

Twitter bots, democratic deliberation and social accountability: the case of #OccupyWallStreet

Dean Neu and Gregory D. Saxton
*Accounting Area, Schulich School of Business, York University,
Toronto, Canada*

Twitter bots,
democracy,
and #Occupy-
WallStreet

Received 1 January 2023
Revised 7 August 2023
13 November 2023
Accepted 18 November 2023

Abstract

Purpose – This study is motivated to provide a theoretically informed, data-driven assessment of the consequences associated with the participation of non-human bots in social accountability movements; specifically, the anti-inequality/anti-corporate #OccupyWallStreet conversation stream on Twitter.

Design/methodology/approach – A latent Dirichlet allocation (LDA) topic modeling approach as well as XGBoost machine learning algorithms are applied to a dataset of 9.2 million #OccupyWallStreet tweets in order to analyze not only how the speech patterns of bots differ from other participants but also how bot participation impacts the trajectory of the aggregate social accountability conversation stream. The authors consider two research questions: (1) do bots speak differently than non-bots and (2) does bot participation influence the conversation stream.

Findings – The results indicate that bots do speak differently than non-bots and that bots exert both weak form and strong form influence. Bots also steadily become more prevalent. At the same time, the results show that bots also learn from and adapt their speaking patterns to emphasize the topics that are important to non-bots and that non-bots continue to speak about their initial topics.

Research limitations/implications – These findings help improve understanding of the consequences of bot participation within social media-based democratic dialogic processes. The analyses also raise important questions about the increasing importance of apparently nonhuman actors within different spheres of social life.

Originality/value – The current study is the first, to the authors' knowledge, that uses a theoretically informed Big Data approach to simultaneously consider the micro details and aggregate consequences of bot participation within social media-based dialogic social accountability processes.

Keywords Bots, Dialogic accounting, Social accountability, Twitter

Paper type Research paper

1. Introduction

Alternative accounting scholars have long taken a broad view of the notion of accountability and emphasized not just the *giving* of accounts of conduct but also the *demanding* of such accounts (Roberts, 1991), including demands made by the less powerful (e.g. Awio *et al.*, 2011; Rodrigue, 2014). The emergence of social media has encouraged grassroots activists and non-governmental organizations to utilize social media in the effort to hold governments and their allies accountable (She and Michelon, 2019; Saxton *et al.*, 2021; Al Mahameed *et al.*, 2021). These *social accountability* processes assume that social media provides, first, a site where

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

The authors gratefully acknowledge financial support from the Canadian Social Sciences and Humanities Research Council.



Accounting, Auditing &
Accountability Journal
Emerald Publishing Limited
0951-3574
DOI 10.1108/AAAJ-01-2023-6234

social actors can express their opinions about modes of government (Calhoun, 2011); second, a mostly unconstrained public space where a deliberative conversation stream about existing problems and potential solutions can emerge (Castells, 2008); and third, a public opinion aggregation mechanism that potentially encourages social change (Fraser, 1990).

While social media has the potential to facilitate large-scale public engagement in accountability-demanding efforts (Neu *et al.*, 2020, 2022; Al Mahameed *et al.*, 2021), there is evidence that it is not just human actors that are engaging but also automated and non-human *bots*. Estimates run as high as 9%–15% of all active Twitter accounts are being run by automated computer algorithms (Varol *et al.*, 2017). This is problematic insofar as these automated, algorithm-driven bots may have negative and unintended consequences on the social accountability conversation stream (Broniatowski *et al.*, 2018). Research in other fields suggests bots may be problematic if, by design, their purpose is not to add to the discussion but rather to maneuver it – whether to drive traffic to a blogger, to gain followers for a Twitter user or to sell a product, service or idea (Ferrara *et al.*, 2016). In other situations, bots may be designed to spread misinformation or unduly influence human discussion networks. For instance, in studies of elections, bots have been found to “have a tangible effect on the tweeting activity of humans” that ultimately enhances political polarization (Gorodnichenko *et al.*, 2021, p. 1). Building upon previous social accountability research on the dialogic potential of accounting (cf. Brown, 2009; Brown *et al.*, 2015; Velázquez *et al.*, 2017), we assume that social media participants – bots included – may start from an ethical stance (Kockelman, 2004, 2005) regarding existing forms of neo-liberal governance and a normative vision of what should change. This said, we propose that the problem with bots is not that they have a normative vision but rather that other conversation participants cannot easily identify bots and this, coupled with the ability of bots to speak “with volume” creates a situation where the participation of bots may unduly influence the conversation stream. For these very reasons, it is worthwhile to understand how the participation of a potentially influential social actor – in this case an algorithm-driven social actor – may both facilitate and constrain social accountability conversations (cf. Brown, 2009, p. 336).

Focusing on the Occupy Wall Street movement – and specifically the 9.2M+ tweets in the #Occupy conversation streams that took place on Twitter between August 2011 and July 2012 – we consider the consequences of bots on social accountability discourse via two broad research questions. The first research question examines whether the speech patterns of bots differ from other Twitter participants, whereas the second research question analyzes whether bots and their preferred speech topics unduly influence the trajectory of the social accountability conversation stream. Taken together, the analyses both document and help us to understand the consequences associated with Twitter bot participation in the Occupy Wall Street conversation stream.

It is important to note at the outset that we employ a naïve classification model based on the volume of tweets sent to identify whether the user is a bot. While this detection methodology is consistent with our theoretical framing and seems to be effective in identifying patterns indicative of automation, it may not unequivocally differentiate between high-volume human traffic and actual bots. While we address this concern via additional analyses based on an alternative, machine learning-based bot classification approach, we acknowledge that our main approach only identifies *likely* bots.

The study has two motivations. First, the study is motivated to understand how the participation of bots impacts on social media-based social accountability processes. Publicly interested accounting research has long assumed that accounting-informed dialogic conversations within the public sphere have the potential to shift the balance of forces sufficiently to achieve positive social and environmental outcomes (Gray *et al.*, 1997; Neu *et al.*, 2001; Bebbington *et al.*, 2007; Brown and Tregidga, 2017; George *et al.*, 2023). At the same time, this research has also recognized that there are limits to dialogue (Brown and Dillard,

2013). The current study complements and extends this line of research by examining an increasingly important site of public democratic conversation. The data that we have gathered provides us with the opportunity to not only analyze how the speech patterns of bots differ from other participants but also how bot participation impacts the trajectory of the Occupy Wall Street conversation stream. The current study is the first, to our knowledge, that uses a theoretically informed Big Data (Arnaboldi *et al.*, 2017) approach to consider the micro details of bot participation within dialogic democratic processes. In doing so, the study pulls back the curtain on this often invisible but increasingly important participant within social media-based social accountability conversations.

Second, the study is motivated to contribute to our understanding of dialogic accounting practices. Prior research has foregrounded how the potential of dialogic accounting practices as well as the limitations of these practices is tied to the social actors who participate in social accountability conversations and the ways that conversation venues simultaneously facilitate and constrain a free-flowing and relatively undistorted conversation (Gray *et al.*, 1997; Neu *et al.*, 2001; George *et al.*, 2023). Previous research acknowledges that not all people are allowed to speak and that not all people, even if they are allowed to speak, will likely be heard (Brown, 2009, p. 326). The current study picks up on these themes and considers how the participation of algorithm-driven social actors within an apparently unconstrained public space impacts on the ability of these other social actors to speak and be heard. In doing so, the study highlights how dialogic social accountability practices are never perfect since they are always enabled and constrained by the public space in which the conversations occur (Neu *et al.*, 2020). At the same time, the study reiterates that social accountability is a dialogic practice that depends on participants believing that, by participating in social accountability conversations, the possibility always exists that the conversation will result in felicitous perlocutionary consequences for oneself and for others (Butler, 2015).

The outline for our paper is as follows. In section 2 we present our starting theoretical premises concerning the role of social media in social accountability movements, while in section 3 we outline the theoretical issues surrounding the potentially deleterious role of bots in social media-based social accountability conversations. Section 4 presents our methods, section 5 our analyses and section 6 concludes with a discussion of our core contributions along with practical implications, limitations and promising avenues of future research.

2. Background: the role of social media in dialogic social accountability processes

Before presenting the theoretical role of bots in the Occupy movement in section 3, it is important to lay out the role of social media in social accountability movements. The emergence of Internet and social media platforms during the 1990s and early 2000s created a new public space where people could more easily talk to each other (Bellucci and Manetti, 2017; Castelló and Lopez-Berzosa, 2023) and where corporations – including transnational media organizations – have less control over what is said and done (Agostino and Sidorova, 2017; Brivot *et al.*, 2017). Internet technologies both shortened the transmission time of information from distant, geographically dispersed sites and bypassed the content filters of traditional print media. Furthermore, it allowed people to speak, and have a conversation, about the events of the day. Events such as the “live” reports during the Zapatista uprising in Chiapas, Mexico during the mid-1990s both captured the imagination of the public and demonstrated the potential power of this new medium (Froehling, 2013, p. 169).

In the years that followed the Zapatista Internet event, Internet-based technologies have been increasingly used to incite and sustain public conversations about important topics. A variety of social actors – from the smallest of grassroots organizations to the largest of international organizations – have enlisted these technologies to promote chains of activity

aimed at holding governments and their allies accountable (*cf.* Fieseler and Fleck, 2013). For instance, the individuals and grassroots organizations that helped to organize the Arab Spring uprising (Gerbaudo, 2012; Tufekci, 2017) and the *Indignados* movement in Spain (Anduiza *et al.*, 2014; Castañeda, 2012) used social media to hold public accountability-demanding conversations. Social media platforms can serve to channel and aggregate individual voices into a collective conversation that provides grassroots participants with the ability to demand social accountability in a way that potentially registers with politicians, governments and their business allies (Butler, 2015). Large international institutions like the World Bank (Agarwal *et al.*, 2009; O'Meally, 2013) and the UNDP (2013) have thus also been encouraged to include social media components in their social accountability initiatives. And whistle-blowing organizations such as Wikileaks (Heemsbergen, 2015) and the International Consortium of Investigative Journalists (ICIJ) have sought to make public previously private information with the purpose of inciting public conversation that leads to positive social change (ICIJ, 2018; Neu *et al.*, 2020, 2022). Overall, social media platforms such as Twitter have begun to fundamentally alter the practice of social accountability (Gomez-Carrasco and Michelin, 2017; Saxton *et al.*, 2021; She and Michelin, 2019).

The aforementioned social accountability initiatives are all quite different, but they share a belief in the importance of democratic dialogic participation (Bebbington *et al.*, 2007, p. 362; Brown and Tregidga, 2017, p. 2; George *et al.*, 2023) and the potential of social media to both provide a space for social accountability-focused public conversation and to have these conversations potentially lead to positive social change [1]. Discussions about the importance of democratic dialogic participation have a long intellectual lineage; however, it is the writings of Habermas – notably *The Transformation of the Public Sphere* (1962) and *The Theory of Communicative Action* (1984) – that are often used as a touchstone for current-day thinking (George *et al.*, 2023, p. 4 footnote 10). This literature is useful in that it provides us with a set of theoretically informed precepts about how the structural features of a communication venue like social media facilitates and/or impedes broad-based, democratic social accountability conversations. The remainder of this section outlines the three aspects of this perspective that are relevant to the current study.

First, this perspective starts from the premise that conversations within public spaces *can* inform and feed into democratic processes (Bebbington *et al.*, 2007, p. 367; Gerhards and Schäfer, 2010, pp. 144–145; Brown and Tregidga, 2017, p. 1 footnote 2). Researchers working in this tradition assume that public spaces are fundamental to democratic processes because such spaces allow individuals to come together to talk about and question the modes of government that organize their daily lives (Ravazzani and Mazzei, 2018, p. 188; George *et al.*, 2023). As Castells (2008, p. 78) notes, “between the state and civil society is the public sphere.” It is within this space of betweenness that individuals come together and “struggle to reconcile individual self-interest with the achievement of an ethical community” (Calhoun, 2011, p. 312). Furthermore, it is through these acts of coming together in conversation, and coming together in action, that create the possibility that governments and their allies can be held accountable. Calhoun summarizes this line of thought, stating that “the public sphere was crucial to the hopes of democracy in that it connected civil society and the state through the principle that public understanding could inform the design and administration of state institutions to serve the interests of all citizens” (2011, p. 312).

Second, attention to the structure of communication channels and an assessment of the impact of these structures on the characteristics of communication is important. From this vantage point, “better” conversations are inclusive, relatively non-distorted and mostly unconstrained (Sabadoz and Singer, 2017, pp. 185–186). This notion of better starts from the writings of Habermas on “ideal speech acts” to think through the necessary conditions for such conversations. For Habermas, ideal speech situations are those where there are reasoned exchanges, reflexivity, turn-taking, sincerity and transparency, formal inclusion and

discursive equality, and autonomy from state and corporate power (Dahlberg, 2005). Conversation streams that fall closer to the ideal end of the spectrum should result in conversations that are more free-flowing and more likely to result in vibrant and creative yet unpredictable democratic deliberation (cf. Butler, 2016).

Twitter bots,
democracy,
and #Occupy-
WallStreet

Third, this viewpoint stresses that democratic, dialogic conversation is not a panacea, nor is it conflict free. Furthermore, these conversations are simultaneously deliberative – in the Habermasian cognitive sense – but also affective in that they allow social media participants to express their emotions and perform their demands for social accountability. In this regard, social accountability conversations are not devoid of different opinions, conflict and drama since people think about and feel about the precarity that comes with globalization and neo-liberalism differently (Butler, 2015, pp. 151–152). This said, what characterizes dialogic democratic social accountability conversations is the willingness to hear the opinions and emotions of others as well as the willingness to work through the differences that characterize dialogic conversations (Kim *et al.*, 1999, p. 363; Brown, 2009, p. 320; Juris *et al.*, 2012). Furthermore, as Butler (2015) reminds us, social accountability demands do not always “spill into the street” and result in large-scale permanent changes but they do have performative consequences for the people who participate in dialogic social accountability conversations (2015, p. 208). We agree with Butler that it is the possibility for felicitous societal level change *plus* the opportunity for individual citizens to participate in and perform democracy via dialogic conversation that are important.

3. The problem with bots

On the surface, social media is an ideal site for democratic dialogic conversation because there are fewer barriers to participation compared to traditional public spheres. For example, in the salons of the eighteenth century that Habermas describes, participation was constrained by geography as well as social class. Within traditional print media, from the time of Habermas up until the present, newspaper editors continue to operate as gatekeepers filtering which narratives receive print space (Gerhards and Schäfer, 2010). In contrast, social media spaces are often assumed to be less constrained, but the reality is that social media is owned and operated by large public corporations who have incentives to generate profits (Brivot *et al.*, 2017; Neu *et al.*, 2020). Furthermore, as we propose below, the unconstrained participation of bots does impact on social media-based democratic deliberations. The remainder of this section discusses these impacts.

Before starting, it is important to state that social media spaces are digital spaces that use technology to facilitate showing up as well as speaking (Graham and Ackland, 2017). These technologies facilitate public speech but also *amplify* and *create* speech communication. For example, a Twitter participant can amplify one’s message by saving a tweet on the computer and periodically resending the same tweet to ensure more audience members see the tweet. Alternatively, a Twitter participant can write a computer program to periodically send the same tweet or a differently worded but similar tweet. If he/she is proficient enough, he/she can program the computer to respond to the tweets of other participants and even “learn” from the communication stream to generate a different tweet message. In these cases, computer technologies facilitate, to different degrees, the participation of social actors within the conversation [2].

Goffman’s (1981) research on speech participation roles is useful in thinking about these different modes of participating. More specifically, Goffman distinguishes among *animators*, *authors* and *principals*, noting that speakers play multiple, sometimes simultaneous, roles. The animator, in this case the bot, is active in the role of utterance production: “it is the one who speaks the words, ‘the talking machine’” (Manning and Gershon, 2013, p. 111). The author, in turn, is “someone who has selected the sentiments that are being expressed and the

words in which they are encoded” (Goffman, 1981, p. 144). Finally, the principal is someone whose position is established by the words spoken, someone who is “committed to” what the words say (Manning and Gershon, 2013, p. 111). Goffman’s different speech roles highlight that while a bot speaks, the action of speaking is connected to a human social actor who programs (i.e. selects) what will be said and whose values inform the words spoken. Furthermore, this author has a normative vision that is articulated through the ways that the bot has been programmed to speak and react. From this vantage point, it is a human social actor that is responsible for the words that bots animate (see also Johnson, 2015, p. 708).

Goffman’s distinction between animator and author/principal is useful to the current study because it reminds us not only that bots, qua animators, are the progeny of human social actors but also that the computerization of speech utterances is a matter of degree in that all social media interaction enlists some amount of computer technology. From this vantage point, the issue is not whether bots are non-humans but rather whether bots result in speech communication that is *less ideal* than the types of speech communication that occurs in other public spaces (see also Graham and Ackland, 2017).

Returning to the list of criteria regarding impediments to ideal speech communication that was presented in the preceding section, we propose that there are three potential concerns with bot participation. First, bots have a greater ability to speak *with volume* than do other participants. As mentioned previously, dialogic conversation presumes that participants take turns speaking and that there is a temporal space for hearing, thinking and feeling about the utterances of the participants. Arguably, speaking with volume both violates the norm of turn taking and fills up the temporal spaces thereby making dialogic conversation more difficult.

In part, the deliberative aspects of conversation are more difficult because audiences cannot easily see *how often* a speaker is speaking. Social media sites allow for mostly anonymous participation compared to other public spheres, but this partial anonymity may be positive because it facilitates a less constrained conversation where the class, race, gender etc. of the speaker does not distort how the message is heard (cf. Witschge, 2004). Rather, the problem with bots is that other participants may not realize that they are seeing the same, or very similar, tweets from the same sender multiple times. In this way, the participation of bots undermines the possibility for a transparent conversation.

Second, the ability of bots to speak with volume impacts on the ability of other participants to be heard within the conversation. While social media sites such as Twitter do not have explicit limits on how many tweets can be sent at a given moment in time, the probability that a tweet will be read decreases as the aggregate volume of tweets increases. For example, there were more than 20 days within the Occupy Wall Street conversation when the volume of tweets exceeded 50,000. This works out to more than 2,000 tweets an hour and 35 tweets a minute. With this amount of traffic, it is almost impossible for all tweets to be read, thought about and responded to. This amount of traffic, along with the difficulty for the audience to quickly identify participants who speak with volume, has the potential to result in a situation where each individual tweet receives less attention but where the repetitive messages sent by a bot receives greater aggregate attention. Stated differently, the “speaking with volume” strategy of bots crowds out the ability of other participants to be adequately heard.

Third, after a bot has been programmed, the amount of marginal effort to send an additional tweet is less than for participants who are mechanically typing their tweets. This differential effort not only makes it easier to speak with volume but also increases the probability that the tweeting activities of bots persist within a topic for longer than for other participants. Previous research notes that Twitter utilizes computerized algorithmic sorting to sustain aggregate participation levels (Graham and Ackland, 2017, p. 6). These techniques include “what’s trending” algorithms as well as tailoring content for different users. The

algorithms help to ensure that participants continue to use Twitter, in part by enticing users to move to new topics before they become satiated. It is this combination of computerized participation techniques – those of the bot and the social media platform – that skew that nature of who participates in a conversation stream. More specifically, it results in a situation where bots are more likely to persist for longer periods of time within a conversation stream than other users.

Taken together, these three aspects of bot participation result in a situation where the utterances of bots may not only be disproportionately seen within the stream but also change the subsequent speaking patterns of other participants. We assume that the disproportionate seeing of bot tweets is a weak type of influence that distorts the stream but doesn't necessarily change the nature of democratic dialogic conversation. In contrast, the adoption by other participants of the points of view expressed by bots indicates that the strategy of speaking with volume has had illocutionary force. We assume that this is a stronger form of influence.

At the same time, the participation of bots has the potential to *improve* the quality of democratic conversation. For example, bots can quickly collect vast amounts of information from a variety of sources and to learn from the utterances of other Twitter participants. For example, [Graham and Ackland \(2017\)](#) suggest that bots have the potential to break the “filter bubble” that is a consequence of the increasing, and often algorithm-induced, fragmentation of social media. The authors suggest that bots, if successful at breaking such filter bubbles, can contribute to social media-based dialogic democracy by de-fragmenting social media conversations. If so, the ability of bots to learn from other participants, summarize what has been said within the conversation stream, and disseminate messages to many participants may facilitate better forms of democratic conversation (cf. [Graham and Ackland, 2017](#), p. 24). Of course, these outcomes are ultimately empirical questions in that the outcomes depend on both how the bots have been programmed and how other participants respond.

In the subsequent analyses, we focus on two research questions that follow from our theoretical framing regarding the narrative consequences of bot participation. The first research question (RQ1) uses topic modeling techniques to identify what bots speak about and whether this is different from the speaking patterns of other Twitter users. The second research question (RQ2) attempts to identify the ways that the speaking patterns of bots exert weak and strong form influence on the overall Twitter conversation stream.

4. Data and method

In the aftermath of the 2008 financial crisis, the Occupy Wall Street movement emerged as a grassroots protest against neo-liberal modes of governance and the intricate financial securitization structures that uphold such governance ([Barthold et al., 2018](#); [Hardt and Negri, 2011](#)). Spurred on by an email sent on July 13th, 2011 by the anti-consumerist pro-environmental magazine *Adbusters* [3], the call to “#Occupy Wall Street” sparked global protests, both physical and virtual. People occupied Zuccotti Park and public spaces in 950+ cities across 82 countries ([NPR, 2011](#)). In contrast to the anti-globalization protest that occurred in Seattle a dozen years prior, the Occupy Wall Street movement also harnessed the power of newly emergent social media – particularly Twitter – to construct a virtual people's assembly that fueled vibrant citizens' conversations ([Tufekci, 2017](#)). The Twitter conversations during Occupy Wall Street thus hold significance as a research milestone, potentially enriching our insights into the emergence and articulation of collective citizen-driven social accountability conversations.

Building on this opportunity, our data consists of 9.2 M+ English-language tweets that contained some variation of the hashtag *#Occupy* (e.g. *#OccupyWallStreet*, *#OccupyWallSt*, *#OWS*, *#OccupyCanada*, etc.) and that were tweeted between August 1, 2011 and July 31,

2012. These tweets were purchased from Twitter. The data contained approximately 3.9 M original tweets as well as 5.3 M+ tweets that were retweeted by a different participant. Included within the data was metadata consisting of the tweet ID, the user ID, the tweet text, when the tweet was sent and the number of followers that the tweet sender had.

Figure 1 shows the aggregate tweet volume per day over the period that we consider along with a timeline of key dates. As seen in the figure, an incipient conversation had begun by August in response to *Adbuster's* email campaign and promotion of the #OCCUPYWALLSTREET hashtag. However, it was not until around September 17th, as the first physical occupations took place in New York's Zuccotti Park, that the Occupy conversation stream exploded. The conversation stream was at its most active during the first two months, with the biggest peak coming around November 15th in the aftermath of widespread evictions and arrests of protesters in Portland, Oakland and New York City. There was another peak in early January surrounding several smaller events – the attempt to re-occupy Zuccotti Park, a flash mob in Grand Central Station, and an influential series of articles on “Undermining the Case for Capitalism” published in the *Financial Times* – but, after this, the daily tweet volume declined to around 7,000 tweets per day. The final peak occurred around the May Day protests that occurred in major cities across the globe.

The first analysis step was to use the aggregate number of tweets sent by each of the 442,820 unique user IDs to identify bots. We wrote an R script that summarized the quantity of tweets by user ID and then calculated the percentile distribution of the aggregate tweet counts. The tweet count percentile information allowed us to identify user ID's that were in

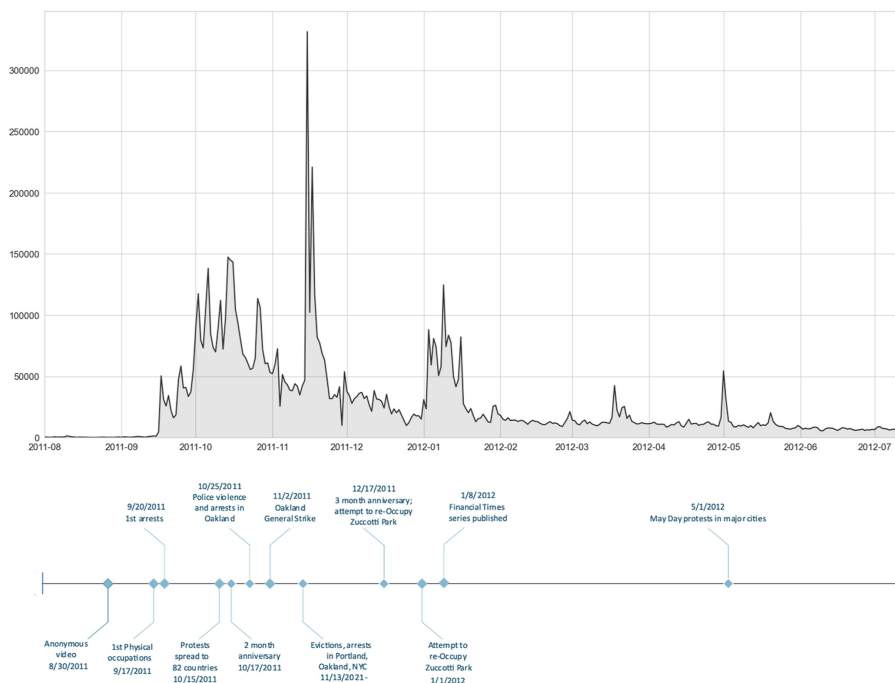


Figure 1.
Daily tweet volume and
main occupy events

Note(s): Figure shows daily number of English-language tweets sent between August 1, 2011 and July 31, 2012 containing some variation of the hashtag #Occupy (e.g., #OccupyWallStreet, #OccupyWallSt, #OWS, #OccupyCanada, etc.)
Timeline below indicates key events in the Occupy movement

Source(s): Authors' own work

the upper tail of the tweet count distribution. A review of the distribution data highlighted that, while the mean user sent 20.91 tweets, the top 30 participants each tweeted more than 21,000 times during the one-year period, whereas the top 50 participants each sent more than 13,000 tweets during the same period. In contrast, more than 432,000 users sent less than 100 tweets each. The initial analysis that follows uses this naïve bot classification method based on tweet volumes where we assume that the top 30 participants are bots [4]. Our decision to use a naïve classification model is based on our theoretical framing which suggests that tweeting with volume is of theoretical interest. The choice of a naïve model, if inappropriate, would bias against finding differences in speaking patterns. While we think that there is a high probability that the 31st to 50th most prolific user IDs are also bots, we decided to focus on the top 30. Once again, if the 31st to 50th grouping are, indeed, bots, this will bias against finding significant differences in the speaking patterns of bots and other participants.

Our bot detection method is predicated on the belief that a defining characteristic of bots is that they speak with volume (e.g. Balaanand *et al.*, 2019). However, we acknowledge that our approach only identifies *likely* bots and may not fully differentiate between bots and highly active human users. To mitigate concerns over this approach, in additional analyses we relax our assumption that tweet volume is the defining characteristic of bots and re-run our analysis using an alternative, machine learning-based bot measure (see Table 5). This subsequent approach validates our assumption that, within the Occupy Wall Street conversation, tweet volume is a defining characteristic of bots.

Using the 21,000 aggregate tweet number as the dividing point to partition the data into bots and non-bots resulted in a corpus of 1.4 M tweets being attributable to bots. Starting from this corpus, we then used *latent Dirichlet allocation* (LDA) topic modeling techniques to identify the latent themes in the speaking patterns of bot and non-bot actors (DiMaggio *et al.*, 2013; Blei *et al.*, 2003). The LDA topic modeling approach is increasingly popular within textual analysis research because it identifies topics within large textual data sets (cf. Brown *et al.*, 2020; Loughran and McDonald, 2016; Mohr and Bogdanov, 2013) and is “one of the most commonly used algorithms for classifying short texts” such as tweets (La Torre *et al.*, 2022, p. 9).

We used the LDA-based structured topic model algorithms of Roberts *et al.* (2014a, b) included in the *STM* package in *R* to identify the topics and to generate a listing of the *high-probability* (HPROB words, as they are referred to in STM) and *frequent-and-exclusive* words (FREX, in STM) associated with each topic. In line with best practices (e.g. DiMaggio *et al.*, 2013; Roberts *et al.*, 2014a; La Torre *et al.*, 2022), we used a combination of quantitative LDA diagnostics (see Appendix) and researcher judgment on the alignment between the automated results and the underlying tweets to determine that the model with 25 topics made the most “real world” sense based on our qualitative interpretation of the data [5]. As the next sections illustrate, the *high-probability* and *frequent-and-exclusive* words for the 25 topics, along with example tweets for each topic, are reasonably evocative as to what the different topics are talking about.

While we think that a topic modeling approach is best for analyzing the differences in speaking patterns between bots and other participants, we also include a word-level analysis of the differences in speaking patterns. More specifically, we employ an XGBoost (“extreme gradient boosting”) random forest machine learning model (Chen and Guestrin, 2016), which utilizes the words in the corpus to distinguish between bots and other speakers. XGBoost is an “ensemble” machine learning technique built on decision trees that has recently been used in accounting research to detect fraud (Sun *et al.*, 2021) [6], among other topics, and works by “boosting” tree-level differences in word usages to classify (in our case) speakers. Random forest and neural network approaches to text classification have been recognized to do an excellent job: the problem has been that these machine learning models have been a black box in terms of interpreting the importance of the different features to the model (Molnar, 2020; Hvitfeldt and Silge, 2021). Recent advances such as the feature importance algorithms

contained within XGBoost (Rozemberczki *et al.*, 2022) have made machine learning models more interpretable.

4.1 Bot speaking patterns – topic model results

Figure 2 shows the frequencies for the 25 topics in the 1.4 million bot tweets with Topic 8 being the most prevalent, occurring in 7.5% of tweets and Topic 25 being least prevalent, occurring in 1.9% of tweets.

To provide more context on the topics, Table 1 provides the *high-probability* and *frequent-and-exclusive* words for the 25 topics and Table 2 provides example tweets for each of the topics. To make it easier to understand the prevalence of the different topics, the topics in all tables are sorted from the most to the least prevalent topic within the corpus of 1.4 M bot tweets.

A review of the *high-probability* and *frequent-and-exclusive* words along with the example tweets provide us with a quick summary of the types of topics that #OccupyWallStreet bots talked about. The three most prevalent topics focused on the events that were unfolding on the streets, including in New York as well as elsewhere in the world. The example tweets for these first three topics illustrate how the tweets were a form of “live” reporting:

“police helicopters have closed air space over occupywallstreet preventing all news helicopters from filming what is happening” ... “Tahrir [square] stood in solidarity with occupyoakland and occupywallstreet” ... “remember with every update photo video you take the world is watching.”

The next three most prevalent topics were less about the live events and more philosophical and deliberative in nature. Once again, the individual tweet examples can be strung together and read as a commentary about the antecedents to Occupy Wall Street and the need for change:

“Clinton taxed the rich, created million jobs, record surplus [and] Bush cut taxes, lost million jobs, record deficit” ... “we must resist the rush toward ever increasing state control of our society” ... “people, voice, purpose, goal, mission possible.”

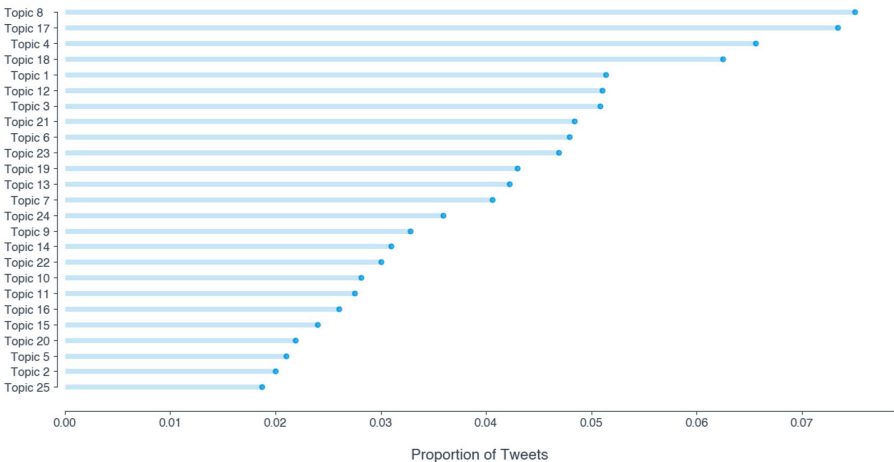


Figure 2.
Topic proportions in
bot tweets

Note(s): Figure shows expected topic proportions within the corpus of 1.4 million tweets sent by bots in English-language tweets sent between August 1, 2011 and July 31, 2012 that contained some variation of the hashtag #Occupy (e.g., #OccupyWallStreet, #OccupyWallSt, #OWS, #OccupyCanada, etc.)
Source(s): Authors’ own work

Twitter bots,
democracy,
and #Occupy-
WallStreet

Topic 8	HPRob: live, video, march, majorityfm, solidar, oakfosh, squar, join FREX: majorityfm, solidar, mayday, mgs, show, march, strike, dontbeaputz
Topic 17	HPRob: polic, protest, park, theother, cop, thinkprogress, nypd, news FREX: spray, pepper, riot, theother, cop, tear, polic, shot
Topic 4	HPRob: occupi, photo, wall, post, protest, osf, opdx, movement FREX: img, post, photo, newyorkc, osf, occupi, paul, opdx
Topic 18	HPRob: thenewd, american, tax, million, vote, economi, ndaa, arrest FREX: million, crash, ndaa, class, bill, rais, tax, bush
Topic 1	HPRob: ronpaul, teaparti, tcot, ronpaulsvoic, amp, state, tlot, freedom FREX: found, form, forc, state, speech, accept, self, seen
Topic 12	HPRob: peopl, ronpaul, govern, liberti, war, teaparti, ronpaulsvoic, constitut FREX: peopl, solv, constitut, grand, voic, reflect, tire, protect
Topic 3	HPRob: temp, forecast, night, high, low, rain, topprog, tcot FREX: temp, forecast, tpp, sunni, ocra, rain, sgp, high
Topic 21	HPRob: nypd, ogundamisi, dustinslaught, protest, offic, shut, just, nigerian FREX: ogundamisi, flash, nigeria, book, lago, gej, remov, nigerian
Topic 6	HPRob: can, one, come, peac, let, make, way, bridg FREX: side, bring, bridg, togeth, possibl, stay, mass, let
Topic 23	HPRob: citizenradio, bank, america, year, last, billion, senatorsand, plan FREX: though, sach, goldman, refund, bank, citizenradio, berni, profi
Topic 19	HPRob: thenewd, don, job, tell, fight, parti, republican, gop FREX: hate, handout, everywher, parti, liber, don, insid, gopwaronwomen
Topic 13	HPRob: need, pleas, help, thank, violenc, rememb, follow, tweet FREX: pleas, rememb, tweet, donat, help, twitter, word, nonviol
Topic 7	HPRob: arrest, report, month, york, chicago, put, old, boston FREX: chicago, nato, happi, month, citibank, john, dispatch, report
Topic 24	HPRob: want, big, think, good, thing, corrupt, read, person FREX: big, institut, corrupt, person, contact, good, want, daddi
Topic 9	HPRob: get, corpor, stop, just, chang, tri, sign, start FREX: tri, corpor, kpop, cours, sign, yes, get, chang
Topic 14	HPRob: today, polit, money, ronpaul, feder, spend, home, politician FREX: spend, polit, reddit, home, reserv, lobbyist, visit, run
Topic 22	HPRob: time, action, nation, everi, continu, alway, destroy, find FREX: action, alway, direct, enough, time, short, continu, everi
Topic 10	HPRob: day, see, end, martin, king, luther, demand, must FREX: king, luther, see, martin, demand, care, begin, day
Topic 11	HPRob: creat, free, countri, without, presid, ronpaul, major, privat FREX: jail, major, privat, market, without, documentari, creat, went
Topic 16	HPRob: mmflint, give, secur, never, work, realli, massiv, dear FREX: massiv, pls, discuss, answer, mmflint, fbi, investig, homeland
Topic 15	HPRob: law, order, econom, court, social, caus, rule, lie FREX: lie, judg, rule, order, suprem, court, law, final
Topic 20	HPRob: livestream, right, viewer, watch, tcoswnuywco, live, check, team FREX: livestream, viewer, watch, tcoswnuywco, videostream, coverag, check, team
Topic 5	HPRob: love, man, govt, sopa, monsanto, antise, wikileak FREX: govt, monsanto, wikileak, antise, learn, feel, eff, beauti
Topic 2	HPRob: right, group, front, away, camp, brooklyn, better, left FREX: shall, front, left, tcoswnqqwt, feed, grant, everyth, done
Topic 25	HPRob: will, take, say, back, just, media, occup, like FREX: take, say, rise, will, school, back, occup, amaz

Note(s): Table shows the *high-probability* (HPRob) and *frequent-and-exclusive* (FREX) words associated with each topic (Roberts *et al.*, 2014a, b). The topics in this and following tables are sorted and ranked from the most to the least prevalent topic within the corpus of 1.4 M bot tweets

Source(s): Authors' own work

Table 1.
Topics (ranked by
most prevalent first)

Table 2.
Sample tweets for
each topic

Topic 8	tahrir stood in solidarity with occupyoakland and occupywallstreet
Topic 17	police helicopters have closed air space over occupywallstreet preventing all news helicopters from filming what's happening
Topic 4	remember with every update photo video you take the world is watching
Topic 18	clinton taxed the rich created million jobs record surplus bush cut taxes lost million jobs record deficit
Topic 1	we must resist the rush toward ever increasing state control of our society
Topic 12	people voice purpose goal mission possible
Topic 3	forecast for nyc monday night through wednesday partly cloudy
Topic 21	min of silence for people who are being shot in nigeria ok
Topic 6	you can't separate peace from freedom because no one can be at peace unless he has his freedom
Topic 23	bank of america received a billion tax refund from the irs last year even though it made billion in profits last year
Topic 19	Don't hate business we want fairness we don't want handouts we want justice
Topic 13	libertysqga kitchen is looking for volunteers please help
Topic 7	mobile tower helped chicago cops during nato protests
Topic 24	how we win. banks have no customers. banks close. corrupt politicians get less votes. they lose. you get the idea
Topic 9	stop at nothing to start something quote on sign in los angeles restaurant
Topic 14	lobbyists wield power over legislators either by promising campaign funds or threatening to support an opponent
Topic 22	in earlier times patriotism meant the willingness courage to challenge gov policies regardless of popular perception
Topic 10	the day we see the truth and cease to speak is the day we begin to die martin luther king
Topic 11	there is no public square it's become private land under our noses
Topic 16	the fbi is tracking your mobile data thanks to carrier
Topic 15	chants of nypd does not respect law and order as cops rush in
Topic 20	cornell west currently visiting occupywallstreet watch live on . . .
Topic 5	if the govt won't stop inequality we will stop the govt
Topic 2	first banner drop for general strike at brooklyn bridge. Nicely done
Topic 25	what was the right wing media saying the other day about ows being dead
Source(s): Authors' own work	

The remaining topics were a mixture of live reporting, deliberative commentary and utterances about the organizational aspects of the physical protests. Included in this amalgam were topics about the weather forecast (“forecast for NYC Monday night through Wednesday partly cloudy”), topics about the public sphere (“there is no public square, it’s become private land under our noses”) and topics about the need for volunteers (“libertysqga kitchen is looking for volunteers, please help”).

The topic model results as well as the example tweets foreground two aspects of bot participation. First, most of the topics do have an ethical stance and normative vision. Furthermore, while this stance and vision is visible in some individual tweets such as “if the govt won’t stop inequality, we will stop the govt,” the ethical stance becomes even more visible when one considers the broader conversation stream. The aforementioned tweet stream examples where we strung together the tweets from the different topics illustrate how a Twitter reader would see the tweets flow by on her/his cellphone. While not all tweets fit perfectly well together – for example, the comment “we must resist increasing state control . . .” is sandwiched between two other tweets – the tweets are closely enough related that an ethical stance is communicated.

Second, the topics and examples highlight that there is not only a variation in topics but also in modes of deliberation and argumentation. The live reporting tweets appear to be less deliberative but, at the same time, communicate an ethical stance. Utterances about solidarity and volunteering do not debate but do communicate a message about what is important:

being pre-figurative (Reinecke, 2018) and giving voice to the notion that one must be the change that one wants to see. Other tweets are minimalist (“people, voice, purpose, goal, mission possible”) but do sketch out both a line of argument and a plan of action. Other tweets are less minimalist, offering a more complete analysis of the problems of current-day forms of government and the need for change. Taken together, the topics and examples illustrate how a social accountability conversation emerges from a seemingly disparate group of tweets.

Twitter bots,
democracy,
and #Occupy-
WallStreet

5. Analysis

5.1 RQ1 – do bots speak differently?

We now turn to analyzing our two primary research questions. This section uses the original tweet data for all users (3.9 M tweets) as well as the entire tweet data (original plus retweeted messages) for all users (9.2 M tweets) to examine whether bots speak differently. To do this, we created a 1/0 indicator variable that was coded 1 if the speaker was a bot (zero otherwise). We also constructed a 1/0 indicator variable for each of the 25 topics based on the *high-probability* and *frequent-and-exclusive* words. We chose five words for each topic from those shown in Table 1 that we thought best represented the topic and which did not appear as words for any of the other 25 topics. The words chosen, when possible, were nouns. This approach to creating a topic variable ensured that the 25 topics were mutually exclusive. If any of the five words for a topic appeared within the tweet, the topic indicator variable was coded one [7].

The above data and variables are used to determine if it is possible to predict whether the speaker is a bot based on the topics mentioned in the tweet. We run two separate analyses: the first with original tweets only and the second that contains all 9.2 M tweets. The results for the original tweets indicate whether there are significant differences in speaking patterns and which type of speaker is more likely to speak about a particular topic (this is indicated by the direction of the individual coefficients as well as the significance of the coefficient). This first set of results allows us to consider what the two groups of speakers talk about when they are sending an original tweet. Velázquez et al. (2017) use a similar method to distinguish between the speaking patterns of bots and non-bots within a data set of 20,000 tweets pertaining to human rights in Mexico.

The second analysis using all 9.2 M tweets is expected to give us slightly different results in that bots can retweet a message from another user as can non-bots. We expect that there will still be significant differences in speaking patterns but that these differences should narrow as the two groups of speakers learn from and copy each other. We are particularly interested to see if the magnitude of the topic coefficients decreases when all the tweets are considered and also whether some coefficients change from negative (i.e. when bots didn’t talk about the topic relative to non-bots) to positive. We assume that a change in the direction of the coefficient indicates that bots learned from the original tweets sent by non-bots and, thus, started to retweet about these topics.

The analyses reported in Table 3 use a series of machine learning classification techniques that seek to predict speaker type (bot or non-bot) based on the speakers’ use of the topics. Because our speaker type variables are relatively stationary speaker traits that are temporal precursors to the selection of character choices (that is, they are “causes” rather than “effects”), the statistics literature suggests that it is best to not use traditional regression methods such as a logit regression (Davis, 1985). However, we can employ recently developed methods from the field of machine learning. There is a growing literature that seeks to identify author characteristics (such as age, gender, regional origin and political orientation) based on Twitter-based textual characteristics (Morgan-Lopez et al., 2017; Rao et al., 2010). Like prior research, we train a machine learning model to predict author characteristics based on characteristics of the text they have written on Twitter.

AAAJ

	Likelihood of bot speaker – original tweets sample		Likelihood of bot speaker – all tweets sample	
	(1)		(2)	
	Relative strength of relationship with bot likelihood (0 → 1)	Direction of relationship with bot likelihood (+/–)	Relative strength of relationship with bot likelihood (0 → 1)	Direction of relationship with bot likelihood (+/–)
Topic 8	0.02	+	0.23	+
Topic 17	0.06	–	0.23	+
Topic 4	0.13	+	0.05	+
Topic 18	0.02	+	0.11	+
Topic 1	0.45	+	0.4	+
Topic 12	0.03	–	0.24	+
Topic 3	1	+	1	+
Topic 21	0.17	–	0.17	+
Topic 6	0.01	+	0.02	+
Topic 23	0.04	–	0.39	+
Topic 19	0.32	–	0.09	–
Topic 13	0.03	–	0.04	–
Topic 7	0.02	–	0.06	+
Topic 24	0.07	+	0.07	+
Topic 9	0.22	+	0.12	+
Topic 14	0.05	+	0.03	+
Topic 22	0	+	0.01	+
Topic 10	0.01	+	0.02	+
Topic 11	0.01	+	0.03	+
Topic 16	0	–	0.59	+
Topic 15	0.01	+	0	–
Topic 20	0.34	+	0.2	+
Topic 5	0.03	+	0.02	+
Topic 2	0.02	–	0	–
Topic 25	0.02	–	0.01	–
N	3,978,103		9,260,886	
Model accuracy	91.3%		83.6%	

Note(s): Table 3 shows results from a support vector machine (SVM) classification model that predicts speaker type (bot or non-bot) based on the speakers’ use of the 25 topics. *Model Accuracy* shows the accuracy of the model in predicting bots on a 20% “hold-out” sample employing 5-fold cross-validation. The *Relative Strength* column shows the SVM coefficients “normalized” on a 0 → 1 scale, such that the variable with the strongest predictive values is given a score of “1,” the variable with the weakest effect is given a score of “0,” and the remaining variables a score between 0 and 1 based on their strength relative to the weakest and strongest variables. The column thus permits a comparison of the relative predictive effect of the 25 topic models. The *Direction of Relationship* column shows the positive or negative direction of each topic’s association with bots

Source(s): Authors’ own work

Table 3. Machine learning predictive models of bot likelihood given speech topics

Table 3.
Machine learning
predictive models of
bot likelihood given
speech topics

To train and test the model we utilized machine learning techniques conducted in the Python programming language, relying heavily on Python’s machine learning library *Scikit-learn* (Pedregosa *et al.*, 2011). Our first step, in line with common machine learning practice (e.g. Amani and Fadlalla, 2017; Guyon and Elisseeff, 2003), is to train and test our classification model, using data on the usage of the characters and concepts (that is, the 25 topics) to classify speakers as bots and non-bots. In line with previous research (Morgan-Lopez *et al.*, 2017; Rao *et al.*, 2010), we employ a classifier in the support vector machine (SVM) family of techniques, namely a linear support vector classification classifier (Ho and Lin, 2012) using Scikit-learn’s linearSVC algorithm with the *L1* norm penalty [8]. Model accuracy for original tweets, as

indicated in Table 3, is over 90% and is based on the accuracy of the model in predicting bots on a 20% “hold-out” sample [9]. Similar to a logit regression, the output of the SVM algorithm identifies which of the independent variables (the 25 topics) are most associated with bots. To facilitate comparisons of the relative predictive power of the variables, we normalize the coefficients shown in the *Relative Strength* column on a 0–1 scale, such that the variable with the strongest effect in each model is given a score of “1,” the variable with the weakest effect is given a score of “0,” and the remaining variables are given a score between 0 and 1 based on their strength relative to the weakest and strongest variables. The *Direction of Relationship* column also indicates the \pm direction of the relationship with bot speakers. Table 3 presents the results for both the original tweet and all tweet (original plus retweeted) segments.

Table 3 results are mostly consistent with our expectations. For example, column 1 (the original tweets only) demonstrates that the speaking patterns of bots differ from non-bots. More specifically, the most prevalent topic was a bot-preferred topic about live on-the-ground reporting (e.g. topic 8), whereas the second-most prevalent topic was a non-bot topic about the police (topic 17). The next three most prevalent topics (topics 4, 18, 1) were bot-preferred topics and consisted of utterances like “remember to update photo videos,” “we must resist . . . increasing state control,” and “Clinton taxed the rich” (see Table 2). While we cannot claim that these example tweets are representative of the topic (because we have no way of selecting a representative example from such a large dataset), Table 3 results indicate that the topics that bots talk about are different than non-bots.

A comparison of column 2 with column 1 shows that, for some topics, the coefficient values narrowed and even reversed when retweets are included. For example, topics that had a negative coefficient in column 1 (meaning that bots were less likely to talk about the topic) changed to having a positive coefficient when the entire corpus was considered. This shift occurred with the second-most most prevalent topic (topic 17, as reflected in the example tweet “Police helicopters . . .”) as well as a half dozen others. Given that column 2 includes both the original tweet data (i.e. the basis for the results in column 1) and the retweet data (i.e. new data), the changed direction of the coefficients implies that bots were overwhelming more likely to retweet these messages than were non-bots.

While bots adjusted their messaging, so did non-bots. Topics such as “we must resist . . . state control” (topic 1) and “if the govt won’t stop inequity . . .” (topic 5) had positive coefficients in column 1 and smaller, positive coefficients in column 2. Like the previous line of reasoning, it appears that non-bots were more likely to retweet these messages than bots.

Taken together, the results document that bots and non-bots speak about different topics but that these differences decrease as both groups retweet the messaging of the other group. This set of results suggests that bot participation does impact on the conversation stream because the topics that bots speak about are different than non-bots. At the same time, the column 2 results show that bots did learn from and did help to disseminate messages about topics that were important to other participants. The next section uses these results as a starting point for examining the different ways that bots influence the conversation stream.

5.2 RQ2 – influencing the conversation stream

The “Problem with Bots” section raised the question as to whether bots change the overall conversation stream trajectory. We suggested that two types of influence are possible: first, that bot participation and that messaging by bots about their preferred topics increases over time, and second that bot-preferred topics are increasingly taken up by non-bots. At the same time, we noted that bots may learn from non-bots and may increasingly speak about non-bot-preferred topics. If this is the case, these learning activities will act as a counterbalance. To foreground these aspects, we consider how the Occupy Wall Street conversation stream changed over time.

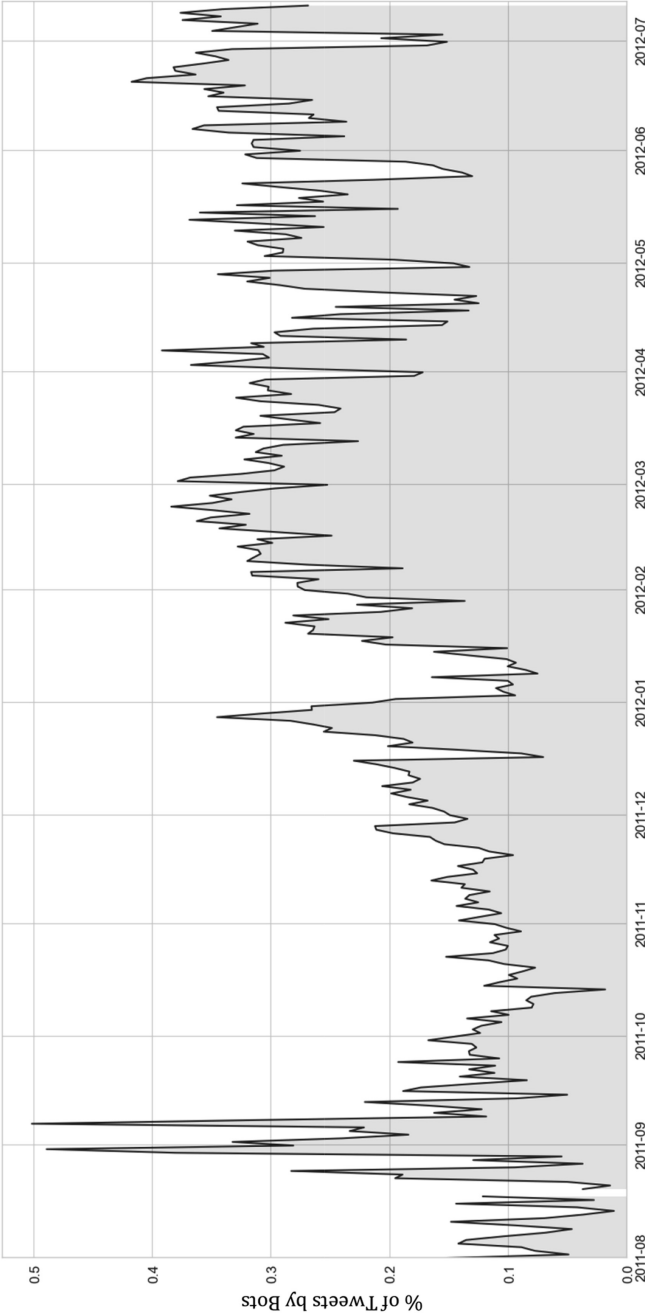


Figure 3.
Percentage of bot
participation per day

Note(s): Figure shows daily proportion of #Occupy tweets (e.g., those containing the hashtag #Occupy, #OccupyWallStreet, #OccupyWallSt, #OWS, #OccupyCanada, etc.) sent by bots over the August 1, 2011 to July 31, 2012 period
Source(s): Authors' own work

Figure 3 shows how the percentage of tweets attributable to bots changes over time: in particular, how the percentage of bot tweets gradually begins to increase after January 15th. This is the same time period where total aggregate tweet volumes were decreasing (see Figure 1). Figure 3 results indicate that, as measured by the percentage of overall tweets, the influence of bot messaging within the conversation stream was increasing. This trend is consistent with our previous suggestion that bots are likely to be more persistent in their tweeting activities than non-bots.

To examine the second type of influence – that is, the impact on non-bot speaking patterns – we provide another set of regressions. In this group of regressions, we re-arrange the data and use *tweetday* (the day the tweet was sent, ranging from 1 to 366) as the dependent variable and the 25 topics as the independent variables. The intuition is that we want to see if the presence of a topic (especially bot-preferred topics) can predict the temporal timing of the tweet. Positive coefficients will indicate that the topic increases in prevalence over time.

Additionally, the regression model includes variables for the topic–speaker interaction for each of the topics. The interaction variables are coded one (zero otherwise) if the tweet contains a particular topic *and* if the tweet was sent by a non-bot. The main effect variable for a topic, along with the interaction variable (topic \times non-bot speaker), tells us whether the prevalence of the topic is increasing/decreasing over time (the main effect) and whether, within this general trend, the prevalence of non-bots speaking with the topic is increasing/decreasing. As mentioned above, we are particularly interested in the topics that had positive coefficients in column 1 of Table 3: that is, the topics that bots were more likely to speak with within the original tweets (i.e. Topics 8, 4, 18, 1). A positive coefficient for the interaction variables within Table 4 for these topics suggest that, over time, the topics preferred by bots are increasingly spoken about by non-bots. Given that we are interested in these interaction variables, and to improve the readability of the table, we exclude the main (non-interacted) effects from Table 4 but do include them in the regression model.

Table 4 also contains two columns. Column 1 only uses the original tweets (the 3.9 M tweets), whereas column 2 uses both original tweets and retweets (the 9.2 M tweets). Column 1 foregrounds what happens to the speaking patterns of non-bots in terms of original tweets only, whereas Column 2 considers how these non-bot speaking patterns change across both original tweets and retweets. We expect that the interaction terms in Column 2 will differ from Column 1 because both bots and non-bots will be learning from, and retweeting, different topics. In this regard, the Column 2 coefficients consider the tweeting and retweeting activities of *both* groups of participants.

Table 4 illustrates how non-bots (i.e. humans) increasingly spoke about some of the bot-preferred topics as time progressed. Column one shows how non-bots were increasingly likely to talk about topics 8, 4 and 18 (but not topic 1), whereas column two indicates that, when all tweets are considered, non-bots are likely to increasingly talk about topics 8 and 18 (but not topics 4 and 1). At the same time, non-bots continued to talk about topic 17, which Table 3 indicated was a non-bot-preferred topic. Taken together, the results suggest that bot participation does influence the conversation stream by influencing the speaking patterns of non-bots. This said, it appears that there are two counterbalances. First, non-bots continue to talk about their most preferred topic (topic 17) and, second, the retweeting activities of bots – especially the retweeting of messages from non-bots – partially mitigates the overall impact on the conversation stream. For these reasons, the overall impact is less than what would otherwise be expected.

5.3 Robustness test – alternative botometer coding of bots

In our theoretical framing, we assumed that speaking with volume was a key distinction between bots and other participants. At the same time, we acknowledge that perhaps not all bots speak with volume. In this section we use an alternative measure for bots to see if the

	Dependent variable: tweetday	
	Original tweets (1)	All tweets (2)
Topic 8 × indiv	34.386*** (0.928)	28.724*** (0.335)
Topic 17 × indiv	9.496*** (2.014)	34.401*** (0.391)
Topic 4 × indiv	2.805*** (0.324)	-6.625*** (0.189)
Topic 18 × indiv	7.195*** (0.636)	5.997*** (0.320)
Topic 1 × indiv	-9.491*** (0.300)	-29.671*** (0.211)
Topic 12 × indiv	-36.041*** (14.633)	-48.364*** (3.049)
Topic 3 × indiv	-2.487*** (0.676)	-25.561*** (0.487)
Topic 21 × indiv	34.945*** (3.421)	24.645*** (0.501)
Topic 6 × indiv	9.874*** (1.040)	1.919*** (0.570)
Topic 23 × indiv	11.697*** (1.514)	-5.479*** (0.363)
Topic 19 × indiv	17.613*** (0.964)	-9.891*** (0.428)
Topic 13 × indiv	2.963*** (1.510)	-3.821*** (0.671)
Topic 7 × indiv	2.303* (1.244)	1.689*** (0.425)
Topic 24 × indiv	2.545*** (0.785)	5.701*** (0.440)
Topic 9 × indiv	20.269*** (0.859)	-5.771*** (0.599)
Topic 14 × indiv	-5.223*** (0.768)	-3.789*** (0.494)
Topic 22 × indiv	4.262*** (0.696)	-3.528*** (0.353)
Topic 10 × indiv	-12.214*** (0.775)	-12.412*** (0.377)
Topic 11 × indiv	27.471*** (1.599)	12.206*** (0.829)
Topic 16 × indiv	10.034*** (1.884)	34.241*** (0.469)
Topic 15 × indiv	12.130*** (1.075)	7.450*** (0.544)
Topic 20 × indiv	-3.448*** (0.599)	-2.223*** (0.399)
Topic 5 × indiv	31.873*** (0.946)	-2.467*** (0.517)
Topic 2 × indiv	-11.087*** (1.271)	-3.012*** (0.495)
Topic 25 × indiv	-8.074*** (0.539)	2.020*** (0.238)
indiv	-52.592*** (0.214)	-24.647*** (0.088)
constant	202.280*** (0.210)	173.106*** (0.083)
Observations	3,978,103	9,260,886
F	8235.36***	13072.07***
Adj. R ²	0.095	0.067

Note(s): * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Model (1) includes original tweets only; Model (2) includes original plus retweeted tweets

Table shows results from OLS regressions of the time variable *tweetday* on the 25 binary topic variables along with the 25 topic variables interacted with *indiv*, a binary variable with values of “1” indicating tweets sent by humans rather than bots. The coefficients on the 25 interaction terms are our coefficients of interest and thus, for ease of presentation, all non-interacted terms have been omitted from the table

Source(s): Authors’ own work

Table 4.
Topic persistence and
the role of individuals
(non-bots)

results change. We re-ran the analyses for original tweets from [Table 3](#) using an alternative, machine learning-based coding of bot likelihood. Specifically, we used Python code to run all conversation participants through the Botometer ([Davis et al., 2016](#)) application programming interface (API). The Botometer machine learning algorithm uses over 1,000 pieces of information from a user’s tweets and Twitter profile to assign a classifier score from 0 to 1, with higher scores indicating a greater likelihood the user is a bot. Based on verifications from the Botometer team and our own manual verifications ([Varol et al., 2017](#); [Wojcik et al., 2018](#)), we used classifier scores of 0.80 and greater as our threshold for considering a user to be a bot. Column 1 of [Table 5](#) contains results from our replication of [Table 3](#)’s test for original tweets. The results are largely consistent with [Table 3](#) and thus confirm our earlier analysis. While we prefer our naïve model due to its simplicity and lower need for “training,” these results do suggest Botometer scores are a useful alternative approach to bot identification. Furthermore, the Botometer coding of bots suggests that some

bots are not high-volume speakers. We return to this observation in the subsequent Discussion section.

Twitter bots,
democracy,
and #Occupy-
WallStreet

5.4 Bot types

One of our core premises is that bots are a new social actor in social accountability narratives. There is evidence in recent research that bots are not a unitary actor in that there are different types of bots (Cresci *et al.*, 2017; Heidari and Jones, 2020; Wojcik *et al.*, 2018). For example, Cresci *et al.* (2017) found evidence of *promoter bots*, *URL spam bots* and *fake followers* – each designed to target different audiences. We provide brief ad hoc analyses here to shed light on whether there are differences in preferred topics for various types of bot actors.

Along with an overall bot likelihood score, the Botometer API includes scores for five different categories of bots: (1) *Astroturf*, politically oriented bots with generally high levels of following and deleting content; (2), *Fake Follower*, bots purchased with the intent to increase the number of followers; (3) *Financial*, bots that post using stock-based *cashtags*; (4) *Self-declared*, bots as declared on botwiki.org, and (5) *Spammer*, bots labeled in existing datasets as being spambots. Columns 2 through 6 of Table 5 contain results for the 5 bot types. The results suggest some noticeable differences in the bot speaking behaviors according to type. For example, financial bots are the only ones more likely to talk about topics 8, 17 and 25. Self-declared bots, meanwhile, have a topic profile quite similar to the average bot, while the astroturf and spammer bots each have half a dozen topics they are more or less likely to focus on. The main takeaway is that bot activity is partially contingent on the type of bot; future research could pick up on this insight and dig into the characteristics of different bot types.

5.5 Word-level distinctions between bots and other participants

The previously provided analysis of bot and non-bot speaking patterns used LDA topic modeling to generate a listing of topics: these identified topics were then used to see if there were differences in speaking patterns. As mentioned previously, we decided to use a topic modeling approach because, similar to prior research, we assumed that conversation is oriented around topics rather than individual words. In this section, we relax this assumption and examine whether the choice of words can distinguish between bots and non-bots. As described earlier, we ran an XGBoost model to determine which *words* can predict whether a speaker was a bot vs a non-bot [10]. The XGBoost model ranks the most important “features” (words) in this prediction model, which in Table 6 we use to make comparisons with our 25 topics. Specifically, Table 6 uses the feature importance results from the XGBoost model to show which words within the previously reported topic model (Tables 1 and 3) were important predictors of bot speaking patterns [11].

To identify influential words vis-a-vis the previously identified topics, we used the 200 most important features from the XGBoost model. The results reported in Table 6 show the topic words that the XGBoost model identified as important along with the word’s ranking position in brackets [12]. For topics where there is a word-level influential feature (e.g. Topic 17, Topic 4, Topic 18, etc.), we can assume that it is these words that are driving the SVM results reported in Table 3. For the topics that do not contain a word-level influential feature (e.g. Topic 8, Topic 13, etc.), we can assume that it is the more holistic-level topic that drives Table 3 results. Taken together, the results reported in Table 6 suggest that the word-level XGBoost model complements the topic modeling results by indicating that sometimes it is individual words that are important and in other cases it is the broader topic that matters. We return to this insight in the discussion section.

Table 5.
Machine learning
predictive models of
likelihood of bot types
given speech topics

Direction of relationship with bot likelihood (+/-)					
All bots	Astroturf	Fake follower	Financial	Self-declared	Spammer
Topic 8	-	-	+	-	-
Topic 17	-	-	+	-	-
Topic 4	-	-	-	-	-
Topic 18	+	+	-	+	-
Topic 1	+	-	-	+	+
Topic 12	-	-	-	-	-
Topic 3	-	-	-	+	+
Topic 21	-	-	-	-	-
Topic 6	+	-	-	-	-
Topic 23	-	-	-	-	-
Topic 19	+	+	-	-	-
Topic 13	+	+	-	-	+
Topic 7	+	+	+	-	+
Topic 24	-	-	-	-	-
Topic 9	+	-	-	+	-
Topic 14	+	+	-	+	-
Topic 22	-	-	-	-	-
Topic 10	+	-	-	-	-
Topic 11	+	-	-	+	-
Topic 16	+	+	-	+	-
Topic 15	+	-	-	+	-
Topic 20	+	-	-	+	-
Topic 5	+	-	-	+	-
Topic 2	-	-	-	+	-
Topic 25	-	-	-	-	-
N	3,978,103	3,978,103	3,978,103	3,978,103	3,978,103
Model accuracy	86.16%	99.44%	99.97%	90.64%	99.09%
Note(s): Table shows results from an SVM classification model that predicts type of bot (see header row for type tested in each column) based on the speakers' use of the 25 topics. <i>Model Accuracy</i> shows the accuracy of the model in predicting bots on a 20% "hold-out" sample employing 5-fold cross-validation. The <i>Direction of Relationship</i> column shows the positive or negative direction of each topic's association with each bot type (for brevity, the SVM coefficients are omitted from the table). The bot types are derived from the Botometer API (Davis et al., 2016)					
Source(s): Authors' own work					

Topic	Overlap of LDA topic words and XGBoost top 200 features
Topic 8	
Topic 17	theother (10), polic (162)
Topic 4	post (5), photo (2), wall (56), movement (91)
Topic 18	tax (37), million (79)
Topic 1	teaparti (4), tcot (14)
Topic 12	govern (26), liberti (54), peopl (29)
Topic 3	temp (30), forecast (7)
Topic 21	
Topic 6	bring (124), peace (90), bridg (61)
Topic 23	bank (178)
Topic 19	gop (35), job (73), liber (54)
Topic 13	
Topic 7	arrest (18), happi (118)
Topic 24	contact (48)
Topic 9	kpop (22), chang (94)
Topic 14	money (123), politician (71)
Topic 22	
Topic 10	
Topic 11	
Topic 16	
Topic 15	court (171)
Topic 20	livestream (25), viewer (9)
Topic 5	monsanto (69)
Topic 2	
Topic 25	

Twitter bots,
democracy,
and #Occupy-
WallStreet

Note(s): Table compares words in our LDA-based topics to words identified through an XGBoost random forest machine learning model (Chen *et al.*, 2016). The XGBoost model tells us which *words* in the entire tweet corpus distinguish bots from other speakers, and Table 6 shows which of the top 200 words from the XGBoost model are also found in our 25 LDA-based topic words (numbers in parentheses indicate XGBoost feature rank). The results allow a comparison of the intersection of the word-level and topic-level analysis of what distinguishes bot from non-bot speaking patterns. For the topics that do not contain a word-level influential feature (e.g. Topic 8), we can assume that it is the more holistic level topic that drives Table 3 results rather than a specific word(s)

Source(s): Authors' own work

Table 6.
Word-level (XGBoost)
and topic model (LDA)
speech patterns

6. Discussion

This study has examined the consequences of Twitter bot participation in the Occupy Wall Street conversation stream. Starting from previous research on dialogic social accountability, we considered two research questions: (1) do bots speak differently than non-bots and (2) does bot participation influence the emergent social accountability conversation? In simple terms, the results indicate that bots do speak differently, exerting both “weak form” and “strong form” influence. This said, the results show that bots also learn from and adapt their speaking patterns to emphasize the topics that are important to non-bots. Furthermore, the speaking patterns of non-bots are influenced by bots but, at the same time, non-bots continue to talk about the topics that are important to them. These last two findings are important counterbalances to concluding that bot participation is unequivocally negative.

The provided findings contribute to our understanding of social media-based social accountability processes. As we mentioned in the introduction, publicly interested accounting researchers have long assumed that accounting-informed dialogic conversations can shift the balance of forces sufficiently to achieve positive social and environmental outcomes (Gray *et al.*, 1997; Neu *et al.*, 2001; Bebbington *et al.*, 2007; Brown and Tregidga, 2017; George *et al.*, 2023). Researchers have also recognized that moments of

technological change simultaneously create the opportunity for stronger and more effective social accountability demands and create new ways for opponents to undermine social accountability conversations (Al Mahameed *et al.*, 2021). Arguably, the emergence of social media created an opportunity, whereas the emergence of bots, other forms of artificial intelligence, and digital forms of surveillance provide new ways to fragment and weaken social accountability. It is within this context that the current study contributes. More specifically, our analysis of the impact of bot participation within the Occupy Wall Street Twitter conversation documents how bots participate and how the speaking patterns of other participants are affected by bot participation. The Occupy Wall Street Twitter conversation is but one social accountability conversation and the tweeting behaviors of bots undoubtedly have become more sophisticated in the last decade. This said, the analysis provides us with a starting point for understanding and assessing the impact of bots on social media-based social accountability conversations.

Second, the study contributes to our understanding of dialogic accounting more generally. Within the social accountability literature, Brown and colleagues have provided us with a series of incredibly detailed and useful studies on the attributes and dynamics of dialogic accounting. These studies implicitly and explicitly acknowledge that the quality of dialogic conversation depends on who participates in the conversation as well as the structural attributes of the conversation venue. The current study complements and extends this research by examining how the participation of a not-totally-human social actor (bots) influences the characteristics of the resultant communication stream, including the impact on the speaking patterns of other participants. We expect that the finding that other participants both change their speaking patterns and continue to talk about topics that motivated them to initially participate in the conversation is relevant to other dialogic accounting settings. Furthermore, the emergence of new forms of AI such as ChatGPT (Li *et al.*, 2023) and deep-fake video and audio recordings (Dack, 2019) suggest that tweets by bots will not be the only not-totally-human participant in dialogic accounting conversations. For these reasons, the current study complements and extends our understandings regarding the future of dialogic accounting.

Despite these contributions, our study and the provided results are shaped and limited by our research choices and our data. For example, we started from the assumption that a key attribute and defining characteristic of bots was that they spoke with volume. This assumption motivated us to define bots based on the number of tweets and resulted in a measure that is easy to understand but that may conflate high-volume bots and high-volume human participants. Furthermore, it resulted in a measure that may obscure the fact that not all bots speak with volume and that not all bots are programmed to speak about the same things. While we still think that this mode of identifying bots is appropriate, the later sections of our analyses relaxed this assumption and used a non-volume-based machine-learning measure to identify bots. These additional tests help validate our assumption about tweet volume as a defining characteristic of bots and mitigate concerns regarding our bot detection method. That said, we must be cautious in noting that our approach only identifies *likely* bots. We also consider the types of bots that appear to be participating in the conversation. These later analyses both round out the provided analysis and draw attention to the importance of digging deeper into the bots that participate in social media-based social accountability conversations. And, as mentioned previously, social media and artificial intelligence have rapidly changed over the last decade. Additional research on more recent social media conversation venues will undoubtedly contribute to our understanding of the impacts of technology on social media-based social accountability phenomena.

Second, the study focuses on differences in tweeting behavior: it was not designed to shed light on how online conversations respond to external events vs influencing an accountability

narrative. We do not consider what gives rise to a specific tweet – whether it is a response to a time-sensitive external event, a comment on the ongoing global economic crisis or a dialogic response to a previous tweet – all three types are presumed to be important for the emergence of the accountability narrative. This said, future research that explicitly considers how different participants respond to external events vs engage in a self-contained “conversation” would be fruitful. Future research could also investigate further characteristics that distinguish bots and bot-like entities from human users, such as variations in timing, preferences for specific offline behaviors and differing responses to “informational” vs “call to action” prompts. Additionally, it would be beneficial to expand upon our findings by examining the possibility of a discernible “periodization” within discussion networks. This could involve identifying phases, like an initial “informing” stage that precedes a “call to action” phase aimed at mobilizing participants for street-level activities. Likewise, future research could explore the issue of the timing of bot activity; for example, do bots wait until some algorithm-driven “triggering event” or threshold level of non-bot activity occurs before jumping heavily in the conversation? And why is it that bots slowly assume a greater share of the conversation activity? Future studies could also pick up on our ad hoc analyses that found distinct speaking patterns of several different types of bots. In effect, while bots appear to play a key role in social accountability narratives, they do not appear to be a unitary social actor, and future research could delve more deeply into this idea.

Finally, it would be well worth building on our study’s insights and extending existing research on the practical implications of bots for democratic societies in such areas ranging from political scandals, democratic deliberation and elections to public health crises (e.g. [Gorodnichenko et al., 2021](#); [Heidari and Jones, 2020](#)). The study also raises the question of whether some type of “algorithmic unfairness” could be contributing to new forms of bot-driven injustices, digital divides or inequalities ([Zhou et al., 2022](#)). Overall, bots are likely to continue to be a “disruptive” social actor ([Hinings et al., 2018](#)), and new accountability tools – and perhaps even regulatory tools – may be needed to help preserve the integrity of democratic participation and public conversations ([Arnaboldi et al., 2017](#)).

In summary, the provided study is both empirical and normative. More specifically, the authors believe in the importance of democratic dialogic conversation and are concerned about the increasing importance of apparently nonhuman actors within different spheres of social life. Prior research, for example, has considered the ethical implications of marketing bots and how such bots influence the actions of human social actors (cf. [Seele et al., 2021](#); [Song et al., 2019](#)). In the spirit of the previous research, the current study encourages us to consider not only the ethical consequences of bots but also the distinction between bots and non-bots. Following from [Goffman \(1981\)](#), we suggested that bots be viewed as a conversation animator – “a speaking machine” – that speaks the sentiments of a human author. We also suggested, following from [Graham and Ackland \(2017\)](#), that bots have the potential to increase the quality of dialogic conversation. Our finding that bots can learn from the utterances of other users is consistent with the view that bots are like an incredibly efficient appendage to the human author who can more quickly summarize and disseminate the gist of the conversation. At the same time, such a conclusion effaces the question as to whether this learning is actually under the supervision of the human author ([Johnson, 2015](#), p. 708) or whether the learning and uttering of bots is a much more “Frankenstein-esque” social actor.

The current study cannot answer this question, but it does draw attention to both the ethical consequences of bot participation within social media-based democratic deliberation and the importance of continuing to study the role of bots within all facets of social life. Like [Castelló and Lopez-Berzosa \(2023\)](#), we believe that the current study contributes to “the debate about social media platforms operating as public spaces for stakeholder engagement and deliberation” (2021, p. 27). This said, future research on the activities of bots, especially

the temporal tightness of the linkage between human authors and their progeny's actions, will help us to better understand the possibilities and dangers associated with the increasing "bot-ization" of social life.

Notes

1. The authors acknowledge that a belief in deliberative democratic processes is perhaps idealistic (Brown and Dillard, 2013, p. 8). At the same time, the authors agree with Butler (2015) that such idealism can have potential perlocutionary consequences for ourselves, the others around us and for broader social change.
2. Prior research distinguishes among bots, cyborgs and humans depending on the degree of automation and human supervision (cf. Chu *et al.*, 2012; Efthimion *et al.*, 2018, p. 4).
3. The email contained an image of a female in a ballerina pose standing on top of a bull along with the following simple call to action: "What is our one demand? #OccupyWallStreet, September 17th. Bring tent."
4. Prior research has used a variety of different naïve, community detection and machine learning techniques to identify bots (e.g. Balaanand *et al.*, 2019; Beskow and Carley, 2018; Heidari and Jones, 2020).
5. LDA topic modeling algorithms use a Bayesian approach to generate the *best division of words given a specified number of topics* (Di Maggio *et al.*, 2013). Specifically, LDA topic models first generate a range of the number of "best" topics for a given corpus (Blei *et al.*, 2003). The researcher must then select the optimal number of topics (the k value), since the LDA topic model generates an optimum solution for a given value of k (which in our case is 25). The process of selecting the optimum k value occurs by utilizing a combination of quantitative and conceptual guides (e.g. DiMaggio *et al.*, 2013). On the quantitative side, it is necessary to utilize a series of LDA diagnostic tests to identify a range of acceptable solutions (Roberts *et al.*, 2014a, b). The statistical optimization diagnostic results contained in the Appendix indicate that the optimal number of topics falls somewhere around 25 topics because this is the best trade-off among the different diagnostic criteria. Following best practices (e.g. DiMaggio *et al.*, 2013) we qualitatively examined each topic model solution across this range of 20–30 topics, including reading random samples of the tweets to assess the alignment between the generated topics and the underlying tweet text, in order to assess the internal consistency and conceptual relevance of the generated topics.
6. As summarized by Sun *et al.* (2021, p. 3), "The basic idea of XGBoost is to combine several tree models . . . into one model with high accuracy."
7. To verify the validity of this approach the authors also ran tests where our topic variables are continuous variables coded based on the presence of any top words regardless of whether they were nouns. These robustness tests overwhelmingly corroborate our main results. For example, compared to our main results for original tweets in Table 3, only one topic undergoes a sign change (Topic 16).
8. The L1 model is also commonly known as the Least Absolute Shrinkage and Selection Operator (LASSO) method (Tibshirani, 1996). LASSO is a form of what are known alternatively as *shrinkage*, *regularization* or *penalization* methods, so-called because they introduce constraints that bias the solution to select fewer features (variables) with the goal of achieving a sparse (or parsimonious) model, one that averts "overfitting" while retaining predictive power through the selection of the most relevant variables.
9. The authors specifically employed *5-fold cross-validation* for the testing, such that the data are split into the 80/20 training/test split five times, and then the five accuracy scores averaged, in order to ensure higher validity of the accuracy score.
10. The 10 most important word features as calculated using the model gain algorithm within XGBoost are *nyc*, *photo*, *http*, *teaparty*, *posted*, *anonymous*, *forecast*, *reet*, *viewers* and *theother*. The XGBoost model diagnostics – a Cohen's Kappa score of 0.59 and an accuracy score of over 90% – indicate the model does a strong job of classifying bot from non-bot speakers.

11. The authors also calculated feature importance using Shapley value methods (see [Rozemberczki et al., 2022](#)). The results for the built-in feature importance algorithm contained in XGBoost and the Shapley value approach are almost identical.
12. Note that the feature importance algorithms do not identify whether the individual features are positively or negatively associated with the dependent variable.

References

- Agarwal, S., Heltberg, R. and Diachok, M. (2009), *Scaling-up Social Accountability in World Bank Operations*, The World Bank, Washington, DC, pp. 1-12, No. 51469.
- Agostino, D. and Sidorova, Y. (2017), "How social media reshapes action on distant customers: some empirical evidence", *Accounting, Auditing & Accountability Journal*, Vol. 30 No. 4, pp. 777-794, doi: [10.1108/aaaj-07-2015-2136](#).
- Al Mahameed, M., Belal, A., Gebreiter, F. and Lowe, A. (2021), "Social accounting in the context of profound political, social and economic crisis: the case of the Arab Spring", *Accounting, Auditing & Accountability Journal*, Vol. 34 No. 5, pp. 1080-1108, doi: [10.1108/aaaj-08-2019-4129](#).
- Amani, F. and Fadlalla, A. (2017), "Data mining applications in accounting: a review of the literature and organizing framework", *International Journal of Accounting Information Systems*, Vol. 24, pp. 32-58, doi: [10.1016/j.accinf.2016.12.004](#).
- Anduiza, E., Cristancho, C. and Sabucedo, J.M. (2014), "Mobilization through online social networks: the political protest of the indignados in Spain", *Information, Communication & Society*, Vol. 17 No. 6, pp. 750-764, doi: [10.1080/1369118x.2013.808360](#).
- Arnaboldi, M., Busco, C. and Cuganesan, S. (2017), "Accounting, accountability, social media and big data: revolution or hype?", *Accounting, Auditing & Accountability Journal*, Vol. 30 No. 4, pp. 762-776, doi: [10.1108/aaaj-03-2017-2880](#).
- Awio, G., Northcott, D. and Lawrence, S. (2011), "Social capital and accountability in grass-roots NGOs: the case of the Ugandan community-led HIV/AIDS initiative", *Accounting, Auditing & Accountability Journal*, Vol. 24 No. 1, pp. 63-92, doi: [10.1108/09513571111098063](#).
- Balaanand, M., Karthikeyan, N., Karthik, S., Varatharajan, R., Manogaran, G. and Sivaparthipan, C. (2019), "An enhanced graph-based semi-supervised learning algorithm to detect fake users on Twitter", *The Journal of Supercomputing*, Vol. 75 No. 9, pp. 6085-6105, doi: [10.1007/s11227-019-02948-w](#).
- Barthold, C., Dunne, S. and Harvie, D. (2018), "Resisting financialisation with Deleuze and Guattari: the case of occupy Wall street", *Critical Perspectives on Accounting*, Vol. 52, pp. 4-16, doi: [10.1016/j.cpa.2017.03.010](#).
- Bebbington, J., Brown, J., Frame, B. and Thomson, I. (2007), "Theorizing engagement: the potential of a critical dialogic approach", *Accounting, Auditing & Accountability Journal*, Vol. 20 No. 3, pp. 356-381, doi: [10.1108/09513570710748544](#).
- Bellucci, M. and Manetti, G. (2017), "Facebook as a tool for supporting dialogic accounting? Evidence from large philanthropic foundations in the United States", *Accounting, Auditing & Accountability Journal*, Vol. 30 No. 4, pp. 874-905, doi: [10.1108/aaaj-07-2015-2122](#).
- Beskow, D. and Carley, K. (2018), "Bot-hunter: a tiered approach to detecting and characterizing automated activity on Twitter", *SBP-BRiMS: International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, Vol. 3, p. 9.
- Blei, D., Ng, A. and Jordan, M. (2003), "Latent Dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.
- Brivot, M., Gendron, Y. and Guénin, H. (2017), "Reinventing organisational control: meaning contest surrounding reputational risk controllability in the social media arena", *Accounting, Auditing & Accountability Journal*, Vol. 30 No. 4, pp. 795-820, doi: [10.1108/aaaj-06-2015-2111](#).

- Broniatowski, D., Jamison, A., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. and Dredze, M. (2018), "Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate", *American Journal of Public Health*, Vol. 108 No. 10, pp. 1378-1384, doi: [10.2105/ajph.2018.304567](https://doi.org/10.2105/ajph.2018.304567).
- Brown, J. (2009), "Democracy, sustainability and dialogic accounting technologies: taking pluralism seriously", *Critical Perspectives on Accounting*, Vol. 20 No. 3, pp. 313-342, doi: [10.1016/j.cpa.2008.08.002](https://doi.org/10.1016/j.cpa.2008.08.002).
- Brown, J. and Dillard, J. (2013), "Critical accounting and communicative action: on the limits of consensual deliberation", *Critical Perspectives on Accounting*, Vol. 24 No. 3, pp. 176-190, doi: [10.1016/j.cpa.2012.06.003](https://doi.org/10.1016/j.cpa.2012.06.003).
- Brown, J. and Tregidga, H. (2017), "Re-politicizing social and environmental accounting through Rancière: on the value of dissensus", *Accounting, Organizations and Society*, Vol. 61, pp. 1-21, doi: [10.1016/j.aos.2017.08.002](https://doi.org/10.1016/j.aos.2017.08.002).
- Brown, J., Dillard, J. and Hopper, T. (2015), "Accounting, accountants and accountability regimes in pluralistic societies: taking multiple perspectives seriously", *Accounting, Auditing & Accountability Journal*, Vol. 28 No. 5, pp. 626-650, doi: [10.1108/aaaj-03-2015-1996](https://doi.org/10.1108/aaaj-03-2015-1996).
- Brown, N., Crowley, R. and Elliott, W. (2020), "What are you saying? Using topic to detect financial misreporting", *Journal of Accounting Research*, Vol. 58 No. 1, pp. 237-291, doi: [10.1111/1475-679x.12294](https://doi.org/10.1111/1475-679x.12294).
- Butler, J. (2015), *Notes toward a Performative Theory of Assembly*, Harvard University Press, Cambridge, MA.
- Butler, J. (2016), "We, the people': thoughts on freedom of assembly", in Badiou, J., Butler, G., Didi-Huberman, S., Khiari, J., Rancière and Bourdieu, P. (Eds), *What is a People?*, Columbia University Press, New York, pp. 49-64.
- Calhoun, C. (2011), "Civil society and the public sphere", in Edwards, M. (Ed.), *The Oxford Handbook of Civil Society*, Oxford University Press, New York, pp. 311-323.
- Castañeda, E. (2012), "The indignados of Spain: a precedent to occupy Wall street", *Social Movement Studies*, Vol. 11 Nos 3-4, pp. 309-319, doi: [10.1080/14742837.2012.708830](https://doi.org/10.1080/14742837.2012.708830).
- Castelló, I. and Lopez-Berzosa, D. (2023), "Affects in online stakeholder engagement: a dissensus perspective", *Business Ethics Quarterly*, Vol. 33 No. 1, pp. 180-215, doi: [10.1017/beq.2021.35](https://doi.org/10.1017/beq.2021.35).
- Castells, M. (2008), "The new public sphere: global civil society, communication networks, and global governance", *The Annals of the American Academy of Political and Social Science*, Vol. 616 No. 1, pp. 78-93, doi: [10.1177/0002716207311877](https://doi.org/10.1177/0002716207311877).
- Chen, T. and Guestrin, C. (2016), "XGBoost: a scalable tree boosting system", *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794.
- Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S. (2012), "Detecting automation of Twitter accounts: are you a human, bot, or cyborg?", *IEEE Transactions on Dependable and Secure Computing*, Vol. 9 No. 6, pp. 811-824, doi: [10.1109/tdsc.2012.75](https://doi.org/10.1109/tdsc.2012.75).
- Cresci, S., Pietro, R., Petrocchi, M., Spognardi, A. and Tesconi, M. (2017), "The paradigm-shift of social spambots: evidence, theories, and tools for the arms race", *Proceedings of the 26th International Conference on World Wide Web Companion*, Perth, April 3-7, 2017, pp. 963-972.
- Dack, S. (2019), *Deep Fakes, Fake News, and what Comes Next*, The Henry M. Jackson School of International Studies, Washington, DC.
- Dahlberg, L. (2005), "The Habermasian public sphere: taking difference seriously?", *Theory and Society*, Vol. 34 No. 2, pp. 111-136, doi: [10.1007/s11186-005-0155-z](https://doi.org/10.1007/s11186-005-0155-z).
- Davis, J. (1985), *The Logic of Causal Order*, Sage, Thousand Oaks, CA.
- Davis, C., Varol, O., Ferrara, E., Flammini, A. and Menczer, F. (2016), "BotOrNot", *Communications of the ACM*, Vol. 59, pp. 273-274.

-
- DiMaggio, P., Nag, M. and Blei, D. (2013), "Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of US government arts funding", *Poetics*, Vol. 41 No. 6, pp. 570-606, doi: [10.1016/j.poetic.2013.08.004](https://doi.org/10.1016/j.poetic.2013.08.004).
- Efthimion, P., Payne, S. and Proferes, N. (2018), "Supervised machine learning bot detection techniques to identify social Twitter bots", *SMU Data Science Review*, Vol. 1, 5.
- Ferrara, E., Varol, O., Davis, C., Menczer, F. and Flammini, A. (2016), "The rise of social bots", *Communications of the ACM*, Vol. 59 No. 7, pp. 96-104, doi: [10.1145/2818717](https://doi.org/10.1145/2818717).
- Fieseler, C. and Fleck, M. (2013), "The pursuit of empowerment through social media: structural social capital dynamics in CSR-blogging", *Journal of Business Ethics*, Vol. 118 No. 4, pp. 759-775, doi: [10.1007/s10551-013-1959-9](https://doi.org/10.1007/s10551-013-1959-9).
- Fraser, N. (1990), "Rethinking the public sphere: a contribution to the critique of actually existing democracy", *Social Text*, Vols 25/26, pp. 56-80, doi: [10.2307/466240](https://doi.org/10.2307/466240).
- Froehling, O. (2013), "Internauts and guerrilleros: the Zapatista rebellion in Chiapas, Mexico and its extension into cyberspace", in Crang, M., Crang, P. and May, J. (Eds), *Virtual Geographies*, Routledge, Milton Park, pp. 173-186.
- George, S., Brown, J. and Dillard, J. (2023), "Social movement activists' conceptions of political action and counter-accounting through a critical dialogic accounting and accountability lens", *Critical Perspectives on Accounting*, Vol. 91, 102408, doi: [10.1016/j.cpa.2021.102408](https://doi.org/10.1016/j.cpa.2021.102408).
- Gerbaudo, P. (2012), *Tweets and the Streets: Social Media and Contemporary Activism*, Pluto Press, London.
- Gerhards, J. and Schäfer, M. (2010), "Is the internet a better public sphere? Comparing old and new media in the USA and Germany", *New Media & Society*, Vol. 12 No. 1, pp. 143-160, doi: [10.1177/1461444809341444](https://doi.org/10.1177/1461444809341444).
- Goffman, E. (1981), *Forms of Talk*, University of Pennsylvania Press, Philadelphia, PA.
- Gomez-Carrasco, P. and Michelon, G. (2017), "The power of stakeholders' voice: the effects of social media activism on stock markets", *Business Strategy and the Environment*, Vol. 26 No. 6, pp. 855-872, doi: [10.1002/bse.1973](https://doi.org/10.1002/bse.1973).
- Gorodnichenko, Y., Pham, T. and Talavera, O. (2021), "Social media, sentiment and public opinions: evidence from #Brexit and #USElection", *European Economic Review*, Vol. 136, 103772, doi: [10.1016/j.eurocorev.2021.103772](https://doi.org/10.1016/j.eurocorev.2021.103772).
- Graham, T. and Ackland, R. (2017), "Do socialbots dream of popping the filter bubble? The role of socialbots in promoting participatory democracy in social media", in Gehl, R.W. and Bakardjieva, M. (Eds), *Socialbots and Their Friends: Digital Media and the Automation of Sociality*, Routledge, New York, pp. 187-206.
- Gray, R., Dey, C., Owen, D., Evans, R. and Zadek, S. (1997), "Struggling with the praxis of social accounting: stakeholders, accountability, audits and procedures", *Accounting, Auditing & Accountability Journal*, Vol. 10 No. 3, pp. 325-364, doi: [10.1108/09513579710178106](https://doi.org/10.1108/09513579710178106).
- Guyon, I. and Elisseeff, A. (2003), "An introduction to variable and feature selection", *Journal of Machine Learning Research*, Vol. 3, pp. 1157-1182.
- Habermas, J. (1962), *Structural Transformation of the Public Sphere* (T. Burger, Trans.) MIT Press, Cambridge.
- Habermas, J. (1984), *The Theory of Communicative Action*, T. McCarthy, Trans., Vols 1 and 2, Polity, Cambridge.
- Hardt, M. and Negri, A. (2011), "The fight for 'real democracy' at the heart of Occupy Wall Street", *Foreign Affairs*, Vol. 11, pp. 301-320.
- Heemsbergen, L.J. (2015), "Designing hues of transparency and democracy after WikiLeaks: vigilance to vigilantes and back again", *New Media & Society*, Vol. 17 No. 8, pp. 1340-1357, doi: [10.1177/1461444814524323](https://doi.org/10.1177/1461444814524323).

- Heidari, M. and Jones, J. (2020), "Using BERT to extract topic-independent sentiment features for social media bot detection", *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, pp. 0542-0547.
- Hinings, B., Gegenhuber, T. and Greenwood, R. (2018), "Digital innovation and transformation: an institutional perspective", *Information and Organization*, Vol. 28 No. 1, pp. 52-61, doi: [10.1016/j.infoandorg.2018.02.004](https://doi.org/10.1016/j.infoandorg.2018.02.004).
- Ho, C.-H. and Lin, C.-J. (2012), "Large-scale linear support vector regression", *Journal of Machine Learning Research*, Vol. 13, pp. 3323-3348.
- Hvitfeldt, E. and Silge, J. (2021), *Supervised Machine Learning for Text Analysis in R*, CRC Press, New York.
- ICIJ (2018), available at: <https://www.icij.org/> (accessed 6 March 2018).
- Johnson, D. (2015), "Technology with no human responsibility?", *Journal of Business Ethics*, Vol. 127 No. 4, pp. 707-715, doi: [10.1007/s10551-014-2180-1](https://doi.org/10.1007/s10551-014-2180-1).
- Juris, J.S., Ronayne, M., Shokooh-Valle, F. and Wengronowitz, R. (2012), "Negotiating power and difference within the 99%", *Social Movement Studies*, Vol. 11 Nos 3-4, pp. 434-440, doi: [10.1080/14742837.2012.704358](https://doi.org/10.1080/14742837.2012.704358).
- Kim, J., Wyatt, R. and Katz, E. (1999), "News, talk, opinion, participation: the part played by conversation in deliberative democracy", *Political Communication*, Vol. 16 No. 4, pp. 361-385, doi: [10.1080/105846099198541](https://doi.org/10.1080/105846099198541).
- Kockelman, P. (2004), "Stance and subjectivity", *Journal of Linguistic Anthropology*, Vol. 14 No. 2, pp. 127-150, doi: [10.1525/jlin.2004.14.2.127](https://doi.org/10.1525/jlin.2004.14.2.127).
- Kockelman, P. (2005), "The semiotic stance", *Semiotica*, Vol. 157, pp. 233-304, doi: [10.1515/semi.2005.2005.157.1.4.233](https://doi.org/10.1515/semi.2005.2005.157.1.4.233).
- La Torre, M., Di Tullio, P., Tamburro, P., Massaro, M. and Rea, M. (2022), "Calculative practices, social movements and the rise of collective identity: how #istayathome mobilised a nation", *Accounting, Auditing & Accountability Journal*, Vol. 35 No. 9, pp. 1-27, doi: [10.1108/aaaj-08-2020-4819](https://doi.org/10.1108/aaaj-08-2020-4819).
- Li, X., Zhang, Y. and Malthouse, E.C. (2023), "A preliminary study of ChatGPT on news recommendation: personalization, provider fairness, fake news", arXiv preprint arXiv: 2306.10702.
- Loughran, T. and McDonald, B. (2016), "Textual analysis in accounting and finance: a survey", *Journal of Accounting Research*, Vol. 54 No. 4, pp. 1187-1230, doi: [10.1111/1475-679x.12123](https://doi.org/10.1111/1475-679x.12123).
- Manning, P. and Gershon, I. (2013), "Animating interaction", *HAU: Journal of Ethnographic Theory*, Vol. 3, pp. 107-137, doi: [10.14318/hau3.3.006](https://doi.org/10.14318/hau3.3.006).
- Mohr, J. and Bogdanov, P. (2013), "Introduction—topic models: what they are and why they matter", *Poetics*, Vol. 41 No. 6, pp. 545-569, doi: [10.1016/j.poetic.2013.10.001](https://doi.org/10.1016/j.poetic.2013.10.001).
- Molnar, C. (2020), *Interpretable Machine Learning*, Leanpub.com, Victoria, BC.
- Morgan-Lopez, A., Kim, A., Chew, R. and Ruddell, P. (2017), "Predicting age groups of Twitter users based on language and metadata features", *PLoS One*, Vol. 12 No. 8, e0183537, doi: [10.1371/journal.pone.0183537](https://doi.org/10.1371/journal.pone.0183537).
- Neu, D., Cooper, D. and Everett, J. (2001), "Critical accounting interventions", *Critical Perspectives on Accounting*, Vol. 12 No. 6, pp. 735-762, doi: [10.1006/cpac.2001.0479](https://doi.org/10.1006/cpac.2001.0479).
- Neu, D., Saxton, G., Everett, J. and Rahaman, A. (2020), "Speaking truth to power: twitter reactions to the Panama Papers", *Journal of Business Ethics*, Vol. 162 No. 2, pp. 473-485, doi: [10.1007/s10551-018-3997-9](https://doi.org/10.1007/s10551-018-3997-9).
- Neu, D., Saxton, G. and Rahaman, A. (2022), "Social accountability, ethics, and the occupy Wall street protests", *Journal of Business Ethics*, Vol. 180 No. 1, pp. 17-31, doi: [10.1007/s10551-021-04795-3](https://doi.org/10.1007/s10551-021-04795-3).

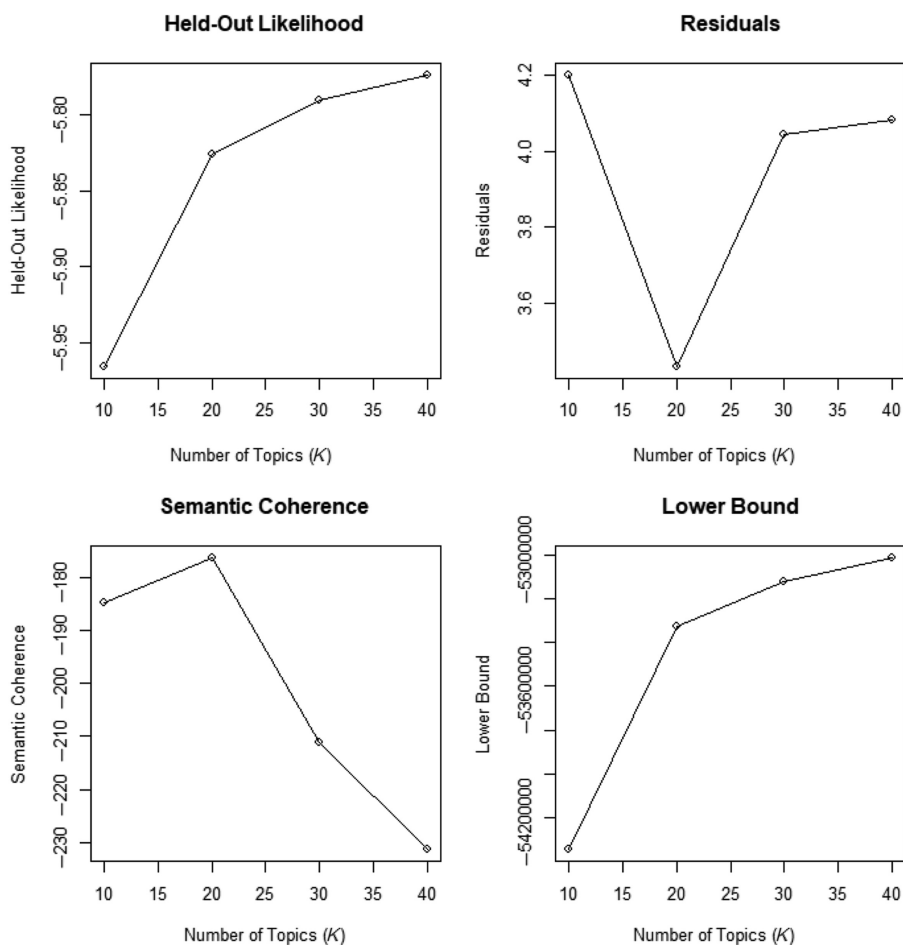
-
- NPR (2011), "Occupy Wall Street inspires worldwide protests", available at: <https://www.npr.org/2011/10/15/141382468/occupy-wall-street-inspires-worldwide-protests> (accessed 5 January 2021).
- O'Meally, S. (2013), *Mapping Context for Social Accountability: A Resource Paper*, World Bank, Washington, DC.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011), "Scikit-learn: machine learning in Python", *Journal of Machine Learning Research*, Vol. 12, pp. 2825-2830.
- Rao, D., Yarowsky, D., Shreevats, A. and Gupta, M. (2010), "Classifying latent user attributes in Twitter", *International Conference on Information and Knowledge Management, Proceedings*, pp. 37-44.
- Ravazzani, S. and Mazzei, A. (2018), "Employee anonymous online dissent: dynamics and ethical challenges for employees, targeted organisations, online outlets, and audiences", *Business Ethics Quarterly*, Vol. 28 No. 2, pp. 175-201, doi: [10.1017/beq.2017.29](https://doi.org/10.1017/beq.2017.29).
- Reinecke, J. (2018), "Social movements and prefigurative organizing: confronting entrenched inequalities in Occupy London", *Organization Studies*, Vol. 39 No. 9, pp. 1299-1321, doi: [10.1177/0170840618759815](https://doi.org/10.1177/0170840618759815).
- Roberts, J. (1991), "The possibilities of accountability", *Accounting, Organizations and Society*, Vol. 16 No. 4, pp. 355-368, doi: [10.1016/0361-3682\(91\)90027-c](https://doi.org/10.1016/0361-3682(91)90027-c).
- Roberts, M., Stewart, B., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S., Albertson, B. and Rand, D. (2014a), "Structural topic models for open-ended survey responses", *American Journal of Political Science*, Vol. 58 No. 4, pp. 1064-1082, doi: [10.1111/ajps.12103](https://doi.org/10.1111/ajps.12103).
- Roberts, M., Stewart, B. and Tingley, D. (2014b), "stm: R package for structural topic models", *Journal of Statistical Software*, Vol. 10, pp. 1-40.
- Rodrigue, M. (2014), "Contrasting realities: corporate environmental disclosure and stakeholder-released information", *Accounting, Auditing & Accountability Journal*, Vol. 27 No. 1, pp. 119-149, doi: [10.1108/aaaj-04-2013-1305](https://doi.org/10.1108/aaaj-04-2013-1305).
- Rozemberczki, B., Watson, L., Bayer, P., Yang, H., Kiss, O., Nilsson, S. and Sarkar, R. (2022), "The Shapley value in machine learning", arXiv preprint arXiv:2202.05594.
- Sabadoz, C. and Singer, A. (2017), "Talk ain't cheap: political CSR and the challenges of corporate deliberation", *Business Ethics Quarterly*, Vol. 27 No. 2, pp. 183-211, doi: [10.1017/beq.2016.73](https://doi.org/10.1017/beq.2016.73).
- Saxton, G.D., Ren, C. and Guo, C. (2021), "Responding to diffused stakeholders on social media: connective power and firm reactions to CSR-related Twitter messages", *Journal of Business Ethics*, Vol. 172 No. 2, pp. 229-252, doi: [10.1007/s10551-020-04472-x](https://doi.org/10.1007/s10551-020-04472-x).
- Seele, P., Dierksmeier, C., Hofstetter, R. and Schultz, M. (2021), "Mapping the ethicality of algorithmic pricing: a review of dynamic and personalized pricing", *Journal of Business Ethics*, Vol. 170 No. 4, pp. 697-719, doi: [10.1007/s10551-019-04371-w](https://doi.org/10.1007/s10551-019-04371-w).
- She, C. and Michelon, G. (2019), "Managing stakeholder perceptions: organized hypocrite in CSR disclosures on Facebook", *Critical Perspectives on Accounting*, Vol. 61, pp. 54-76, doi: [10.1016/j.cpa.2018.09.004](https://doi.org/10.1016/j.cpa.2018.09.004).
- Song, R., Kim, H., Lee, G. and Jang, S. (2019), "Does deceptive marketing pay? The evolution of consumer sentiment surrounding a pseudo-product-harm crisis", *Journal of Business Ethics*, Vol. 158 No. 3, pp. 743-761, doi: [10.1007/s10551-017-3720-2](https://doi.org/10.1007/s10551-017-3720-2).
- Sun, Y., Ma, Z., Zeng, X. and Guo, Y. (2021), "A predicting model for accounting fraud based on ensemble learning", *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, pp. 1-5.
- Tibshirani, R. (1996), "Regression shrinkage and selection via the LASSO", *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 58 No. 1, pp. 267-288, doi: [10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x).

-
- Tufekci, Z. (2017), *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, Yale University Press, New Haven, CT.
- UNDP (2013), *Reflections on Social Accountability: Catalyzing Democratic Governance to Accelerate Progress toward the Millennium Development Goals*, United Nations Development Programme, New York, NY.
- Varol, O., Ferrara, E., Davis, C., Menczer, F. and Flammini, A. (2017), "Online human-bot interactions: detection, estimation, and characterization", *Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017*.
- Velázquez, E., Yazdani, M. and Suárez-Serrato, P. (2017), "Socialbots supporting human rights", arXiv preprint arXiv:1710.11346.
- Witschge, T. (2004), "Online deliberation: possibilities of the Internet for deliberative democracy", in Shane, P.M. (Ed.), *Democracy Online*, Routledge, New York, pp. 129-142.
- Wojcik, S., Messing, S., Smith, A., Rainie, L. and Hitlin, P. (2018), *Bots in the Twittersphere: Methodology*, Pew Research Center, Washington, DC, available at: <https://www.pewresearch.org/internet/2018/04/09/bots-in-the-twittersphere-methodology> (accessed 2 April 2022).
- Zhou, N., Zhang, Z., Nair, V., Singhal, H. and Chen, J. (2022), "Bias, fairness and accountability with artificial intelligence and machine learning algorithms", *International Statistical Review*, Vol. 90 No. 3, pp. 468-480, doi: [10.1111/insr.12492](https://doi.org/10.1111/insr.12492).

Corresponding author

Gregory D. Saxton can be contacted at: gsaxton@yorku.ca

Diagnostic Values by Number of Topics



Note(s): Identification of the optimal number of topics in a corpus requires a combination of statistical optimization diagnostics and researcher familiarity with the data, the topic, and prior research (DiMaggio *et al.*, 2013). The above tests show that the optimal number of topics falls somewhere between 15 and 35 topics. As noted in the main text, our review of the discursive themes across topic model solutions showed that the $k = 25$ topic model represented the best trade-off among the different diagnostic results

Source(s): Authors' own work

Figure A1.
Latent Dirichlet topic
modeling –
diagnostics tests