Journalists on Twitter: self-branding, audiences, and involvement of bots

Onur Varol & Ismail Uluturk

Journal of Computational Social Science

ISSN 2432-2717

J Comput Soc Sc
DOI 10.1007/s42001-019-00056-6

Journal of Computational Social Science
https://doi.org/10.1007/s42001-019-00056-6

RESEARCH ARTICLE

Journalists on Twitter: self-branding, audiences, and involvement of bots

Onur Varol[1] · Ismail Uluturk[2]

Received: 10 May 2019 / Accepted: 9 September 2019

Abstract

Spread of news and misinformation on social networks have been a topic of extensive study in the recent years. There are concerns about the possibility of ongoing information operations, which has lead to studies on a wide scope including the truthfulness of content and the participation of social bots in the process. Studying how online entities of journalists is embedded in the Twitter network is crucial for understanding the core of this problem, since they hold a valuable broadcast platform in informing the public. In this work, we collected over 290,000 accounts that self-identify as a journalist or a reporter and analyzed their professional and follower networks on the platform. Twitter follower composition of journalists reflect their potential audiences and who disseminates their messages further on the network. It is essential for a journalist to reach a broad, organic readership as opposed to a following of bots and bot-assisted accounts. We looked at the followers of journalists for an analysis of the composition and evolution of their audiences, particularly looking out for social bot involvement. We found the trends for verified and non-verified accounts to be opposite of each other; among verified accounts bot follower tend to target more popular ones, whereas unverified accounts have a higher fraction of bot followers early on when they have fewer followers, possibly indicating attempts at boosting apparent popularity artificially. Outcomes of this research emphasize the importance of editorial oversight and that the prestige of journalists should not be confused with their apparent popularity online.

Keywords Social bots · Online journalism · Self-branding · Social media

✉ Onur Varol
ovarol@northeastern.edu

Ismail Uluturk
uluturki@mail.usf.edu

1 Center for Complex Network Research, Northeastern University, Boston, USA

2 University of South Florida, Tampa, USA

Introduction

Widespread use of social media is bringing significant changes in many aspects of our daily lives, one of the most striking being the way we access information. In the Internet age, we have instant access to information from virtually anywhere in the world. Social media, in particular, makes it so much easier that we can hear about breaking news and most recent developments even before journalists had a chance to cover them [1–4]. As the internet changes the way we communicate and exchange information, Twitter has taken the function of a microphone to the masses [5–7].

Traditionally, journalists had an important role in curating, validating, and reporting important events around the world. Rise of the social media has been changing these traditional journalistic practices and public's expectations from news agencies and reporters [8]. The need for concise communication through micro-posts prevalent in social media platforms is effecting how ideas are now framed and presented by journalists [9, 10]. Journalism as a whole is also changing into a service from a product, where journalists are starting to use tools like live-stream events, and work to disseminate the voices of their audiences [8].

In addition to the news agencies increasing their adaptation to social media, journalists themselves are also improving their online presence [11]. Efforts such as personal branding and audience engagement has become more significant for their careers [12–14]; however, this may not necessarily lead to good journalistic practices. Therefore, analyzing how journalists engage with their audiences has paramount importance in understanding the impact of journalism in the age of social media. A recent Gallup survey found that only 40% of Americans trust mass media outlets to report news fully and accurately [15].

Social media are demanding for our attention, and we can only select a limited number of accounts to follow and engage with [16]. Our information reach is also subject to similar limitations, and we can only pay attention to a select few number of journalists and news sources. Criteria for this selection can follow certain heuristics such as topical alignment, popularity, or perceived authority of an account. Twitter addressed the problems related to fabrication of authority and impostor accounts by introducing "Verified" accounts in Summer 2009.[1] Verification is mainly used to establish authenticity of the identities on Twitter, and verified accounts display a badge (✓) next to their names on the platform. An analysis by Triggertrap found that journalists, with their 25% fraction among all verified accounts, are the largest and the most active category [17].

Researchers have reported that 9–15% of all accounts on Twitter are social bots [18], and they have been found to be used for infiltrating political discourse, manipulating the stock market, stealing personal information, and spreading misinformation [19]. It is possible that some of these bots may be purchased from 3rd-party services to follow journalist accounts. One possible motivation behind such an action can be artificially boosting the popularity of an account catering to the selection biases of

[1] https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts.

audiences, and apparent popularity in these cases do not imply trustworthiness in our information-rich world. An alternative explanation can be made from the perspective of a malicious actor that is trying to shape the perception of journalists on certain debates online and steering their attention towards misleading content.

In this work, we study social media with a focus on journalists. Personal branding and online popularity has been becoming increasingly important for journalists as their profession moves in the direction of becoming a service from a product, where its performative aspects are increasing in significance. This leads us to expect a heightened presence of social bots in the social networks of journalists. We aim to characterize this presence and detect distinct patterns, and present them as a precursor to further research in this area. We also aim to characterize the effects of the perceived authority and the "verified" badge program of Twitter, as social bot followers are one of the easiest ways to boost perceived authority and popularity online.

Contributions and outline

We present a detailed analysis of journalist accounts in the Twitter ecosystem. Considering their popularity and roles as authority figures, they are prime targets for information operations. Social bots, in particular, are commonly used for both increasing the apparent popularity of these accounts and attempting to direct their attention to take advantage of their central role on disseminating information online. In this paper, we make the following contributions:

- We identified journalist and reporter accounts on Twitter, collected information regarding to their user meta-data, how they present themselves online, and their social network structure.
- We analyzed followers of journalist accounts and studied the prevalence of social bots in their audiences. We identified distinct follower patterns, and present them with examples.
- We analyzed journalist accounts with Twitter "Verified" badges on their profiles, and how their audiences differ compared to their non-verified counterparts.

Data collection

Our analysis is based on data collected from Twitter. On this platform, users can provide free-text descriptions about themselves. This field is visible on the user profile, and accounts can use hashtags to indicate their interests (#tech, #politics, etc.), mentions to link relevant accounts such as employers, schools they attended, or professional organizations they are a part of (@nytimes, @columbiajourn, etc.), and URLs for linking external webpages. Examples of Twitter accounts together with their profile information including descriptions, friend and follower counts, and other user meta-data are displayed in Fig. 1.

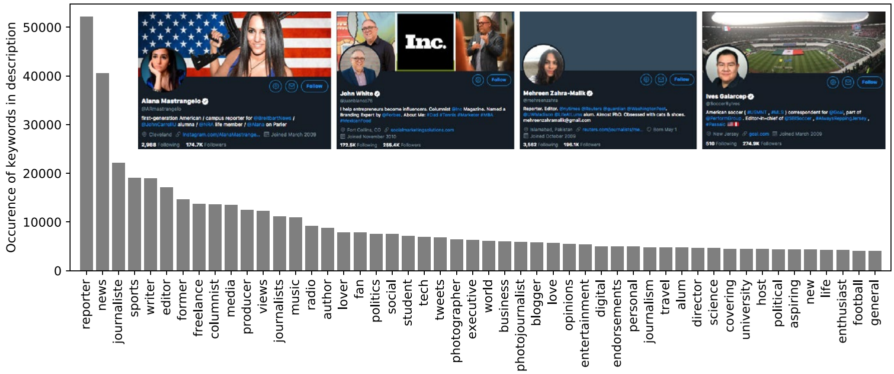


Fig. 1 Ranked list of top-50 frequent words in profile descriptions. Profile examples are selected among the most popular verified accounts

Identification of journalist accounts

The data set for our study got collected from a 10% random sample of all tweets streamed in real time, which was stored and post-processed after collection. We scanned the user meta-data in the collection of tweets for the period of 6 months between June and December 2018 to find the accounts that self-identify as journalists. Professional Twitter users often take advantage of the profile description field to share information on their interests, education history, and current employers. The practice of self-branding for journalists have been demonstrated for Australian journalists, where 95% of the accounts identified themselves as journalists and 90% of them provided their current employer [13]. Based on this evidence, we used the set of keywords "reporter", "journalist", "executive producer", "columnist", and "news editor" as a filter on profile descriptions, which lead to 294,500 unique accounts. We further narrowed the selection down to accounts that have at least 10 entries in our data set within the observation period.

Accounts in our sample use their profile descriptions often, and we observe that 44% use this field to share at least one mention, hashtag, or URL. In Fig. 1, we present the top-50 keywords used for mentioning journalists' professional and personal interests, job titles, and links to their emails and websites. Some examples of sample profile cards can also be seen in this figure.

Spatial and language distributions

We analyzed the meta-data provided for each account to characterize the profiles of journalist accounts. We studied the language, location, and profile descriptions for our collection of accounts. Geographical distribution of journalists and their corresponding languages are presented in Fig. 2.

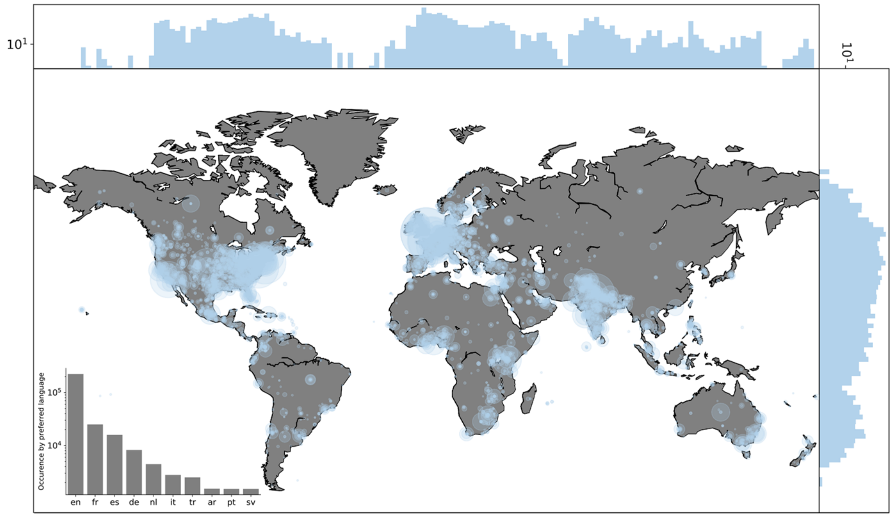


Fig. 2 Location information of journalists collected from profile meta-data through reverse geocoding. Size of the points are proportional to the number of accounts assigned to that specific locations. In the inset image, a histogram of user-defined languages are presented

We analyzed the accounts' preferred languages in our samples. Not surprisingly, we found that over 75% of the accounts were using English as their language setting, followed by French, Spanish, and German.

On Twitter, location of accounts can be identified either by GPS coordinates provided in tweet meta-data or by user-defined profile details. Although GPS coordinates provide high-resolution information on user mobility, the majority of the users do not share their location for privacy concerns. User-defined location information ranges from a resolution of city level details to just the country name. We collected the locations extracted from profile information and applied reverse geocoding using OpenStreetMap's API for collecting details on the country level.[2] Our collection of location information leads to 70,134 unique places, with 15.2% of accounts not providing any location information. There are 4508 unique locations that are referenced at least five times, which covers the 86.5% of all accounts. Remaining less frequent location names contain entries like "Somewhere and Anywhere" or "where the news goes" do not refer to actual places, and some journalists also like to share a link to their other online presences (email, website, social media) as their location. When all the valid location entries that were mentioned at least three times are considered, we observe that USA, India, and UK contribute to 27.6%, 7.8% and 7.2% of all the collected journalist accounts, respectively.

[2] https://nominatim.openstreetmap.org/.

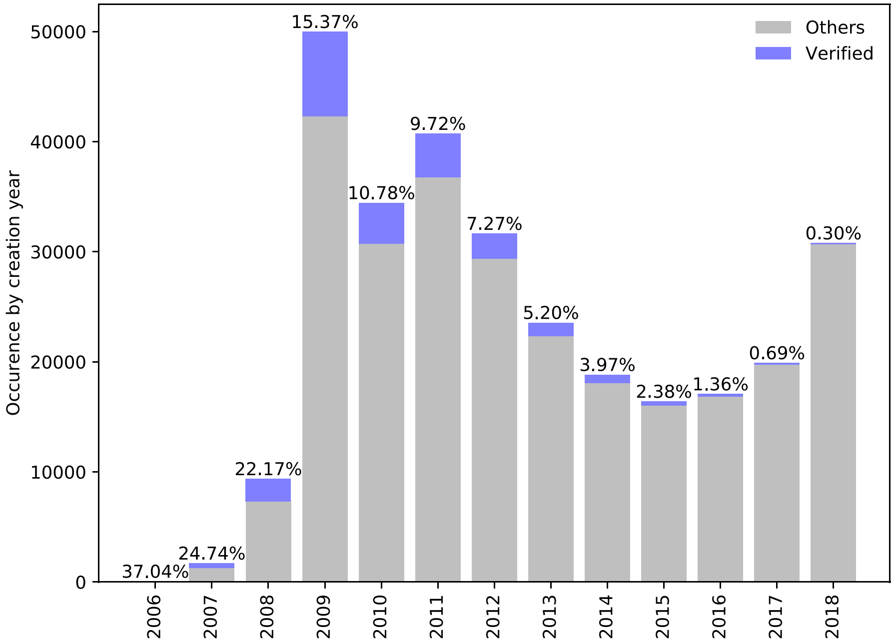


Fig. 3 Number of self-identified journalist accounts observed in our data set based on their year of creation. Using profile meta-data, verified accounts are highlighted and percentages of verified accounts are reported on top

Results

In this section, we are presenting our results for analyzing verified journalist accounts and their presence over time, co-mention networks of journalists and their apparent prestige on social media platforms, and the bot follower patterns of journalists with exemplar accounts highlighting distinct patterns of interest, where some of them exhibit concerning behavior.

Verified journalist accounts

In our analysis, we are especially interested in studying the "Verified" accounts, since they have to make an application and satisfy certain community standards to obtain this status. Twitter awards blue verified badges ( ) to let users know that the accounts for people of public interest are authentic. Although Twitter notes that verifying an account is not an endorsement, verified accounts are still seen as important in public's perspective.

We scanned our collection of journalist accounts to identify verified accounts using the meta-data and found almost 23,000 verified accounts in our data set. Since our analysis focuses on characterizing journalist accounts and their audiences, we

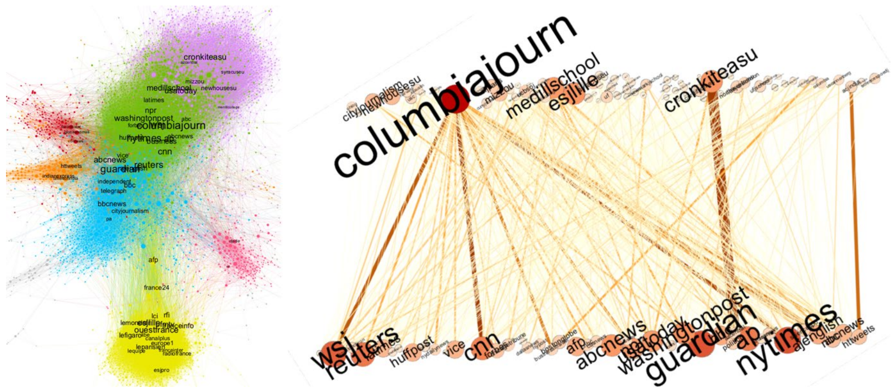


Fig. 4 Co-occurrence network of mentioned Twitter accounts in the journalist profile descriptions (left). Only items observed more than five times are kept in the network. ForceAtlas2 layout and Louvain modularity method for community detection is used for visualization. Bipartite network of news institutions and schools for accounts with English language settings (right). Edge thickness and colors reflect the co-occurrence between two institutions, and nodes with degrees under 10 are hidden

analyzed verified accounts separately to compare them with the non-verified journalist accounts. Comparison based on verification status provides opportunities to interpret the effects of external confirmation on authenticity and prestige of the accounts.

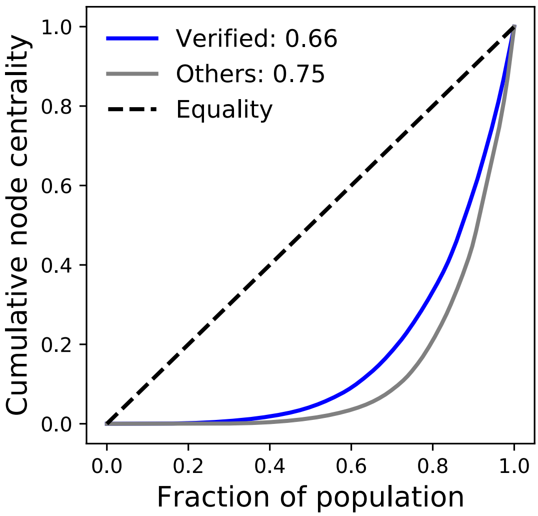
Veracity of the verified accounts are of interest when journalists start engaging with the social media platforms regularly. In Fig. 3, we present the number of accounts created each year that self-identify as journalists since 2006. Just the enrollment from year 2009 covers over 20% of our entire data set, and there is a consistent yearly increase at the rate of 20,000 or more accounts after that point.

An interesting observation of the growing journalism ecosystem on Twitter is the contradictory trend of the number of verified accounts. The fraction of verified accounts monotonically decreases from 37 to 0.3%, over 100 fold, starting from 2006. There may be multiple explanations for this observation: (i) more prestigious journalists of public interest verifying their accounts earlier; (ii) Twitter applying stricter rules when granting verified badges; (iii) increasing numbers of low-quality accounts presenting themselves as journalists. Considering the increase in account creation since 2016, we might suggest an increase in the low-quality journalism accounts. In addition to this observation, we also observe an increase in creation of new journalist accounts, where we should expect that most journalists to have already created their accounts.

Journalist profile co-mentions

Journalists in our data set frequently use the profile description field to link their work and education information by mentioning Twitter accounts of those organizations. We built a network of co-mentions using this profile information, as presented in Fig. 4.

Fig. 5 Lorenz curve for quantifying inequalities for journalist accounts with different verification statuses. The node with the highest eigenvector centrality that was mentioned in the profile description is used to compute the Gini coefficient



The co-mention network reveals communities that are relevant both geographically and contextually. We observe a distinct French community (yellow) separated from the other English speaking clusters. Organizations in the US can be distinguished as mainly news organizations like @washingtonpost, @cnn, etc. in the green cluster and the purple cluster contains most of the top journalism schools. Geographically, we can see British (blue), Canadian (red), and Indian (orange) communities as well.

We can identify organizations that exhibit higher importance, or prestige, in this co-occurrence network. For instance, @columbiajourn, @nytimes, and @guardian ranked at the top based on their eigenvector centrality in this network. We introduce a definition for the apparent online prestige of a journalist as the highest centrality out of the accounts mentioned in their profiles. Using this definition of journalist prestige based on associated organizations, we compute the Gini coefficient and Lorenz curves to quantify the uniformity, or equality, of the distribution of prestige among journalists. Gini coefficient is commonly used to measure social inequalities and values close to zero indicate a perfect equality, while values larger than 0.5 point to increasing inequality.

Verified accounts are found to have a more uniform prestige distribution in the co-mention network, while non-verified accounts exhibit higher levels of inequality, as shown in Fig. 5. The Gini coefficients for verified and non-verified journalist accounts are computed as 0.75 and 0.66, respectively. We observed lower levels of inequality in the verified group, and found them to be twice as likely to be associated with the top-5 venues (4.3% compared to 2% for the population) in terms of co-mention network centrality. Result of this analysis were consistent when repeated using pagerank and degree centrality as well.

We can also investigate the relationship between the schools journalists have graduated from and their workplaces, as shown in Fig. 4. This network highlights the significant institutions on the online journalism landscape on Twitter. Columbia

Journalism School (@columbiajourn) is one of the oldest journalism schools in the world, and the first one to offer a graduate program in the United States. Its strong influence over this landscape is evident, where it is strongly tied with a large number of significant news institutions, followed by Medill School of Journalism at Northwestern University (@medillschool), Walter Cronkite School of Journalism and Mass Communication at Arizona State University (@cronkiteasu), and ESJ Lille in France (@esjlille). We would like to point out that a small number of schools appear to have a strong impact on a large number of news institutions with significant Twitter presences, which may in turn help their graduates appear more prestigious online.

Measuring bot followers of journalists

Journalists hold an important position for information dissemination on Twitter. In addition to publishing in their regular venues, they are also expected to engage in online conversations and share their thoughts on current issues. Their authority puts them in a central position in the social network. Most journalists in our data set have over 1000 followers and the most popular one reaches an audience of over 5 million. Considering their authority and centrality in the network, journalist accounts are ideal targets for information operations seeking to deceive the public and spread misinformation. Once their attention is captured by malicious entities to retweet a post or reply to their content, journalists' actions are broadcast to their entire audiences.

We employed a publicly available social bot detection tool called Botometer[3] for detecting bot followers of accounts to characterize their follower compositions. Botometer is a supervised machine learning system that utilizes over a thousand features extracted from user and content meta-data to estimate the likelihood that the account is automated [18, 20]. On accuracy and performance of Botometer, it is reported to have an AUC score of 0.95 and the performance of the method have also been verified by third-party research studies, most recently by a study conducted by Pew research center reporting a precision of 0.82 and a recall of 0.86 on detecting bots in their sample [21].

Due to Twitter's API rate limits, collecting the information required by Botometer API for every account following the self-identified journalist accounts in our sample is not practical. To carry out the analysis in this section, we sampled 1000 accounts from our collection at random and collected their entire friends and followers through the Twitter API. We recorded the information of followers ranked by the recency of their following. We then collected the bot scores of a random sample of up to 1000 followers and analyzed the scores estimated for these followers to characterize the audiences of the journalist accounts.

Popularity and productivity of an account are closely related to how long they have been on Twitter. Recent accounts tend to have fewer followers, but if they are

[3] http://botometer.iuni.iu.edu.

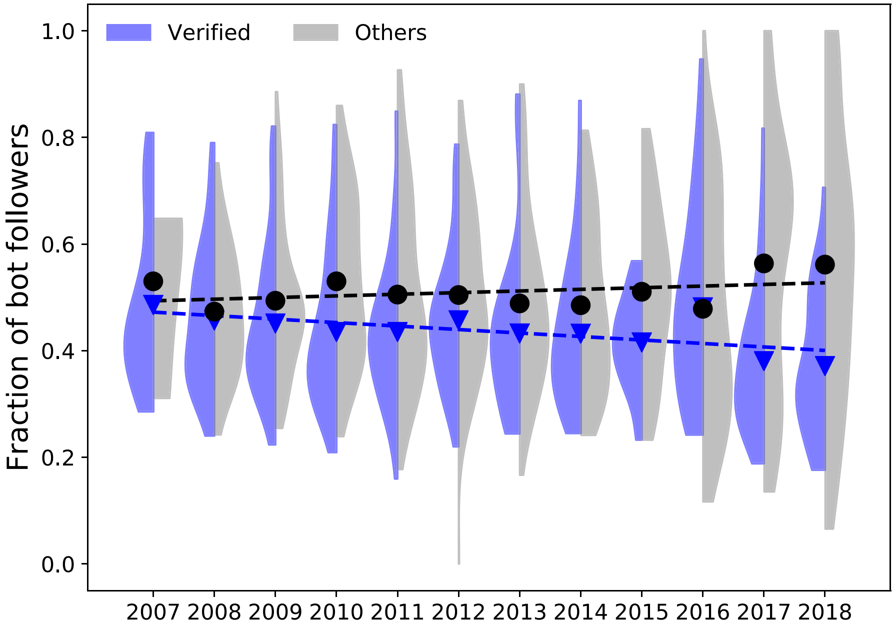


Fig. 6 Journalists grouped by their account creation year and their fraction of bot followers, compared by their verification status

already well known through other platforms, they can gain a significant number of followers quickly. Similarly, more established accounts might have a more loyal audience and their follower numbers might be more stable, or grow more gradually. Keeping the account age in mind, we analyzed the fraction of bot followers conditioned on the account creation date, as shown in Fig. 6.

We observe that the fraction of bot followers is lower for the newer verified accounts, while non-verified accounts are having an increasing trend of bot followers as they get younger. In this comparison of 10 years, the largest difference between the followers of verified and non-verified accounts is observed after 2016. We also observe a bi-modal distribution for non-verified accounts, indicating that the number of low-quality journalism accounts with high fractions of bot followers have been increasing in the recent years.

Figure 7 shows an analysis of bot followers for verified and non-verified journalist accounts. Verified accounts tend to have more followers overall as expected, and we observe a distinct difference between these groups in the number of bot followers. Non-verified accounts have higher fractions of bots in their audiences compared to verified accounts, and their bot follower fractions reach more extreme values.

We analyzed the relationship between the fraction of bot followers and the number of followers to investigate whether popular accounts attract more bot followers. Figure 7 shows this relationship, estimated separately for the two groups based on their verification statuses. We observe contradicting patterns in this figure: verified accounts attract more bot followers as they get more popular; however, non-verified accounts have a higher fraction of bot followers when they are less popular.

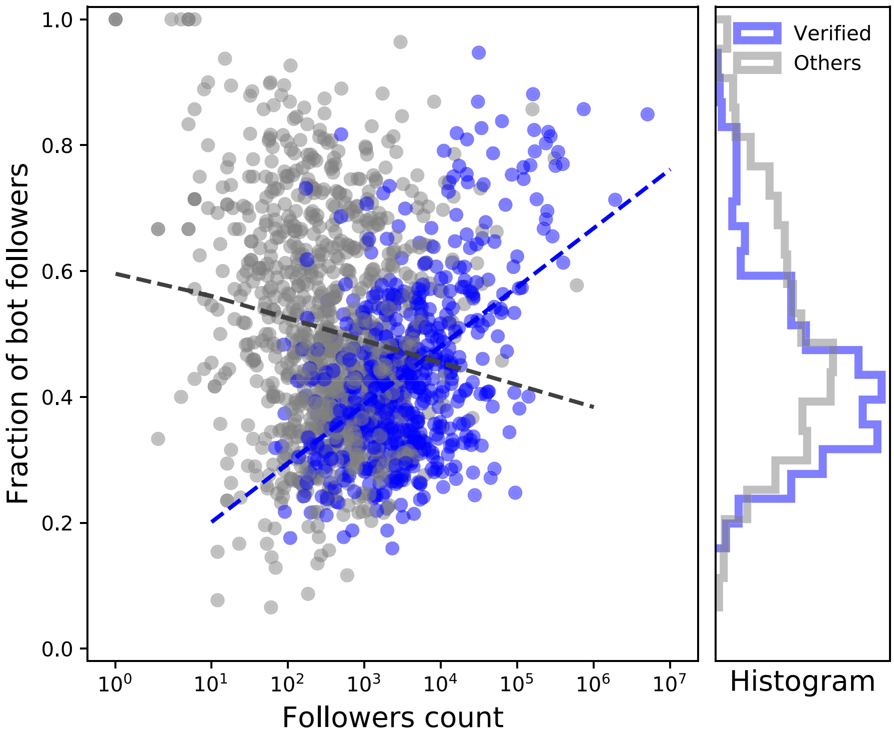


Fig. 7 Fraction of bot followers for verified and non-verified journalist accounts. Trends between bot fraction and follower count for the two groups are contradictory (left). Distribution of bot fraction have lower variance and average score for verified accounts (right)

We suspect that bots are targeting verified accounts to capture their attention and have them disseminate their messages, while non-verified journalist accounts might be employing social bots to increase their apparent popularity online. We provide a more detailed analysis of patterns we observed through individual journalists in the following section.

Bot follower analysis for exemplar accounts

We can explore follower patterns of individual accounts further by analyzing the temporal information of the time of creation for following accounts. An analysis conducted by NYTimes, The Follower Factory, monitors such patterns to highlight suspicious follower growth.[4] This analysis reveals a market for fake followers and automated audience growth for celebrity accounts. We extend their analysis by introducing bot scores of randomly sampled followers and analyzing the temporal trends

[4] https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html.

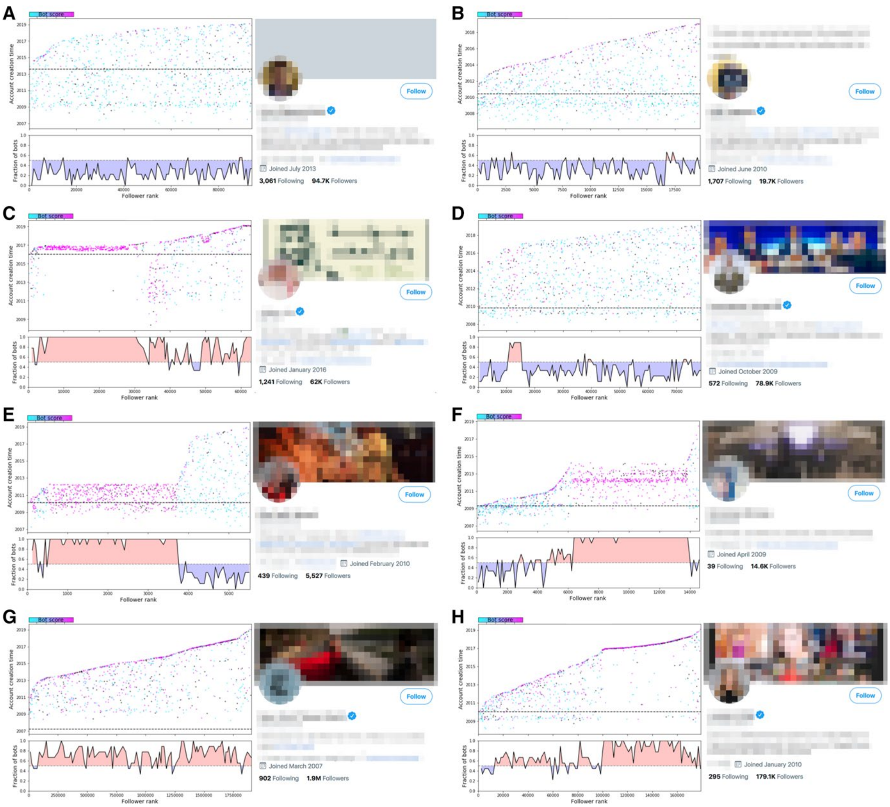


Fig. 8 Analysis of bot followers for exemplar accounts. A random sample of 1000 followers are displayed (colors indicate bot scores) based on their account creation date and rank in following the account. Dashed black line in the top panel indicates account creation date for the journalist. Average bot score within a sliding window is shown in bottom panel to capture the temporal evolution of bot scores

of these bot scores. This analysis, supported by the predicted bot scores for each follower, provides insightful observations on different patterns of bot engagement.

We present profile details and bot follower analysis results for eight different accounts in Fig. 8, six of which are verified accounts. Concerned with the privacy of these journalist accounts, we blurred their profile pictures, names, and profile details. We will refer to these accounts with their labels in the figure; such as `Account-A`.

We chose two popular journalist accounts at the top row to introduce desirable patterns of followers. Both of these accounts exhibit low fractions of bot followers and high numbers of total followers. `Account-A` has 94.7k followers and only 13.7% of which are likely to be bots. The other authentic journalist account, `Account-B`, has almost 20k followers, where 21.2% of them exhibit bot behavior. Examples presented here are some of the lowest bot fractions observed among popular accounts. We present accounts with more anomalous results below.

One of the most suspicious patterns is observed for `Account-C`, having 71.2% of their followers exhibiting bot behavior and a very uniform pattern of bot followers. The first 30,000 followers of this account exhibits high bot scores and they were all created in late 2017. Considering the journalist's account was created in January 2016, having nearly the first 50% of all followers in the first year of the accounts 3 year history is far from normal.

Another pattern frequently observed among journalist accounts is a couple thousand bot followers following these accounts early in their careers, and mostly around certain milestones such as reaching the first thousand or so followers. An example of such a pattern is observed for `Account-D`. Despite their mostly human audience (17.5% bot), their profile points to high volumes of bot engagement between their ten and fifteen thousandth followers. While it is not possible to be certain, this observation points to a purchase of bot followers from a third-party service, or conversely, an effort to create an orchestrated attack towards the account owner and/or their network.

We observed anomalous patterns more frequently for non-verified accounts. Two examples of such accounts are presented in the third row of Fig. 8. Both these accounts are owned by journalists that have relatively small numbers of followers. We observed that the first ~ 3000 followers of `Account-E`, a video gaming TV presenter, have very high bot scores and all 3000 (60% of all followers) of them started following the account within a short time frame in 2012. An example of late-term bot engagement is observed for the account `Account-F`. They gain more than half of their followers in 2014, and most of these follower exhibit bot behavior and were created in 2013. Despite having an organic growth early on, they later gain mostly bot followers increasing their fraction of bot followers to 57.8%.

We have also observed popular accounts with large fractions of bot followers such as `Account-G` and `Account-H`, alongside the accounts with suspicious temporal patterns. These users are attracting bot accounts that were created recently, and they tend to have very few aged bot accounts following them. We suspect that bot accounts follow them during the account creation process. Twitter requires users to follow at least 5 account when they are creating their Twitter accounts. These popular accounts might be the ones that are suggested by Twitter to those who are creating new accounts for bot farms. We noticed that the `Account-H` became more attractive to bot accounts when it reached 100,000 followers and suspect that might align with the time Twitter might start recommending their account.

Since it is practically impossible for researchers to trace the source of these bot followers and who orchestrated them to follow these sample accounts, we choose to avoid speculation and accusations, and use the accounts presented here merely to illustrate certain patterns of interest. We can, however, discuss the implications of the heavy involvement of bots with journalist accounts. First, and the most readily obvious observation is that bot followers are purchased to boost the apparent popularity of accounts on social media platforms. Malicious actors with access to large numbers of bots can also orchestrate them to influence and capture the attention of journalists. Their attention can be captured by manufacturing public interest around certain topics, while mentioning and engaging with the journalists. A more passive way to possibly influence entire careers of journalists is using large bot networks to provide shaped feedback to them through "engagement metrics", followers, retweets, mentions, click rate, etc., that are used to measure success on social media platforms and often closely monitored by account holders or managers. Journalists can be presented with positive feedback through increased engagement when they are working, and tweeting, on desirable topics and they can be similarly discouraged from focusing on undesirable topics. This can have a particularly significant effect on journalists early in their careers, or new on social media platforms, who may be willing to form their professional interests to optimize their engagement metrics and create their own niche audiences so that they can stand out in a saturated news environment.

Related work

Since its inception in 2006, Twitter has been a proxy to monitor the most significant events in the world. Social media has been used for political mobilization and information dissemination during events such as Arab Spring, Gezi movement, and occupy wall street [22–25]. The nature of social media makes it an ideal platform to search for breaking news and makes it possible to access information before traditional media can even mention it.

Journalism is one of the industries going through a significant transformation due to the advent of social media. While news agencies found early Twitter to be a useful advertisement tool [26], later social media transformed how journalists communicate the breaking news with their audiences [1–3].

Recent literature addresses the challenges of identifying different professions on Twitter, including journalists and reporters [27]. Once identified, further research has been on investigating what they post online and how they engage with their audiences [28–31].

Journalists have been acting as the gatekeepers of the means of information dissemination to the public, and the advent of social media have shifted these power dynamics. Researchers have been studying the effects of Twitter on journalistic practices. Lee et al. present that Korean journalists are less likely to share their opinions on Twitter about controversial issues when they expect a discrepancy between their opinions and the opinions of their Twitter audiences [32]. This example of the “spiral of silence” has been a common concern, and other researchers have been studying how online platforms impact traditional journalistic norms, such as objectivity and gatekeeping [33, 34].

Lasorsa et al. analyzed over 22,000 tweets, and found that journalists express their opinions more freely and engage with their followers more when they are working for more “prestigious” news agencies [35]. Competition among journalists for online popularity is encouraging them to develop personal brands and a significant online presence. Journalistic profession has been transforming into a performance, where engaging with the audience and conveying messages in an original way is starting to become more important for increasing the journalists’ “market value” [12]. Hanusch et al. analyzed over 4000 Australian journalists and found that journalists self-identify primarily through their professional characteristics, yet a significant fraction also provides personal details [13]. Our analysis of the most frequent keywords used in the profile descriptions aligns with this observation. Our analysis also points to the use of mentions to share employer and education information as a very succinct version of their resume. Molyneux et al. indicate that the personal and institutional branding are common for establishing authenticity, and the professional decorum is common for establishing credibility [14].

Pressure for authenticity and breaking the news first is affecting how journalists are practicing their jobs and their consideration of ethical issues. Recent research investigates how external mechanisms such as popularity of the content and journalist accounts indirectly pose editorial biases on their work. Jürgens et al. analyzed the dissemination of political information during 2009 German general election and found that a small fraction of accounts held critical gatekeeping positions for information dissemination, and they tended to filter the information based on their own political bias [36]. Similarly, expectations about practices of "good journalism" demands engagement between the audience and the journalists, and higher personal interaction with journalists leads to lower levels of perceived bias [37]. Rogstad et al. identify five distinct classes of journalists using cluster analysis—the skeptics, the networkers, the two-faced, the opiners and the sparks—based on their behavior online [38].

Automating content dissemination became a common practice on most venues; however, there have been efforts to automate content creation as well [39, 40]. Computational journalism aims to help journalists create engaging content more efficiently and apply automation to generate data-driven narratives algorithmically. Researchers have been raising concerns on automated content generation and its consequences on credit allocation, legal liabilities, and algorithmic biases [41, 42].

A recent Pew research center report found that one third of adults and a much larger fraction of young population use social media to access news [43]. Nowadays, most prestigious newspapers have a social media presence and Twitter is one of the leading platforms for disseminating news online [44].

Platforms like Twitter foster content production and dissemination. Automation seems like an innocent way of spreading news periodically by reducing the manual effort in online journalism [45]. However, we have been starting to understand how automation can also be weaponized to manipulate public discourse and electoral systems [19, 46, 47].

Conclusion

In this work, we systematically identified journalist accounts on Twitter and collected information regarding their profiles and audiences. Our analysis investigates how journalist accounts present themselves online and shows results in agreement with the literature that states journalists' self-branding efforts have become more prominent recently.

The popularity game between journalist accounts have become crucial in gaining authority and credibility. Journalist are the largest group in terms of having verified profiles [17] and yet we still observed that the fraction of verified journalists have been decreasing significantly over time. In the past 10 years, the fraction of accounts that received verified status badges from Twitter has reduced 100-fold. We would also like to note that Twitter has suspended general applications for the verified account program since November 2017.[5] Twitter is no longer assigning verified badges to new requests, but they are keeping the existing ones.

Our analysis points to significant differences between verified and non-verified journalist accounts in terms of their audiences and popularity. We found that fraction of bot followers increase as follower counts of verified accounts grow; however, non-verified accounts are observed to have larger fractions of bots when they are less popular. Similarly, perplexing patterns are observed as when we looked at accounts based on their creation date. Recent verified accounts tend to have significantly smaller bot followings compared to non-verified accounts, and this discrepancy between groups increases over time. These results point to positive implications of the verification process.

Considering the significant role of journalists on social media, malicious entities might target them to manipulate them into sharing unsubstantiated information by employing social bots to create an illusion of popularity. Examples of such targeting strategies have been observed during the 2016 US presidential election [47]. Historically, authoritarian figures have been used to manipulate the public opinion, and nowadays social media enables such information operations to be automated and made in a larger scale [48]. When we analyzed individual journalist accounts, we also found concerning patterns that might indicate purchase of social bots to boost apparent popularity on social media. It is also important to note that non-verified accounts tend to have larger numbers of bot followers, and they tend to be associated with lower prestige organizations.

In recent years, increase in automation and misuse of online networks have created new issues such as online information operations and targeted campaigns [46, 49]. Prevalence and impact of misinformation and fake news have attracted the attention of researchers [47, 50]. Such environments also require journalists to be more alert about the truthfulness of their sources [51]. Despite the importance of fact-checking, studies by Coddington et al. and Brandtzaeg et al. suggest that journalists and commentators tend to share more opinionated content on social media, and they are not familiar with the fact-checking and verification services [52, 53].

Social bots are a major tool for disseminating misinformation or disrupting online conversations [19, 54]. A survey conducted by Pew research center highlights interesting findings on the public perception about social bots: (i) two-thirds of Americans have heard about social bots, but only 40% is confident that they can identify social bots; (ii) over 80% of the respondents believe that some of the news in social media comes from bots; and (iii) more knowledgeable individuals are less likely to accept the use of social bots, especially by celebrities, political parties, and news organizations [55].

Considering the impact of social bots and the concerns among the social media users, it is crucial to set high standards for online journalism and exhibit caution about information obtained online. Platforms like Hoaxy[6] and

[5] https://twitter.com/TwitterSupport/status/930926225517719552.

[6] https://hoaxy.iuni.iu.edu/.

OSoMe[7] help researchers, journalists, and public to track, monitor, and analyze the spread of claims [56, 57]. In addition to developing systems that can help journalists and the public, it is also essential to be pro-active in the fight against the misuse of social media and reclaim these online platforms from automation [54].

From a broader perspective, we can start a discussion of the implications of journalists purchasing social bot followers, and what that means for journalism in general. Are journalists willing to boost their online popularity artificially? Is their selection of subject matters getting influenced by a pursuit of further online fame or engagement? How is the performance aspect of online journalism affecting their journalistic practices? These are all valuable questions as the transition of journalists from traditional to online media is still in progress. Twitter's verification program appears to serve as a good proxy for separating journalists based on their prestige and the involvement of bots in their audiences. However, what it means for readers when a journalist has a verified badge is still not clear. We would like to end our study on a positive note—in an open system like Twitter, researchers and public get to have a significant role in self-regulation and can be pro-active to exclude malicious actors from the system, and we have been observing increasing efforts in this direction.

Compliances with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Vis, F. (2013). Twitter as a reporting tool for breaking news: Journalists tweeting the 2011 UK riots. *Digital Journalism*, *1*(1), 27.
2. Hu, M., Liu, S., Wei, F., Wu, Y., Stasko, J., & Ma, K.L. (2012). Breaking news on twitter. In *Proceedings of the SIGCHI conference on human factors in computing systems* (ACM), pp. 2751–2754.
3. Bruns, A., Highfield, T., & Lind, R. A. (2012). Produsing theory in a digital world: The intersection of audiences and production in contemporary theory. *Twitter, and Breaking News: The Produsage of Citizen Journalism*, *80*(2012), 15.
4. Zubiaga, A., Heravi, B., An, J., & Kwak, H. (2019). Social media mining for journalism. *Online Information Review*, *43*(1), 2.
5. Morris, M., & Ogan, C. (1996). The Internet as mass medium. *Journal of Communication*, *46*(1), 39.
6. Murthy, D. (2011). Twitter: Microphone for the masses? *Media, Culture and Society*, *33*(5), 779.
7. Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web* (AcM), pp. 591–600
8. Artwick, C. G. (2013). Reporters on Twitter: Product or service? *Digital Journalism*, *1*(2), 212.
9. Wasike, B.S. (2013). Framing news in 140 characters: how social media editors frame the news and interact with audiences via Twitter. *Global Media Journal: Canadian Edition 6(1).*

[7] http://osome.iuni.iu.edu/.

10. Armstrong, C. L., & Gao, F. (2010). Now tweet this: How news organizations use Twitter. *Electronic News*, *4*(4), 218.
11. Bruns, A. (2012). Journalists and Twitter: How Australian news organisations adapt to a new medium. *Media International Australia*, *144*(1), 97.
12. Brems, C., Temmerman, M., Graham, T., & Broersma, M. (2017). Personal Branding on Twitter: How employed and freelance journalists stage themselves on social media. *Digital Journalism*, *5*(4), 443.
13. Hanusch, F., & Bruns, A. (2017). Journalistic branding on Twitter: A representative study of Australian journalists' profile descriptions. *Digital Journalism*, *5*(1), 26.
14. Molyneux, L., Holton, A., & Lewis, S. C. (2018). How journalists engage in branding on Twitter: Individual, organizational, and institutional levels. *Information, Communication and Society*, *21*(10), 1386.
15. Riffkin, R. (2015). Americans' trust in media remains at historical low, Gallup, September *28*
16. Gonçalves, B., Perra, N., & Vespignani, A. (2011). Modeling users' activity on twitter networks: Validation of dunbar's number. *PloS One*, *6*(8), e22656.
17. Mullin, B. (2015). *Report: Journalists are largest, most active verified group on Twitter*. Poynter, 26. https://www.poynter.org/reporting-editing/2015/report-journalists-are-largest-most-active-group-on-twitter/.
18. Varol, O., Ferrara, E., Davis, C.A., Menczer, F., Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In *Proc. Intl. AAAI Conf. on Web and Social Media (ICWSM)*
19. Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96.
20. Davis, C.A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A system to evaluate social bots. In *In Proc. 25th intl. conf. companion on World Wide Web*, pp. 273–274
21. Wojcik, S., Messing, S., Smith, A., Rainie, L., & Hitlin, P. (2018). Bots in the Twittersphere, Pew Research Center. *Retrieved May*, *22*, 2018.
22. Borge-Holthoefer, J., Rivero, A., García, I., Cauhé, E., Ferrer, A., Ferrer, D., et al. (2011). Structural and dynamical patterns on online social networks: The Spanish May 15th movement as a case study. *PloS One*, *6*(8), e23883.
23. Choudhary, A., Hendrix, W., Lee, K., Palsetia, D., & Liao, W. K. (2012). Social media evolution of the Egyptian revolution. *Communications of the ACM*, *55*(5), 74.
24. Conover, M. D., Ferrara, E., Menczer, F., & Flammini, A. (2013). The digital evolution of occupy wall street. *PloS One*, *8*(5), e64679.
25. Varol, O., Ferrara, E., Ogan, C.L., Menczer, F. & Flammini, A. (2014). Evolution of online user behavior during a social upheaval. In *Proceedings of the 2014 ACM conference on web science* (ACM), pp. 81–90.
26. Ahmad, A. N. (2010). Is Twitter a useful tool for journalists? *Journal of Media Practice*, *11*(2), 145.
27. Bagdouri, M., Oard, D.W. (2015). Profession-based person search in microblogs: using seed sets to find journalists. In *Proceedings of the 24th ACM international on conference on information and knowledge management* (ACM), pp. 593–602
28. Russell, F. M., Hendricks, M. A., Choi, H., & Stephens, E. C. (2015). Who sets the news agenda on Twitter? Journalists' posts during the 2013 US government shutdown. *Digital Journalism*, *3*(6), 925.
29. Molyneux, L. (2015). What journalists retweet: Opinion, humor, and brand development on Twitter. *Journalism*, *16*(7), 920.
30. Sheffer, M. L., & Schultz, B. (2010). Paradigm shift or passing fad? Twitter and sports journalism. *International Journal of Sport Communication*, *3*(4), 472.
31. Canter, L. (2015). Personalised tweeting: The emerging practices of journalists on Twitter. *Digital Journalism*, *3*(6), 888.
32. Lee, N. Y., & Kim, Y. (2014). The spiral of silence and journalists' outspokenness on Twitter. *Asian Journal of Communication*, *24*(3), 262.
33. Parmelee, J. H. (2013). Political journalists and Twitter: Influences on norms and practices. *Journal of Media Practice*, *14*(4), 291.
34. Lawrence, R. G., Molyneux, L., Coddington, M., & Holton, A. (2014). Tweeting Conventions: Political journalists' use of Twitter to cover the 2012 presidential campaign. *Journalism Studies*, *15*(6), 789.

35. Lasorsa, D. (2012). Transparency and other journalistic norms on Twitter: The role of gender. *Journalism Studies*, *13*(3), 402.
36. Jürgens, P., Jungherr, A., & Schoen, H. (2011). Small worlds with a difference: New gatekeepers and the filtering of political information on Twitter. In *Proceedings of the 3rd international web science conference* (ACM), p. 21
37. Gil de Zúñiga, H., Diehl, T., & Ardèvol-Abreu, A. (2018). When citizens and journalists interact on Twitter: Expectations of journalists' performance on social media and perceptions of media bias. *Journalism Studies*, *19*(2), 227.
38. Rogstad, I. D. (2014). Political News Journalists in Social Media: Transforming political reporters into political pundits? *Journalism Practice*, *8*(6), 688.
39. Broussard, M. (2015). Artificial intelligence for investigative reporting: Using an expert system to enhance journalists' ability to discover original public affairs stories. *Digital Journalism*, *3*(6), 814.
40. Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism*, *3*(3), 416.
41. Montal, T., & Reich, Z. (2017). I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism*, *5*(7), 829.
42. Lewis, S. C., Sanders, A. K., & Carmody, C. (2019). Libel by algorithm? Automated journalism and the threat of legal liability. *Journalism and Mass Communication Quarterly*, *96*(1), 60.
43. Mitchell, A., & Holcomb, J. (2015). State of the news media 2015. *Pew Research Center*, *29*
44. Ju, A., Jeong, S. H., & Chyi, H. I. (2014). Will social media save newspapers? Examining the effectiveness of Facebook and Twitter as news platforms. *Journalism Practice*, *8*(1), 1.
45. Lokot, T., & Diakopoulos, N. (2016). News bots: Automating news and information dissemination on Twitter. *Digital Journalism*, *4*(6), 682.
46. Badawy, A., Addawood, A., Lerman, K., & Ferrara, E. (2018). Characterizing the 2016 Russian IRA Influence Campaign, arXiv preprint arXiv:1812.01997
47. Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*(1), 4787.
48. Varol, O., & Uluturk, I. (2018). Deception strategies and threats for online discussions. *First Monday*, *22*(5)
49. Dutt, R., Deb, A., & Ferrara, E. (2018). Senator, We Sell Ads: Analysis of the 2016 Russian Facebook Ads Campaign. In *International conference on intelligent information technologies* (Springer), pp. 151–168
50. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146.
51. Schifferes, S., Newman, N., Thurman, N., Corney, D., Göker, A., & Martin, C. (2014). Identifying and verifying news through social media: Developing a user-centred tool for professional journalists. *Digital Journalism*, *2*(3), 406.
52. Coddington, M., Molyneux, L., & Lawrence, R. G. (2014). Fact checking the campaign: How political reporters use Twitter to set the record straight (or not). *The International Journal of Press/Politics*, *19*(4), 391.
53. Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, *12*(9), 1109.
54. Yang, K. C., 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*,. https://doi.org/10.1002/hbe2.115.
55. Stocking, G., & Sumida, N. (2018). *Social media bots draw public's attention and concern*. Washington, DC: Pew Research Center. https://www.journalism.org/2018/10/15/social-media-bots-draw-publics-attention-and-concern/.
56. Shao, C., Ciampaglia, G.L., Flammini, A., & Menczer, F. (2016). Hoaxy: A platform for tracking online misinformation, In *Proceedings of the 25th international conference companion on world wide web* (International World Wide Web Conferences Steering Committee), pp. 745–750.
57. Davis, C. A., Ciampaglia, G. L., Aiello, L. M., Chung, K., Conover, M. D., Ferrara, E., et al. (2016). OSoMe: The IUNI observatory on social media. *PeerJ Computer Science*, *2*, e87.

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