. Please see Correction (http://dx.doi.org/10.1080/1369118X.2020.1755558).

aftermath of Brexit referendum in UK and 2016 US Presidential elections, the antagonist online participatory practices of sharing, collaborating and organizing collective actions (Jenkins, 2006; Shirky, 2008), which used to be considered the prerogative of democratizing forces fighting established powers, proved to be just as effective in supporting the spread of extremisms, hate speech, violence and false news (Marwick & Lewis, 2017).

Given the high stakes, researchers, governments and supranational institutions have clearly put a lot of effort into clarifying disinformation-related concepts, unraveling the complex, intertwined dimensions of the phenomenon, studying its empirical manifestations and trying to find solutions (HLEG EU Commission, 2018; Jack, 2017; Wardle & Derakhshan, 2017). Unfortunately, despite all efforts, stopping disinformation has proved harder than expected. A serious obstacle has certainly been the difficulty to mark a clear boundary between problematic and non-problematic information (Jack, 2019; Lazer et al., 2018).

More recently, an emerging approach has suggested to circumvent this obstacle by shifting the focus from content to dynamics of information spreading within online networks (Hui et al., 2019; Keller et al., 2019). Online content, indeed, benefits from a multitude of actors that amplify its reach, with a magnitude proportional to their popularity, the budget they can invest in social media ads, and the capacity to activate the platform algorithms that prioritizes better-performing images, videos, and posts, making popular content spread faster. It follows that 'bad actors' may attempt to coordinate their efforts to get the initial plug which, once detected by the algorithm, may ignite the propagation machine and even attract the attention of mainstream media (Phillips, 2018) on the content they spread for profit or propaganda. Although this is not a new phenomenon (boyd, 2017), during the last few years we observed similar practices applied with the aim of enhancing the spread of political news stories.

Despite the promises of this emerging approach and the fact that a leading social media company as Facebook currently employs 'coordinated inauthentic behavior' (Gleicher, 2018) to enforce its policies, there is a shortage of attempts to substantiate the relationship between coordinated behavior and the sharing of problematic information. To start filling the gap, this present paper frames the ill-defined concept of coordinated inauthentic behavior in the existing scientific literature and tests the related 'action-based' approach to disinformation detection through a combination of Facebook data and two datasets of political news stories published in the six months before the 2018 general election and 2019 European election in Italy.

Italy stands out among other countries (Vaccari, 2013) for having proved to be a fertile ground for the rise of digital parties (Gerbaudo, 2018). In the context of a traditional media partisanship, Facebook is reportedly used to get news by 54% of the population and online sources rival TV as the main source of news (Newman et al., 2019). Evidence of organized social media manipulation has also been observed (Bradshaw & Howard, 2019).

While building on several previous works and researches, we make two main contributions: (a) we present a method for identifying networks of pages and groups that coordinately shared the same news items on Facebook, a phenomenon we call 'coordinated link sharing behavior'; (b) we test whether the coordinated social media activity is associated with the spread of problematic information, i.e., we test whether shifting the focus from online content/actor to patterns of actions represents a fruitful approach to problematic information detection.

The paper is structured as follows: the first section frames the main challenges of problematic information detection, media manipulation and 'coordinated inauthentic behavior' in the existing scientific literature. Based on this literature, research questions are formulated, and the datasets and methods used to address them detailed. Following the data analysis, the limitations of the study are enucleated, and the results discussed to draw general conclusions and provide hints for future research.

## Literature review and research questions

The widely recognized risks of misinformed citizens for healthy democracies brought a cohort of scholars to tackle this issue from a range of different perspectives. The deeply and still undergoing transformations of contemporary media ecologies led to a renewed interest in this topic resulting in a rapidly growing body of interdisciplinary scholarly works published during the last few years. While the topic has seen a recent surge of academic interest, many studies build on the long-existing field of automatic detection of problematic content or actors.

Rather than attempting to provide a systematic review of all these studies, the following paragraphs highlight the essential literature that frames our approach, clarify the terminology used and lead to our research questions.

A first necessary step is to highlight the results and limits of existing approaches to detect bad information and malicious actors. Then we describe the media manipulation frame and the concepts of amplification and problematic information. Finally, we analyze the concept of 'coordinated inauthentic behavior' and pinpoint its potential roots in the existing literature.

## Challenges of problematic information detection

While unanimously recognizing misinformation as detrimental for healthy democracies, the existing literature is fragmented when it comes to defining the object of study. The lack of a shared, consistent and operationalizable definition undermines both the attempt to estimate its prevalence over legitimate information (Lazer et al., 2018) and measure the impact of misinformation on citizens' opinions or behavior (Benkler, 2019; Weeks & Gil de Zúñiga, 2019). While the issue of definition is widely recognized by scholars, the prevalent effect of the countless attempts to formally address it by way of new lexicons and taxonomies (Molina et al., 2019; Silverman, 2017; Wardle & Derakhshan, 2017) seems to have mostly dragged the scientific community deeper into the epistemological rabbit hole of 'fake news' (Caplan et al., 2018; HLEG EU Commission, 2018). To avoid falling in this trap, throughout this article we adopt the umbrella terminology of problematic information (Jack, 2017) to reference the whole spectrum of contents that range from deliberately or mistakenly false news to propaganda, gaslighting and satire.

Even when avoiding to differentiate the phenomenon based on the motivations of who creates and distributes problematic information, the simple basic choice of flagging the content as true or false is not always possible or advisable (Giglietto et al., 2019; Marwick, 2018). On the one hand, drawing such a clear distinction requires a significant amount of time, skills and resources for each content. On the other hand, due to the lack of commonly accepted definitions, making such calls is a great responsibility that deeply affects the outcomes of the study. For all these reasons, a large body of studies tend to delegate this crucial process to established external bodies (e.g., list of false content flagged by fact-checkers) (Allcott et al., 2019; Fletcher & Nielsen, 2017; Guess et al., 2019) or adopting narrow definitions (Khan et al., 2019) or features based detections (Reis et al., 2019) that are strict or vague enough to be operationalized in an algorithm (in a certain sense another form of delegation).

Both approaches come however with their own well-known limits. Due to the amount of work required to fact-check a single content, the quota of content flagged as false tends to be a small fraction of the overall false content circulating online. The algorithmic approach has also its shortcomings due to the need for narrowing down the definition enough to make it possible to operationalize the concept and minimize false positives. In both cases, there is a high risk of underestimating the real prevalence of existing false content.[1]

Parallel to the attempts to detect problematic content, it should be mentioned the stream of research that has tried to identify problematic and automated users often observing a strong correlation between problematic content and forms of automation (Badawy et al., 2018; Davis et al., 2016). The idea is that a user/page/account/news outlet that repeatedly published or shared content that was flagged as false, can be deemed as problematic as a whole. This goal has been pursued following different strategies. The most straightforward is using blacklists provided by professional fact-checkers who manually assess the content and the nature of online actors. While widely used, this approach can hardly catch up to the ever-changing strategies of malicious actors (Bastos & Mercea, 2018) and thus carries some risks of biased estimates.

Beside using blacklists, malicious actors may also be automatically detected (Shu et al., 2019). Attempts to automatic detection can be grouped in two streams: on the one side researchers have tried to use specific features of the online accounts (such as the account creation date, activity patterns, type of messages produced) and train machine learning models to identify the inauthentic accounts (Yang et al., 2019). Varol and colleagues (2017) recently provided a good overview of the existing methods and their performance. While offering better results than automatic means to identify false content, bot detection is also far from perfect (Morstatter et al., 2016) and better detection techniques can simply lead to more sophisticated bots that may be difficult and sometimes impossible to detect (Luceri et al., 2019).

A second[2] stream is based on the assumption that automated users will develop an online network aimed at maximizing its effectiveness (e.g., being frequently connected with other automated accounts) rather than fulfilling some social need (Alvisi et al., 2013; Boshmaf et al., 2011). Despite a wider range of applications ranging from online reviews (Ye & Akoglu, 2015) to Twitter (Freitas et al., 2015), this approach is often hindered by the limits posed by platforms API to network data collection (Coscia & Rossi, 2018).

### Media manipulation and false content

To make things even more tricky, creating and distributing false content only represents the tip of the iceberg of the strategies employed by malicious actors to manipulate social media, mainstream media and the public debate (Marwick & Lewis, 2017). Sometimes, even a suitably titled legitimate news stories, opportunely and artificially amplified, can be weaponized to skew the public narrative around certain issues. Both false and real content benefit from a multitude of actors that amplify (intentionally or not) its reach. Depending on the popularity of each actor in the network and the budget they can invest in social media ads, the magnitude of this amplification may change drastically. Furthermore, popular content tends to spread faster on social media due to the effect of algorithms that prioritize better-performing links, images, videos, and posts. These performances depend on an estimate of popularity based on the analysis of quantified attention metrics provided by each platform (likes, reactions, views, shares, etc). Beside the effect of this 'rich will get richer' feedback loop, popular social media content and highly discussed topics are often featured in traditional media, thus benefiting from a significant further spin. The centrality of these metrics offers big rewards to those interested in increasing the visibility of certain content (Zhang et al., 2017). For these reasons, different actors may attempt to coordinate their efforts to get the initial spin which, once detected by the algorithm, may ignite the propagation machine and even attract the attention of mainstream media (Phillips, 2018). This is not at all a new phenomenon. Fans' attempts to coordinate their behavior to push certain hashtags into Twitter trending topics date back to 2011 at least (boyd, 2017). During the last few years, we observed similar practices applied with the aim of enhancing the spread of political news stories. The practice of *hacking the attention economy* can be driven by a range of motivations from ideology to commercial, to gain status or attention or simply for fun (Marwick & Lewis, 2017; Zhang et al., 2013).

Similar campaigns can be identified by looking at the veracity of content or actors involved. For this reason, in this paper we argue for a broader ecological approach that primarily takes into account the collective behavior of malicious actors. While it is in fact perfectly possible that problematic content is published and distributed without any form of attempts of amplifying its reach, it is highly probable that the spread of harmful content is supported by these operations. Detecting the coordinated attempts of multiple actors to increase the visibility of certain content may thus lead to identify networks of potentially inauthentic actors aimed at amplifying problematic content. Following a terminology introduced by Facebook, we describe this type of operation as 'Coordinated Inauthentic behavior'. 'Coordinated Inauthentic Behavior' has been defined by Nathaniel Gleicher, Head of Cybersecurity Policy of Facebook as a case when 'groups of pages or people work together to mislead others about who they are or what they are doing' (Gleicher, 2018). By shifting the attention to deceptive behaviors, the definition deliberately avoids falling into the trap of judging the truthfulness of content: 'The posts themselves may not be false'. Gleicher also provides this example: 'We may take a network down for making it look like it's being run from one part of the world when in fact it's being run from another. This could be done for ideological purposes or can be financially motivated.' Beside the operations undertaken by foreign or local governments, the policy also applies to 'non-state actors, domestic groups and commercial companies' (Gleicher, 2019).

In other terms, the definition comprises the concept of coordination and inauthenticity. Both concepts have been widely studied, albeit rarely in conjunction. In the next paragraphs, we summarize these studies with the aim of grounding the definition of coordinated inauthentic behavior in the existing literature.

### *Coordination*

Coordination can be defined as the act of making people and/or things involved in organized cooperation. Several authors argued that it is a distinctive mark of users' participation within online spaces (Bruns et al., 2013; Jenkins, 2006; Shirky, 2008). Such coordination plays a key role in the online participatory culture described by Henry Jenkins in 'Convergence Culture' (2008). Online fandom, for instance, proved to be capable to organize collective actions with different purposes, as inflate social media attention metrics (likes, retweets, etc.) on a specific topic or to influence the plot of a narrative or the trade of an item.

Online activism benefited from the opportunity of building online communities and coordinating their collective actions allowed by the Internet (Bennett & Segerberg, 2012). While most of the early accounts and scholarly work focuses on the beneficial outcomes of digital mediated forms of collaboration as they empower protest movements to fight established and sometimes oppressive powers (Coleman, 2015; Freelon et al., 2018; Loader & Mercea, 2011), the same infrastructure and organization techniques can be employed by a range of diversely motivated malicious actors (Jenkins et al., 2015; Marwick & Lewis, 2017).

### *Authenticity*

The concept of authenticity is ultimately connected with the identification of malicious or fake actors on the network, that we have reported in the previous paragraphs. Nevertheless, the concept has acquired new relevance in the context of political discussion on social media (Salisbury & Pooley, 2017). Here malicious actors use whatever websites and social media opportunity to propagate ideas while hiding their real identities and intentions (Bastos & Farkas, 2019; Daniels, 2009; Donovan & Friedberg, 2019). Several scholars provided a range of examples of these activities, from anti-abortion sites masked under the pro-choice tag (Daniels, 2014) to false Islamist Facebook pages spreading anti-Muslims content (Farkas et al., 2018).

Besides such 'cloaked websites' (Daniels, 2009), a well-known type of inauthentic online behavior is that of bots and fake accounts (Woolley & Howard, 2016). Bots are widely employed to manipulate online political discussion and boost politicians' followers to generate false impressions of popularity (Bastos & Mercea, 2017; Bessi & Ferrara, 2016; Ratkiewicz et al., 2011; Woolley & Howard, 2016). Paid users are also employed to impersonate fake social media accounts to undermine online public discourse and distract the public from controversial issues (King et al., 2017).

In the seminal work 'The people's choice', Lazarsfeld and colleagues (1944) inquired the role played by personal influence (exposure to casual conversations about politics as opposed to the role played by mass media) on the formation of political opinions, finding that personal influence, compared with traditional media, is able to reach more frequently undecided voters and catch the audience less prepared against influence. Given the effect of accidental exposure to political content on social media on online participation (Valeriani & Vaccari, 2016), malicious social media entities aimed at influencing political opinion may have strong incentives to do so without revealing their authentic motivation and identity. Furthermore, leveraging the Internet affordance to gather together people based on personal interests (Ito et al., 2010), it is much easier to build a large follower base by presenting the entity, in order to appeal to a wider audience, as dedicated to entertainment or popular culture than politics. Once the follower base is established, the pages and groups can be used to convey political content to a largely unguarded audience.

### *Research questions*

Building on the aforementioned approaches to detect problematic and inauthentic online behavior, companies such as Facebook constantly improve their strategies and have recently focused increasing attention to the aspect of inauthentic coordination (Gleicher, 2018). Nevertheless, there is a shortage of scholarly evidence on the effectiveness of this approach in terms of surfacing malicious actors and problematic information. To address this gap, we put the idea to test it by analyzing Facebook shares of political news stories published in the run up of two Italian elections. Using an original method described in the next section, we detected several networks of coordinated and inauthentic actors that cooperated to boost certain political news stories in the lead up of both 2018 and 2019 elections. We thus formulated the following research questions:

> RQ1: Did these coordinated networks share problematic content in the months preceding 2018 and 2019 Italian elections?

The evidence available in the literature clearly describes a range of motivations pushing actors to coordinate their activities to artificially boost the popularity of certain online content. Given this different range of motivations, we expect that networks entirely composed by openly political entities (pages, groups and verified profiles belonging to political actors and/or presenting themselves as a venue to get information and discuss politics) and networks also composed, instead, by inauthentic entities that shared political contents under a misleading non-political identity, would differ in terms of typology of content shared and structure of the network. We thus formulated the following research questions:

> RQ2a: Did political and non-political coordinated networks employ different link sharing strategies?

Considering that existing research on online coordinated information spreading (Del Vicario et al., 2016) has suggested that specific network configurations might be more effective, and thus preferable to achieve broader dissemination of content, we finally settled on analyzing the structure of the coordinated networks. Based on these studies it was asked:

> RQ2b: Are there significant structural differences between political and non-political coordinated networks?

## Data and methods

The analyses presented in this paper are based on two datasets of online Italian political news stories shared on Facebook during the six months preceding the 2018 Italian general election

($N$ = 84,815) and the 2019 European election ($N$ = 164,760). News items were collected from Google News, the Global Database of Society (GDELT) and Twitter Streaming API (filtering for tweets including a link and mention of a leader or a political party).

CrowdTangle API link endpoint (CrowdTangle Team, 2019) was used to collect public Facebook/Instagram shares of the news stories URLs in our datasets performed in a period of seven days after the publication of each piece of news. CrowdTangle is a Facebook owned analytic platform that tracks posts published by 3.M+ ‘influential’ Facebook pages, verified public profiles and public groups. While CrowdTangle coverage of all Facebook public entities is thus incomplete by design, it tracks – according to the latest official figures (Fraser, 2020) – almost the totality of Facebook Pages with over 100,000 Likes/Followers.[3] The resulting datasets consisted of 107,842 shares performed by 6,215 unique entities (2018 election dataset) and 222,877 shares performed by 8,148 unique entities (2019 election).

Detecting the networks of coordinated entities is a two steps process (Figure 1). First, the algorithm[4] estimates a time threshold for identifying all the news items shared near simultaneously by different entities in a short period of time. Subsequently, the coordinated networks are identified by grouping just the entities that repeatedly shared the same news story near simultaneously.

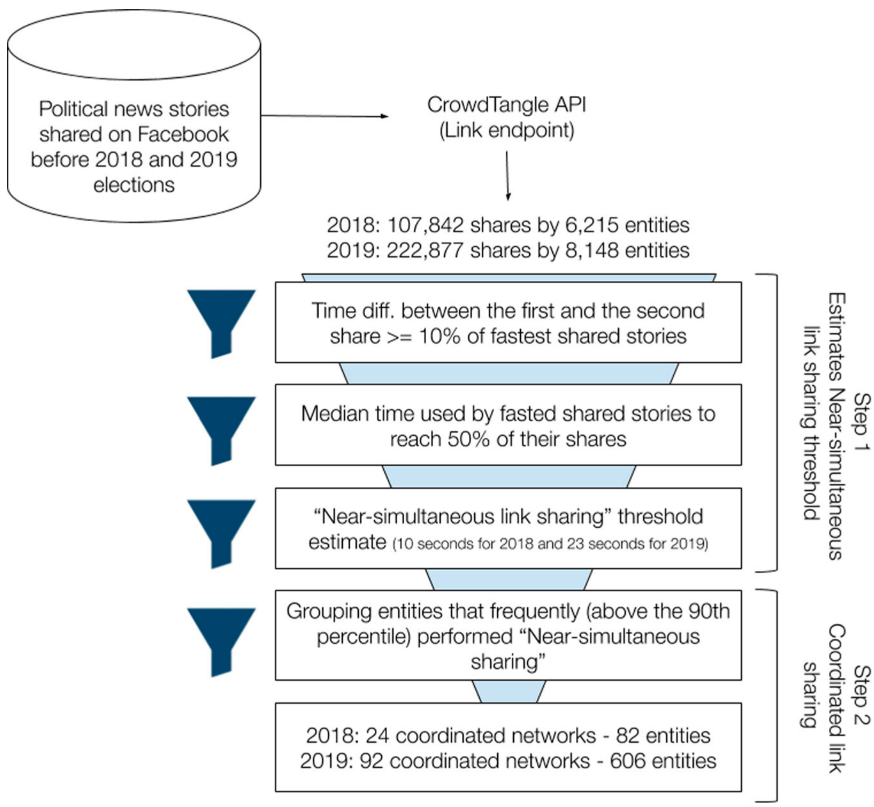


**Figure 1.** Funnel of coordinated link sharing.

While it is ordinary that several entities share the same URLs, it is unlikely that this occurs within a very short time span and repeatedly. Such rapidity and regularity in sharing news items can be a signal of coordinated activity.

The idea is to operationalize, as a first step, the concept of 'near-simultaneous sharing' by finding an appropriate time threshold. Given a CrowdTangle dataset of URLs shares, this threshold is estimated by analyzing the time differences between each share of the same URL ranked by date (i.e., the date-time when the links were shared) to identify a subset consisting of 10% URLs with the shortest time span between the first and second share. We then identified the desired threshold by calculating the median time in seconds used by 10% of the quickest URLs to reach 50% of their total number of shares, assuming that networks aimed at spreading news items are likely to be closely associated with the news sources they spread. We used this threshold to identify a list of entities that performed 'near-simultaneous link sharing'.

Since a regular pattern of activity is a significant signal of an organized structure, as a second step, we derived the entities' networks frequently (above the 90th percentile or more than 4 times for 2018 and more than 3 times for 2019) sharing news links in a coordinated way from the list of entities resulting from the previous step.

By using this method, a total of 24 and 92 strongly coordinated networks that spread political news before the 2018 and 2019 elections, respectively, were identified. The 2018 networks were composed of 82 entities, while the 2019 networks of 606 entities. Given the conservative approach used in estimating the 'near-simultaneous shares' threshold, the entities listed should be considered as the core of potentially larger networks.

The analyses focused on the news stories shared by these highly coordinated networks within a very short time from each other, that is 2,213 news items shared in the 2018 election dataset, and 5,863 in the 2019 election dataset, also comparing them with the news items shared by the non-coordinated entities in both years, which were 38,233 in the 2018 dataset and 66,810 in the 2019 dataset. There is a small overlap in the news items shared by coordinated and non-coordinated entities: in the 2018 election dataset, about 3% of all the news items shared by non-coordinated pages and groups were shared also by coordinated ones, while in the 2019 election dataset the overlap amounts to about 6% (Table 1).

To answer the first research question about whether the coordinated networks identified through the method described above actually spread problematic content online, we checked the domains they shared against blacklists of already identified sources of 'fake' and hyperpartisan news. The list of problematic websites was retrieved from Italian established debunking websites (butac.it, bufale.net, bufalopedia.it and pagellapolitica.it). The first three sources have been previously used to identify and measure Italian problematic information (Fletcher et al., 2018), while pagellapolitica.it is Facebook's Italian

**Table 1.** Descriptive statistics relative to the 2018 and 2019 election datasets.

| | Coordinated | | | | | non-coordinated | | | |
|---|---|---|---|---|---|---|---|---|---|
| Year | Networks | Entities | Median Subscribers | News Items | Median Engagement | Entities | Median Subscribers | News Items | Median Engagement |
| 2018 | 24 | 82 | 32,825 | 2,213 | 581 | 6,133 | 3,205 | 38,233 | 87 |
| 2019 | 92 | 606 | 5,630 | 5,863 | 663 | 7,542 | 3,231 | 66,810 | 62 |

official fact-checking partner. Merging these blacklists, we ended up with a list of 376 problematic news domains (see Appendix 1). Moreover, we checked the coordinated entities against a list of 87 Facebook pages already pointed out as sources of problematic information by the nonprofit organization Avaaz (Di Benedetto Montaccini, 2019; Mastinu, 2019), whose investigative reports have already spotlighted numerous disinformation networks that Facebook subsequently took down (see Appendix 2). Risk Ratio (RR)[5] was used to calculate statistically significant differences in proportions of problematic domains shared by coordinated and non-coordinated networks, as well as in the prevalence of problematic Facebook entities.

Concerning the differences in terms of link sharing strategies (RQ2), the analysis focused on the degree of politicalness of the coordinated networks and the variety of sources they shared. First, we performed a qualitative analysis of the description, cover photo and profile picture of each coordinated entities, so as to characterize their self-presentation strategy and classify them as 'political', if any open reference to a political actor or a particular public issue was present, or 'non-political', when their self-presentation did not include any of these references. Then, a measure of self-declared politicalness ranging from 0 to 1 was computed for each network based on the proportion of openly political entities over the total entities of a network. Afterwards, it was measured how large or narrow was the set of domains shared by each coordinated network. To this end, it was computed the Gini coefficient on the proportions of unique domains they shared.[6] The Spearman's rank correlation coefficient[7] was then used to calculate the correlation between the politicalness and sharing strategies as measured through the Gini coefficient.

In terms of structure of the coordinated networks (RQ3), the analysis was focused on examining coordinated networks to investigate if the political nature of the networks (politicalness) or their editorial strategy, measured through the Gini coefficient, is more frequently associated with a specific degree centralization and/or clustering coefficient. Degree centralization is a metric of degree distribution concentration (Butts, 2006; Wasserman & Faust, 1994) and it has been observed as a measure for authoritarian structures where the opinion of a central node is imposed to and shared with external satellites (Sicilia et al., 2006). We used this metric to measure how much the observed co-sharing network was structured like a star-like network with a clear center of origin. Clustering coefficient (Watts & Strogatz, 1998) measures the degree to which nodes in a network tend to cluster together forming triangles and it has been often associated with the presence of strong community structures (Girvan & Newman, 2002).

Given the nature of the chosen metrics, the analysis performed to address the third research question could only be performed on networks counting more than two entities. This meant that 73 networks composed of only two nodes were removed. Figure 2 shows the density functions of the size of the networks showing how dyads of only two nodes were, by far, the most common size.

## Findings

Statistically significant relations emerged between coordinated activity and the problematicity of both the domains and the Facebook entities (pages or groups) that shared their stories.

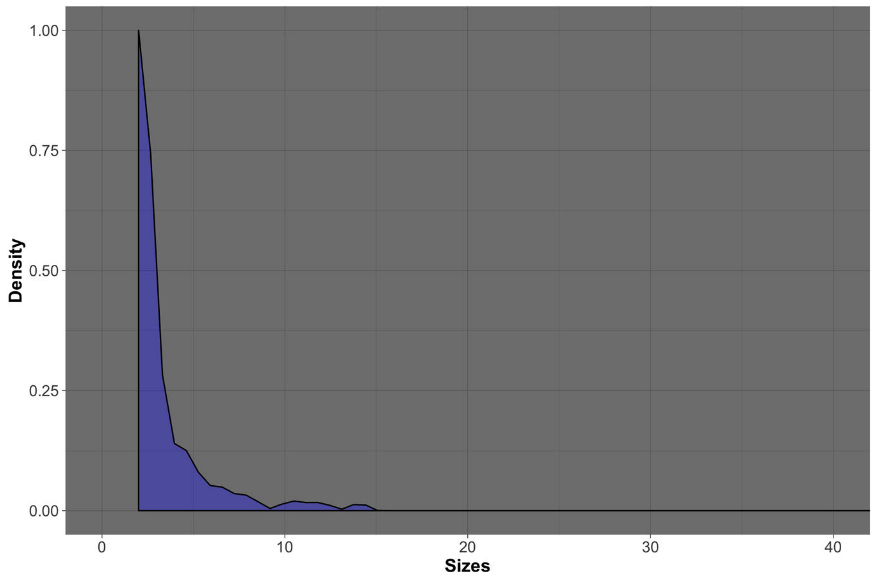


**Figure 2.** Density functions of networks size.

In the 2018 election dataset, problematic domains were published significantly more frequently by the coordinated (5.62%) than non-coordinated entities (3.15%), and the same relation emerged before the 2019 elections when the problematic domains were shared in a coordinated way more frequently (3.87%) than without coordination (1.74%). Problematic domains are 1.79 times (95% CI [1.08, 2.96]) and 2.22 times (95% CI [1.35, 3.67]) more likely to be shared by coordinated than non-coordinated entities, which is a statistically significant difference (Figure 3).

Our lists of coordinated entities also significantly overlap with the Avaaz list of Italian problematic Facebook pages. The 2018 coordinated pages and groups occurred in the list of problematic Facebook pages much more frequently (11%) than the non-coordinated ones (1%). The same emerged in the 2019 election dataset, where the number of

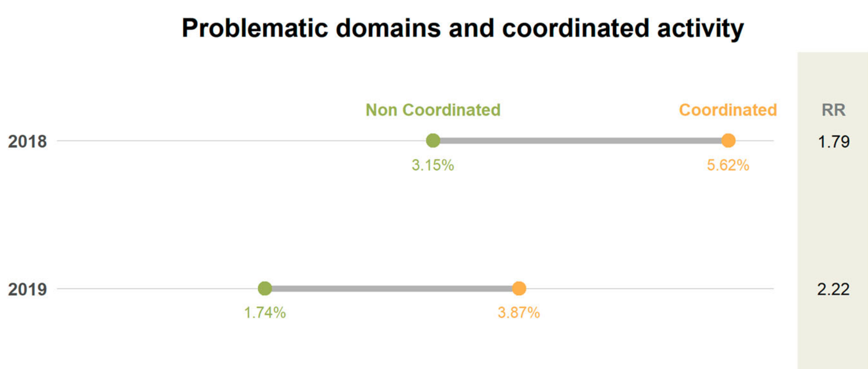


**Figure 3.** Proportion of problematic domains shared by coordinated and non-coordinated entities. The panel on the right displays the risk ratio (RR) values, all statistically significant.

coordinated entities included in the list of problematic pages was larger (6.8%) than that of the non-coordinated ones (0.3%). The problematic Facebook pages in the list are 19.24 times (95% CI [9.56, 38.72]) and 23.19 times (95% CI [13.91, 38.68]) more likely to be found in the coordinated than non-coordinated entities, which is a statistically significant difference (Figure 4).

Based on the abovementioned evidence it was concluded that coordinated entities shared problematic information before the 2018 and 2019 elections in Italy.

As expected, (RQ2), the qualitative inspection of the entities' facade revealed a certain degree of deception. Although all of the pages in the datasets shared political news stories, some of them did not disclose their political nature but, on the contrary, conceal it under the appearance of venues exclusively devoted to entertainment, soft news stories or gossip (Table 2).

Besides the issue of deception, the percentage of political entities in a network was also associated with different strategies in terms of the variety of news domains a coordinated network shared. Considering the coordinated networks that spread news items before both the 2018 and 2019 elections in Italy, a strong relation emerged between the self-declared politicalness of a network and the domains sharing strategies. Indeed, a Spearman correlation found that the more explicit the politicalness of a network, the lower the shares concentration around a few domains, both in the 2018 election dataset, $r_s = -.76$ ($N = 24$, $p < 0.001$), and in the 2019 election dataset, $r_s = -.63$ ($N = 92$, $p < 0.001$).

The analysis of the network structures associated with the coordinated networks (RQ3) revealed a tendency of the networks to assume either one or the other of the two ideal configurations we have identified: highly clustered or highly centralized networks. Figure 5 represents the density functions of the two metrics measured on the networks. It shows

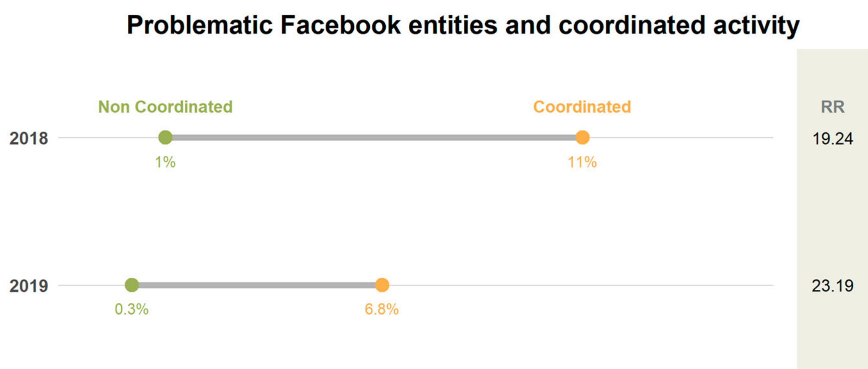


**Figure 4.** Proportion of problematic entities included in the coordinated and non-coordinated entities. The panel on the right displays the risk ratio (RR) values, all statistically significant.

**Table 2.** Percentage of political, non-political and mixed coordinated networks based on their self-description.

| Year | Political | Non-Political | Mixed |
|---|---|---|---|
| 2018 | 44% | 27% | 29% |
| 2019 | 17% | 19% | 64% |

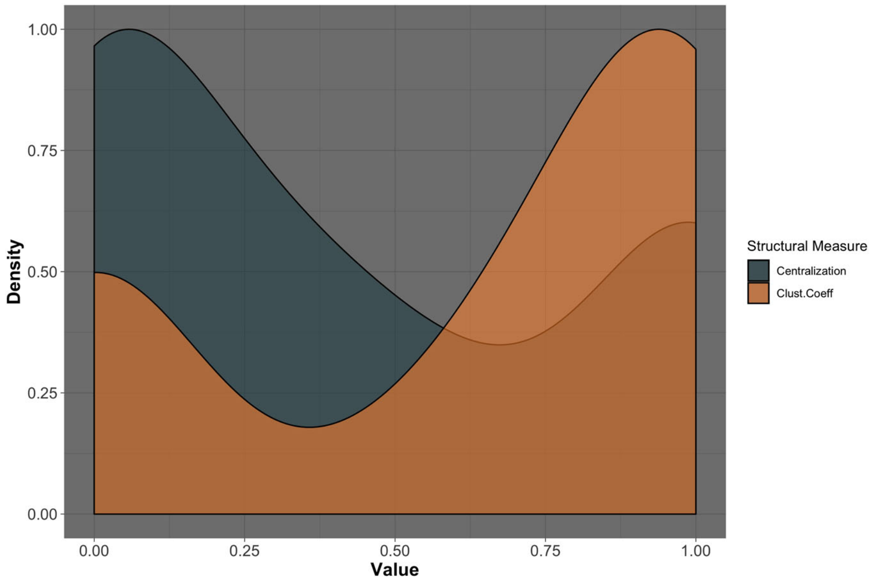


**Figure 5.** Density functions of the Centralization and Clustering Coefficient measured on the strongly coordinated networks.

how the networks seem to be either organized in one way or another, thus clustering into two groups: one dominated by a centralized structure and one dominated by a clustered structure. Obviously, a network cannot have both high clustering coefficient and high degree centralization, but a majority of values either on a single side or in the middle ground was entirely possible and it has not been observed.

Figure 6 shows strongly coordinated networks detected in 2018 and 2019 plotted according to their values of centralization and clustering coefficient. It can be observed that networks present various types of structures with the tendency for the networks to assume either one of the two 'ideal' structures (highly clustered or highly centralized).

We have then explored if there were a correlation between the two structures we identified and specific levels of politicalness or editorial strategy (measured through the Gini coefficient), without finding any significant relation. Figure 7 shows the level of politicalness and the Gini coefficient for each strongly coordinated network plotting according to their clustering coefficient and level of centralization. Inspecting Figure 7, emerged that there was no relation between the structure of the strongly coordinated networks and their politicalness or their Gini index. We observed networks dominated by political pages both with a highly centralized structure and highly clustered structure. Similarly, we found highly centralized pages with extremely high Gini index and as well as highly clustered pages. Thus, the observed dichotomy of the adopted structure is a finding that requires future works to be fully explained.

## Limitations

The algorithm used to detect the 'coordinated link sharing behavior' proved useful to surface subsets with the highest concentration of problematic content and actors across the

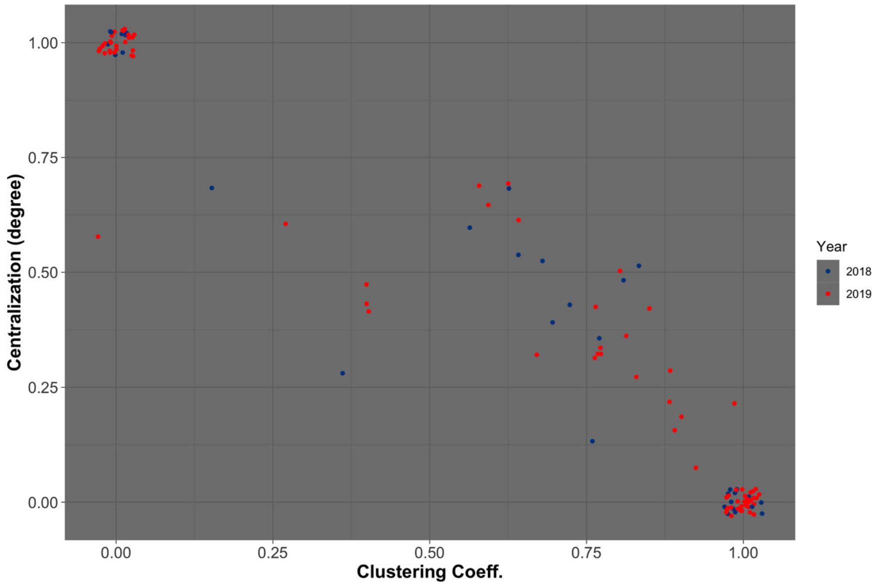


**Figure 6.** Centralization and Clustering coefficient of strongly coordinated networks (to avoid node overlapping a jitter function added minor noise to the plotted data).

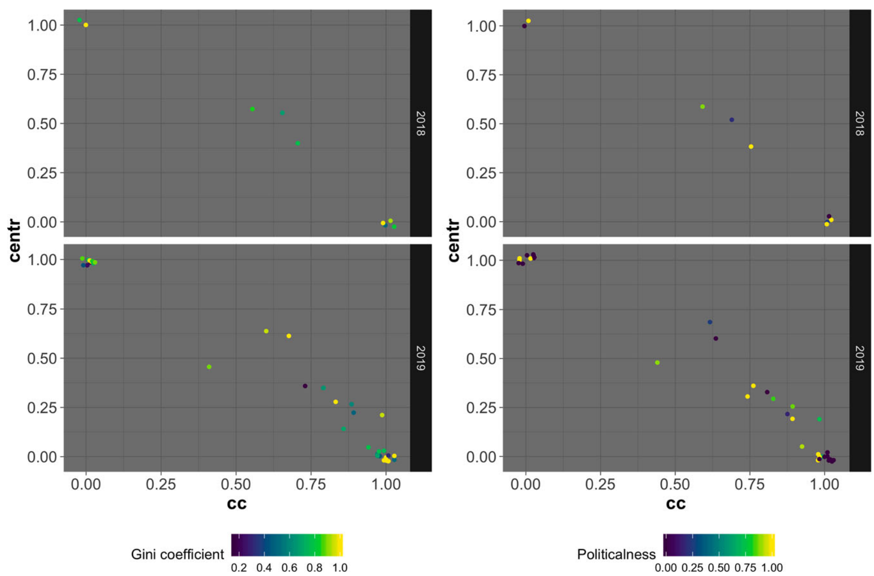


**Figure 7.** Gini Coefficient and Politicalness for strongly coordinated networks plotted according to their level of centralization and clustering coefficient.

two different datasets of CrowdTangle shares. However, additional tests are needed on a wider range of different datasets to fine tune the algorithm. At the same time, while we tried to carefully avoid arbitrary choices when setting time and edges filters by linking these thresholds to the distributions, a certain amount of arbitrariness proved to be unavoidable.

Furthermore, given that CrowdTangle focuses on 'influential' Facebook public entities, the content and actors surfaced by looking at 'coordinated link sharing behavior' will hardly be comprehensive. On the one hand, this limit hinders detecting operations undergoing the audience building phase. On the other hand, it prevents to observe operations where the coordinated activity is performed at the private user's accounts level.

Entities removed by Facebook as the result of a violation of their policies, disappear from CrowdTangle as well. Given the focus on the analysis of potentially malicious actors, we cannot exclude the presence of additional entities or entire coordinated networks at work during both 2018 and 2019 elections. Under this perspective, a public database of the removed entities and the URLs they shared, similar to the one maintained by Twitter (2019), would be helpful for future studies.

## Discussion and conclusion

Given the widely recognized risks posed to democracy by information operations aimed at manipulating the public debate through social media, a wide range of studies attempted to define the phenomenon estimates its prevalence and effects. However, the complexity of the challenge and the wide variety of motivations and strategies adopted by different malicious actors undermined the attempts to establish a sufficiently shared terminology required to build reliable measures of prevalence and impact. Given this lack of common ground, both content-based and actor-based approaches seem unable to provide compelling answers to the challenges at stake.

In this paper, we introduce an additional approach that focuses on the collective behavior of the actors. The contribution is twofold. On the one hand, we frame the concept of 'coordinated inauthentic behavior' in the existing literature on coordination and authenticity. On the other, we assess the reliability of the approach by detecting and analyzing networks of coordinated Facebook entities that boosted political news stories in the lead up of 2018 and 2019 Italian elections.

By analyzing over three hundred thousand Facebook shares of thousands political news stories, we identified hundreds of networks of coordinated entities that cooperated to boost a wide variety of sources under a wide variety of political and non-political identities. As it might be expected, coordinated behavior is enacted by heterogeneity of actors, this includes the Facebook pages of media actors (e.g., newspapers and media companies) as well as pages that exist only on Facebook. As discussed, while some of this activity clearly aims at monetization (Newman et al., 2013, 2019), the analysis of the relation between coordinated activity and authenticity highlights the blurred boundaries between economic and ideological motivations.

Entities feeding political content to their subscribers while hiding their identity and intention are particularly distressing. While the scholarly debate about the role played by social media in fostering more selective or cross-cutting exposure flourished, our understanding of the prevalence and effect of this form of casual exposure requires more work (especially concerning cases where this *casual* exposure is orchestrated by malicious actors).

The Facebook coordinated entities we have detected (as well as the news outlet they share) tend to appear with a frequency well above other news outlets and entities in blacklists compiled by Italian fact-checkers. Furthermore, we show that networks predominantly composed by political entities tend to share a wider variety of news outlets than networks that include entities with deceptive non-political identities. Certain networks only share one specific domain and this domain is often problematic. In other terms, while political networks tend to share news stories from different sources as long as they support their worldview and even sometimes to shame the alternative ones, the entire existence of certain non-political networks is devoted to boost specific news outlets. While the first group are primarily ideologically motivated, the second are mainly commercially motivated. Given the key role played by different motivations to make sense of the spectrum of problematic information (Jack, 2017), this finding seems to be particularly promising.

Thanks to the comparative perspective offered by two subsequent elections, we also observed several differences that may depend on changes in the strategies adopted by these networks or being the effect of the new policies enforced by Facebook before the EU Parliamentarian 2019 election (Woodford, 2019). The ever-changing policies of social media platforms combined with the as much changing adversarial strategies conceived by malicious actors and their networks, pose serious challenges to those who intend to study this phenomenon. At the same time, while we used blacklists to assess the presence of previously known problematic outlets and entities, the comparison between 2018 and 2019 clearly points out that new outlets and entities keep substituting old ones making static blacklists partially ineffective. This finding calls for future studies to employ coordinated behavior to keep this list more easily updated.

The analysis of the structural properties of the strongly coordinated networks produced mixed results. On the one side, the structural properties that have been identified, centralization and level of clustering, appear to be relevant since networks seem to assume one of the two configurations associated with those properties. Nevertheless, our attempt to explain the structures using as explanatory variable the level of politicalness or the Gini index of the networks resulted inconclusive leaving the explanation of the observed duality in structures as a goal for further research.

## Notes

1. Training the algorithm with content flagged as false by professional fact-checker (supervised machine learning) sounds like a promising compromise. However, even computationally and financially resourceful companies such as Google, Facebook or Twitter are still experimenting with this approach when it comes to misinformation.
2. Even if we introduce this stream of research as second, it is probably fair to say that, from a chronological point of view, it can easily be traced back to research that is older than those focused on actor identification.
3. Please see https://help.crowdtangle.com/en/articles/1140930-what-is-crowdtangle-tracking for an overview of what CrowdTangle is tracking. For this study, only Facebook and Instagram platforms have been used.
4. The algorithm is developed in R and the code is available at https://github.com/fabiogiglietto/CooRnet.
5. The Risk Ratio (or Relative Risk) is the ratio between the proportion of occurrence in groups exposed and non-exposed to a risk variable (here the coordinated/non-coordinated link sharing behavior). The value is statistically significant when its 95% confidence interval (CI) does not include 1.
6. The Gini coefficient is a measure of the degree of concentration (inequality) of a variable in a distribution. It ranges between 0 and 1: the more nearly equal the distribution, the lower its Gini index.

7. Spearman's correlation is a measure of the strength and direction of a monotonic relationship between two variables. It ranges between 1 (perfect association) and −1 (perfect negative association), whereas a value of 0 indicates no association between variables.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by Social Science Research Council [Grant number: SSRC-030].

## Notes on contributors

*Fabio Giglietto*, is Associate Professor at the Department of Communication Sciences, Humanities and International Studies at the University of Urbino Carlo Bo. His main research interests are theory of information, communication, and society with a specific focus on the relationship between social systems and digital technologies. On these topics, he has published extensively in journals such as the Journal of Communication, Information, Communication and Society, the Journal of Broadcasting and Electronic Media, Social Media + Society, and the International Journal of Communication.

*Nicola Righetti*, Ph.D., is a postdoctoral research fellow at the Department of Communication Sciences, Humanities and International Studies at the University of Urbino Carlo Bo. His research interests are in the field of sociology of culture and communication, with a focus on digital sociology, digital methods, data and text mining.

*Luca Rossi*, is Associate Professor in the Department of Digital Design at the IT University of Copenhagen. He is active in the field of digital methods for social sciences. His research is interdisciplinary trying to connect traditional sociological approaches with computational approaches. He is working on extending social network analysis techniques for social media analysis, on new approaches for unstructured communities detection, and mapping based on the study of multiplex networks.

*Giada Marino*, Ph.D. She is mainly interested in authenticity and ephemerality topics, with specific reference to social media affordances and user behaviors. She has contributed as junior research assistant to several academic research projects, such as Mapping Italian News Media Political Coverage in the Lead-up of 2018 General Election.

## ORCID

*Fabio Giglietto* http://orcid.org/0000-0001-8019-1035
*Nicola Righetti* http://orcid.org/0000-0002-9257-5113
*Luca Rossi* http://orcid.org/0000-0002-3629-2039

## References

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*(2), https://doi.org/10.1177/2053168019848554

Alvisi, L., Clement, A., Epasto, A., Lattanzi, S., & Panconesi, A. (2013). *Sok: The Evolution of Sybil Defense via social networks*. 2013 IEEE Symposium on Security and Privacy, Berkeley, CA, 382–396. https://doi.org/10.1109/SP.2013.33

Badawy, A., Ferrara, E., & Lerman, K. (2018). *Analyzing the digital traces of political manipulation: The 2016 Russian Interference Twitter Campaign*. 2018 IEEE/ACM International Conference on Advances in social networks analysis and Mining (ASONAM), August 28–31, 258–265.

Bastos, M., & Farkas, J. (2019). "Donald Trump is my President!": The Internet research agency propaganda machine. *Social Media + Society*, *5*(3), https://doi.org/10.1177/2056305119865466.

Bastos, M., & Mercea, D. (2017). The Brexit Botnet and user-generated hyperpartisan news. *Social Science Computer Review*, *37*(1), 38–54. https://doi.org/10.1177/0894439317734157.

Bastos, M., & Mercea, D. (2018). The public accountability of social platforms: Lessons from a study on bots and trolls in the Brexit campaign. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *376*(2128), 2128. https://doi.org/10.1098/rsta.2018.0003

Benkler, Y. (2019). *Cautionary Notes on disinformation and the origins of distrust*. Social Science Research Council. https://doi.org/10.35650/MD.2004.d.2019

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. *Information, Communication and Society*, *15*(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11), https://doi.org/10.5210/fm.v21i11.7090

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). *The socialbot network: When bots socialize for fame and money*. Proceedings of the 27th Annual Computer Security applications Conference, Orlando, Florida, 5–9 December, 93–102.

boyd, d. (2017). Hacking the attention economy. *Data and Society: Points*. Https://points. Datasociety. Net/hacking-the-Attention-Economy-9fa1daca7a37. https://points.datasociety.net/hacking-the-attention-economy-9fa1daca7a37

Bradshaw, S., & Howard, P. (2019). *The global disinformation order: 2019 global inventory of organised social media manipulation (Vol. 31)*. Oxford Internet Institute. https://comprop.oii.ox.ac.uk/research/cybertroops2019/

Bruns, A., Highfield, T., & Burgess, J. (2013). The Arab Spring and social media audiences: English and Arabic Twitter users and their networks. *The American Behavioral Scientist*, *57*(7), 871–898. https://doi.org/10.1177/0002764213479374

Butts, C. T. (2006). Exact bounds for degree centralization. *Social Networks*, *28*(4), 283–296. https://doi.org/10.1016/j.socnet.2005.07.003

Caplan, R., Hanson, L., & Donovan, J. (2018). *Dead Reckoning navigating content moderation after "fake news"*. Data&Society.

Coleman, E. G. (2015). *Hacker, hoaxer, whistleblower, spy: The many faces of anonymous*. Verso.

Coscia, M., & Rossi, L. (2018). *Benchmarking API Costs of network sampling strategies*. 2018 IEEE International Conference on Big data (Big data), Seattle, December 10–13, 663–672.

CrowdTangle Team. (2019). *Crowdtangle API*. CrowdTangle Help. https://help.crowdtangle.com/en/articles/1189612-crowdtangle-api

Daniels, J. (2009). Cloaked websites: Propaganda, cyber-racism and epistemology in the digital era. *New Media & Society*, *11*(5), 659–683. https://doi.org/10.1177/1461444809105345

Daniels, J. (2014). From crisis pregnancy centers to TeenBreaks.com: Anti-abortion activism's Use of cloaked websites. In Martha McCaughey (Ed.), *Cyberactivism on the participatory Web* (pp. 152–166). New York, NY, USA: Routledge.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A system to evaluate social bots. In *arXiv [cs.SI]*. arXiv. http://arxiv.org/abs/1602.00975

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Di Benedetto Montaccini, V. (2019, May 14). Facebook fake news | Elenco di pagine e siti che diffondono notizie false. *TPI*. https://www.tpi.it/tecnologia/facebook-fake-news-pagine-20190514313552/

Donovan, J., & Friedberg, B. (2019). *Source hacking: Media manipulation in practice*. Data&Society. https://datasociety.net/output/source-hacking-media-manipulation-in-practice/

Farkas, J., Schou, J., & Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, *20*(5), 1850–1867. https://doi.org/10.1177/1461444817707759

Fletcher, R., Cornia, A., Graves, L., & Nielsen, R. K. (2018). Measuring the reach of "fake news" and online disinformation in Europe. Reuters institute factsheet.

Fletcher, R., & Nielsen, R. K. (2017). Are news audiences increasingly fragmented? A cross-national comparative analysis of cross-platform news audience fragmentation and duplication. *The Journal of Communication*, *67*(4), 476–498. https://doi.org/10.1111/jcom.12315

Fraser, L. (2020). What data is CrowdTangle tracking? CrowdTangle Help. https://help.crowdtangle.com/en/articles/1140930-what-is-crowdtangle-tracking

Freelon, D., McIlwain, C., & Clark, M. (2018). Quantifying the power and consequences of social media protest. *New Media & Society*, *20*(3), 990–1011. https://doi.org/10.1177/1461444816676646

Freitas, C., Benevenuto, F., Ghosh, S., & Veloso, A. (2015). *Reverse engineering socialbot infiltration strategies in Twitter*. 2015 IEEE/ACM International Conference on Advances in social networks analysis and Mining (ASONAM), Paris, France, August 25–28, 25–32.

Gerbaudo, P. (2018). *The digital party: Political organisation and online democracy* (1st ed.). Pluto Press.

Giglietto, F., Iannelli, L., Valeriani, A., & Rossi, L. (2019). "Fake news" is the invention of a liar: How false information circulates within the hybrid news system. *Current Sociology. La Sociologie Contemporaine*, *67*(4), 625–642. https://doi.org/10.1177/0011392119837536

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7821–7826. https://doi.org/10.1073/pnas.122653799

Gleicher, N. (2018, December 6). Coordinated inauthentic behavior explained. *Facebook Newsroom*. https://newsroom.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/

Gleicher, N. (2019, October 21). How we respond to inauthentic behavior on our platforms: Policy update. *Facebook Newsroom*. https://newsroom.fb.com/news/2019/10/inauthentic-behavior-policy-update/

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

HLEG EU Commission. (2018). *A multi-dimensional approach to disinformation. Report of the independent high level group on fake news and online disinformation*. EU commission. https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

Hui, P.-M., Yang, K.-C., Torres-Lugo, C., Monroe, Z., McCarty, M., Serrette, B., Pentchev, V., & Menczer, F. (2019). Botslayer: Real-time detection of bot amplification on Twitter. *Journal of Open Source Software*, *4*(42), 1706. https://doi.org/10.21105/joss.01706

Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Herr-Stephenson, B., Horst, H. A., Lange, P. G., Mahendran, D., Martinez, K. Z., Pascoe, C. J., Perkel, D., Robinson, L., Sims, C., & Tripp, L. (2010). *Hanging out, messing around, and Geeking out: Kids living and learning with new media (p. 432)*. The MIT Press.

Jack, C. (2017). *Lexicon of lies: Terms for problematic information*. Data & Society. https://datasociety.net/output/lexicon-of-lies/

Jack, C. (2019). Wicked content. *Communication, Culture and Critique*. https://doi.org/10.1093/ccc/tcz043

Jenkins, H. (2006). *Fans, Bloggers, and Gamers: Exploring participatory culture*. NYU Press.

Jenkins, H. (2008). *Convergence culture: Where old and new media collide* (Revised ed.). University Press.

Jenkins, H., Ito, M., & boyd, d. (2015). *Participatory culture in a Networked Era: A conversation on youth, learning, commerce, and politics*. John Wiley & Sons.

Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2019). Political Astroturfing on Twitter: How to coordinate a disinformation Campaign. *Political Communication*, 1–25. https://doi.org/10.1080/10584609.2019.1661888

Khan, J. Y., Khondaker, M. T. I., Iqbal, A., & Afroz, S. (2019). A Benchmark Study on Machine Learning Methods for Fake News Detection. In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/1905.04749

King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *The American Political Science Review*, *111*(3), 484–501. https://doi.org/10.1017/S0003055417000144

Korfiatis, N. T., Poulos, M., & Bokos, G. (2006). Evaluating authoritative sources using social networks: An insight from Wikipedia. *Online Information Review*, *30*(3), 252–262. https://doi.org/10.1108/14684520610675780

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice. 178*. https://psycnet.apa.org/fulltext/1945-02291-000.pdf

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Loader, B. D., & Mercea, D. (2011). Networking democracy? *Information, Communication and Society*, *14*(6), 757–769. https://doi.org/10.1080/1369118X.2011.592648

Luceri, L., Deb, A., Giordano, S., & Ferrara, E. (2019). Evolution of bot and human behavior during elections. *First Monday*, *24*(9), https://doi.org/10.5210/fm.v24i9.10213

Marwick, A. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetownlawtechreview.org*, *2*(2). https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-512.pdf

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

Mastinu, L. (2019, May 14). *TPI propone un nuovo elenco di pagine Facebook dispensatrici di odio e falsità*. Bufale. https://www.bufale.net/tpi-propone-un-nuovo-elenco-di-pagine-facebook-dispensatrici-di-odio-e-falsita/

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). "Fake news" Is Not simply false information: A concept Explication and Taxonomy of online content. *The American Behavioral Scientist*, *20*. https://doi.org/10.1177/000276421987822

Morstatter, F., Wu, L., Nazer, T. H., Carley, K. M., & Liu, H. (2016). *A new approach to bot detection: Striking the balance between precision and recall*. 2016 IEEE/ACM International Conference on Advances in social networks analysis and Mining (ASONAM), San Francisco, California, August 18-21. 533–540.

Newman, N., Dutton, W., & Blank, G. (2013). Social media in the changing ecology of news: The fourth and fifth estates in Britain. *International Journal of Internet Science*, *7*(1), 6–22. .https://ora.ox.ac.uk/objects/uuid:abd7aa83-49fb-47bc-88bf-a71ed8548926.

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. (2019). *Reuters institute digital news report 2019 (Vol. 2019)*. Reuters Institute for the Study of Journalism.

Phillips, W. (2018). *The Oxygen of amplification. Better practices for Reporting on Far right Extremists, Antagonists, and Manipulators*. Data & Society Research Institute.

Ratkiewicz, J., Conover, M. D., Meiss, M., Flammini, A., & Menczer, F. (2011). *Detecting and tracking political abuse in social media*. In Proceedings of the 5th AAAI International Conference on Weblogs and social media (ICWSM'11), Barcelona, Spain, July 17–21. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.646.5073.

Reis, J. C. S., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised learning for fake news detection. *IEEE Intelligent Systems*, *34*(2), 76–81. https://doi.org/10.1109/MIS.2019.2899143

Salisbury, M., & Pooley, J. D. (2017). The #nofilter self: The contest for authenticity among social Networking sites, 2002–2016. *Social Sciences*, *6*(1), 1–24. https://doi.org/10.3390/socsci6010010

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. Penguin Press.

Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019). The role of user profile for fake news detection. In *arXiv [cs.SI]*. arXiv. http://arxiv.org/abs/1904.13355

Silverman, C. (2017, December 31). I Helped Popularize The Term "fake news" And Now I Cringe Whenever I Hear It. *BuzzFeed News; BuzzFeed News*. https://www.buzzfeednews.com/article/craigsilverman/i-helped-popularize-the-term-fake-news-and-now-i-cringe

Twitter. (2019). *Elections integrity*. Twitter About. https://about.twitter.com/en_us/values/elections-integrity.html

Valeriani, A., & Vaccari, C. (2016). Accidental exposure to politics on social media as online participation equalizer in Germany, Italy, and the United Kingdom. *New Media & Society*, *18*(9), 1857–1874. https://doi.org/10.1177/1461444815616223

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). *Online human-bot interactions: Detection, estimation, and characterization*. Eleventh International AAAI Conference on Web and social media, Montréal, Canada, May 15–18. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/viewPaper/15587.

Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe Report*, *27*. http://www.theewc.org/content/download/2105/18430/file/INFORMATION%20DISORDER.pdf

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Weeks, B. E., & Gil de Zúñiga, H. (2019). What's next? Six observations for the future of political misinformation research. *The American Behavioral Scientist*, https://doi.org/10.1177/0002764219878236

Woodford, A. (2019, April 25). Protecting the EU elections from misinformation and Expanding Our fact-Checking Program to New Languages. *Facebook Newsroom*. https://newsroom.fb.com/news/2019/04/protecting-eu-elections-from-misinformation/

Woolley, S. C., & Howard, P. N. (2016). Automation, algorithms, and politics| political communication, computational propaganda, and autonomous agents — Introduction. *International Journal of Communication Systems*, *10*, 4882–4890. https://ijoc.org/index.php/ijoc/article/view/6298

Yang, K., Varol, O., Davis, C. A., Ferrara, E., Flammini, A., & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies*, *1*(1), 48–61. https://doi.org/10.1002/hbe2.115

Ye, J., & Akoglu, L. (2015). Discovering opinion Spammer groups by network Footprints. *Machine Learning and Knowledge Discovery in Databases*, 267–282. https://doi.org/10.1007/978-3-319-23528-8_17

Zhang, J., Carpenter, D., & Ko, M. (2013). *Online Astroturfing: A Theoretical Perspective*. https://aisel.aisnet.org/amcis2013/HumanComputerInteraction/GeneralPresentations/5/

Zhang, Y., Wells, C., Wang, S., & Rohe, K. (2017). Attention and amplification in the hybrid media system: The composition and activity of Donald Trump's Twitter following during the 2016 presidential election. *New Media & Society*. https://doi.org/10.1177/1461444817744390

# Appendices

## *Appendix 1. blacklist of problematic domains signaled by debunking websites*

tg-news24.com
retenews24.it
tg24-ore.com
notizie-24.com
informazionelibera.agency
streamingneuws.com
notizie-24.jimdo.com
oltrenews.com
adessobasta.org
attivo.news
skynew.it
llfatto.it
01.oltrenews.com
gazzettadellasera.it
larepubbica.it
tgnews.altervista.org
inedito.net
leggimiora.com
mondoit.info
tg-quotidiano.com
attivonews.com
conseilnational.blogspot.pe
guardiegiurate.altervista.org
howtodofor.com
ilfattoinedito.com
iloveitalia.net
ilvostropensiero.altervista.org
italia.imigliori.org
italia24.online
koenig2099.wordpress.com
leggiora.info
mondodiannunci.com
notiziarioonline360.com
notiziarioromacapitale.blogspot.com
peoplehavethepower.altervista.org
retenews24.net
sadefenza.blogspot.com
siviaggia.it
verodonna.it
interagisco.net
italiapatriamia.eu
kontrokultura.it
lastampatv.it
lonesto.it
mole24.it
noidelsud.com
nonguardarlo.com
notiziecorriere.it
senegaldirect.net
stampalibera.net
strettoweb.com
telegiornale.news
tg-quotidiano.net
tg-quotidiano.org
tutto24h.info
tuttoilcalcio24.com
tuttoinweb.com
vocidallastrada.org
zapping2017.myblog.it
24orenotizie.com
5ideeinmovimento.blogspot.it
ilradicale.altervista.org
ilrisvegliodellamente.altervista.org
ilsapereepotere.blogspot.it
ilsapereepotere2.blogspot.com
imigliorivideo.co
imolaoggi.it
incazzaticontrolacasta.com
informare.over-blog.it
informarexresistere.fr
informasalus.it
informatitalia.blogspot.com
informatitalia.blogspot.it
informazionebycuriosity.altervista.org
informazionelibera.eu
internapoli.it
ioco.altervista.org
iocritico.com
iovivoaroma.org
iozummo.com
irresponsabile.com
italiainmovimento.it
italiani-informati.com
italianiperlapatria.info
italianosveglia.com
italianotizie24.eu
italiarialzati.com
jedanews.it
jedasupport.altervista.org
jhwhinri.blogspot.it
kevideo.eu
lacucinasiciliana.altervista.org
laleva.org
lanozione.com
lapillolarossa15.altervista.org
larepubbica.it
larepubblica.info
lastella.altervista.org
lastnews24h.com
last-webs.com
laveritadininconaco.altervista.org
liberamenteservo.com
liberogiornale.com
libero-notizie.com
libreidee.org
linktoday.it
livewebit.com
lolandesevolante.net
luogocomune.net
madreterra.myblog.it
mafia-capitale.it
mammainrete.blogspot.com
mattino.ch
medicinenon.it
mednat.org
messaggiero.com
meteoweb.eu
misteriditalia.it
mondodeglianimali.altervista.org
mondonews1.blogspot.it
morasta.it
naturalnews.com

5stellenews.com
accademiadellaliberta.blogspot.it
actiionweb.com
actionweb24.com
agenews.eu
aidaa-animaliambiente.blogspot.it
albainternazionale.blogspot.it
alcyonpleiadi.blogspot.it
alfredolio.wordpress.com
alleanzadellasalute
all-news.tv
allnews24.eu
altrarealta.blogspot.it
anno3000.altervista.org
appunti2008.blogspot.it
aprilamente.info
assis.it
attaccomirato.com
attiviamoci.it
attivistam5snews.blogspot.it
attivotv.it
aurorasito.wordpress.com
autismovaccini.org
benesserefeed.it
benitomussolini.altervista.org
blogdieles.altervista.org
blogdieles2.altervista.org
bloginmovimento.altervista.org
blognews24ore.com
blogopenyoureyes.altervista.org
breaknotizie.com
byoblu.com
catenaumana.com
centrometeoitaliano.it
chesuccede.it
chiaveorgonica.altervista.org
clicknotizie.it
comedonchisciotte.org
comilva.org
compressamente.blogspot.it
condividilo.org
conoscenzealconfine.it
controinformo.altervista.org
controinformo.eu
controinformo.info
controinformoblog.altervista.org
corrieredelcorsaro.it
corrieredellapera
corrieredelmattino
corrieredelmattino.altervista.org
corrieredisera.it
cortesovrana.it
coscienzaitaliana.com
cosmofruttariano3m.altervista.org
curiosandoonline.com
curiosity2013.altervista.org
curiosityalive.altervista.org
cvdiariodelpollinho.altervista.org
delta-leveritanascoste.blogspot.it
devinformarti.blogspot.it
dionidream.com
direttainfo.blogspot.it
direttanews.it
direttanews24.com
direttanfo.blogspot.it
neovitruvian.wordpress.com
nessuna-censura.com
newapocalypse.altervista.org
news24italia.com
newscronaca.it
newsitalys.com
newstime24.altervista.org
nexusedizioni.it
nibiru2012.it
nocensura.com
noinvasione.com
nonsiamosoli.net
notixonline.com
notixweb.com
notiziaitalica.altervista.org
notiziarioeuropeo.altervista.org
notiziariosegreto.wordpress.com
notizie24h.net
notiziedalpopolo.altervista.org
notizieflash.eu
notizieinmovimentonews.blogspot.it
notizieora.it
notiziepericolose.blogspot.com
notiziepericolose.blogspot.it
notizieregionali.com
notiziespericolate.com
nuovamedicinagermanica.it
nwonuovoordinemondiale.wordpress.com
ohmiaitalia.com
olivieromannucci.blogspot.it
pandoratv.it
paoloferrarocdd.blogspot.it
paradisefruit.altervista.org
perdavvero.com
pianetablunews.it
pianetablunews.wordpress.com
pianetax.wordpress.com
piccolocuore.it
piovegovernoladro.info
politicanews.altervista.org
presadicoscienza.altervista.org
questaitalia.wordpress.com
raggioindaco.wordpress.com
rainews24.live
rassegnastampa.eu
realfunny.info
realscience.altervista.org
reattivonews.com
repubblica24.com
ribelli.eu
ripuliamolitalia.altervista.org
riscattonazionale.org
rischiocalcolato.it
rivoluzioneoraomaipiu.com
robadapazzi.com
rollingston.altervista.org
sadefenza.blogspot.it
saltoquantico.org
sapereeundovere.com
saper-link-news.com
scenarieconomici.it
scienzaeconoscenza.it
secretnews.fr
segnidalcielo.it
segretodistato.com

disinformazione.it
disquisendo.wordpress.com
divieto.reattivonews.com
ecocreando.weebly.com
ecologiasociale
ecotricolore.altervista.org
ecplanet.org
effervescienza.com
essenzialmentedonna.altervista.org
essere-informati.it
eticamente.net
europeannews.altervista.org
eurosalus.com
evidenzaliena.altervista.org
facebooknotizia.altervista.org
fascinazione.info
filosofiaelogos.it
frasideilibri.com
freeondarevolution.blogspot.it
funvideosonline.info
futuroacinquestelle.altervista
gazzetta24.com
gazzettadellasera.com
gazzettanews24.com
giornaleilsole.com
globo365.info
grandecocomero.com
guarda.link
guarda.one
guarireoggi.blogspot.it
hackthematrix.it
iconicon.it
ilbazarinformazione.blogspot.it
ilbellocheavanza.com
ilcorriere.cloud
ilcorrieredellanotte.it
ilfarosulmondo.it
ilfattaccio.org
ilfattodalweb.com
ilfattoquotidaino.it
ilgiomale.myblog.it
il-giornale.info
ilgiornaleitaliano.it
ilgiornalenews.com
ilmattoquotidiano.it
ilmeglio-delweb.it
ilmessagero.altervista.org
ilmessaggio.it
ilmeteo.it
ilmisito.blogspot.com
ilmondodinoidonne.com
ilmovimentovaavanti.alteravista.org
ilmovimentovaavanti.altervista.org
ilnavigatorecurioso.it
ilnord.it
ilnotiziario24.com
ilnuovomondodanielereale.blogspot.it
ilpopulista.it
ilprimatonazionale.it
ilpuntosulmistero.it
ilquotidaino.wordpress.com
il-quotidiano.info
semplicipensieri.com
senzacensura.org
siamolagente.altervista.org
siamonapoletani.org
siamorimastisoli.com
siamorimastisoli.today
signoraggio.it
silenziefalsita.it
sitoaurora.wordpress.com
skytg24news.com
socialbuzz.it
socialnotixweb.com
soloitaliani.com
sostenitori.info
sputtaniamotutti.reattivonews.org
stampa-lazio.com
stopeuro.org
succedenelmondo.com
sulatestagiannilannes.blogspot.it
superbamente.com
supernotizie.net
tankerenemy.com
tanker-enemy.com
tankerenemymeteo.blogspot.it
telegrafo.altervista.org
terrarealtime.blogspot.com
terrarealtime2.blogspot.com
tg24h.altervista.org
tg24italia.com
tg24notizie.altervista.org
tg5stelle.it
tgcom24news.com
tgnewsitalia.it
thecancermagazine.blogspot.it
thetonicexpress.com
thuglifevideos.com
ticinolive.ch
tmcrew.org
tuttiicriminidegliimmigrati.com
tutto24.info
ultimora24.it
ununiverso.altervista.org
ununiverso.it
vacciniinforma.it
veritanwo.altervista.org
videoenotizie.it
videoscoop.info
videoviraliweb.com
vivoinsalute.com
vnews24.it
vocedelweb.com
voxnews.info
webitalia360.com
web-news24.com
webnotice.altervista.org
webtg24.com
worldnewsdailyreport.com
wwwblogdicristian.blogspot.it
younetspiegalevele.altervista.org
younetspiegalevele.info
yournewswire.com
zonagrigianews.com

## Appendix 2. blacklist of problematic Facebook entities signaled by Avaaz

5Stelle TV
Adesso Basta
Adesso Basta! Movimento Italiano Contro La Politica Corrotta!
Adesso Italia
Affari condivisi
Amici di Beppe Grillo
Amici di Gianroberto Casaleggio
Anonymous attivisti
Beautiful exotic
CheMusica & News
CheNews e Viaggi.
CheVideo24.org
Conte il Premier Populista
Corriere della notizia
Curiosauro
Daniele Ferrari Attivista M5S
Di Maio Conte Ministri del popolo
Di Matteo-Morra-Gratteri fans Club
Dimissioni dei razzisti del Pd
DonneA5Stelle
E Sti Cazzi
Fans club di Alessandro Di Battista
Fascinazione
Forza Popolare
Fronte dei Popoli
Fronte del Popolo. Contributi e Dibattito.
Gli Attivisti Cambiano il Mondo
Gossip e Cucina
Governo Del Cambiamento
Governo Giallo- Verde Al Servizio Del Paese
Grande Cocomero Classic
Grillino a mia insaputa
Il Riscatto Nazionale
IL PD
Informazione Alternativa
IO SONO populista
IO SOSTENGO ADESSO ITALIA
Italia Patria Mia
Italia Sovrana e fuori dall'Europa
Italia uguale Dittatura
Jeda News
Kontrokultura
L'Angolo della Risata
L'attivista a 5 stelle
La Gazzetta della Sera By KontroKultura
La pagina eventi
La Repubblica Del Cazzeggio
La Verità ci Rende Liberi – Il Risveglio
La voce dell'Italia Onesta.
Lega – Salvini premier News
Leggilo.Org
Liberi e Indipendenti
M5S \ LEGA
M5s- Attivisti Blasonati
M5S InSide
Ma Anche No
Mafia Capitale
Mafia Capitales
MANDIAMO A CACARE L'EUROPA
Mondo Sporco
NERO Dentro
Nessuno di loro
Notizie in Movimento
Notizie Lega Nord
Pianetablunews
Piu'siamo PIU' Rumore Facciamo"
Politici Corrotti, confisca dei beni e carcere
Politici Mafiosi
Prima aiutiamo gli Italiani poi si vede. – Movimento Adesso Italia
Quel che non sai
Rassegne Italia
Rimani Informato
Riprendiamoci La Patria
Riscatto Nazionale
Salute e Alimentazione
Salute Eco Bio
Salvini quasi simile. a Mosè. HA ridato al popolo la forza
Segreto di Stato
Stelle di Vita
Travaglio fans club
UFO e Alieni le Verita' Nascoste
Un'Italia senza Renzusconi – fan club.
Vera Informazione
ViaggiNews e Sentimenti
Virus5Stelle
Viva La Patria
YouMovies.it