UNDERSTANDING BOTS IN SOCIAL MOVEMENTS

by

CAROLINA SALGE

(Under the Direction of Elena Karahanna)

Abstract

Despite the prevalence and potential impact of bots—automated accounts in online social networks—on organizations and society, scant attention has been paid to understanding the effects that they have in social networks. This research focuses on examining bots in the context of social movements. The first study explicates the discovery of bots as central actors in a political protest on Twitter and discusses how neglecting them in online social network studies can threaten research validity. The paper also offers a multi-method approach that scholars can use to best detect bots. The second study explores the role of bots in social movements and derives a theoretical model of *BrokeCasting* (broadcasting + brokerage) which can be applied by social movement organizations and activists to raise awareness of a social cause on social media. The third paper uses topic analysis to identify patterns in the content disseminated by bots during a social movement and proposes a framework of *Botivism* for mobilizing resources on social media. Together these essays shed light on the synergies between bots and social change.

INDEX WORDS: bots; social movements; Twitter; research validity; tagging features; topic modeling; resource mobilization

UNDERSTANDING BOTS IN SOCIAL MOVEMENTS

by

CAROLINA SALGE

B.A., Clemson University, 2008

M.A., Clemson University, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

UNDERSTANDING BOTS IN SOCIAL MOVEMENTS

by

CAROLINA SALGE

Major Professor:	Elena Karahanna
Committee:	Nicholas Berente Weifeng Li Hani Safadi Richard T. Watson

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction & Literature Review

Bots (a short term for social robots) are automated accounts in online social networks (Morstatter, Wu, Nazer, Carley, & Liu, 2016) which have garnered a lot of media attention recently. Much of this attention is unfortunately negative. Facebook and Twitter, for example, were asked last year to testify to Congressional Intelligence Committees about bots and the roles that they played during the 2016 presidential election in the United States. Research suggests that bots were used to manipulate public opinion and thus threaten democracy (Bessi & Ferrara, 2016). While this example emphasizes the importance and potential impact of bots on society, the reality is that we still know very little about them. As Gorwa and Guilbeault (2018) write, "we need to better understand bots before we can really research and write about them" (p. 4). These three essays shed light on this nascent research domain by examining the synergies between bots and social movements.

Bots & Social Movements

Bots can be based on rules or machine learning. Rule bots are quite limited—they can only respond to predefined commands in a script. *101-a-tron* represents an example. When authorized users ask the bot to inform others on something—for example, "@101atron, tell @thisUser about human rights"—it autonomously replies with a link associated with the "human rights" key phrase.[1] Machine-learning bots are different than bots like *101-a-tron* since they understand language, not just predefined commands, and continuously learn from conversational

[1] To learn more, please see https://github.com/FeelTrainCoop/101atron

data fed to them. An example here is Microsoft's Tay. The bot is a project Microsoft developed to "experiment with and conduct research on conversational understanding[2]."

Regardless of whether they are governed by rules or machine learning, bots have a variety of attributes (Freitas, Benevenuto, Ghosh, & Veloso, 2014; Savage, Monroy-Hernández, & Hollerer, 2016). For instance, the account profile of a bot represents its description on a platform. On social media, this usually includes a username, biography, join date, location, photo, number of messages they generate, number of subscribers they have, and number of users they subscribe to. Bots are also designed for a purpose—i.e., created to achieve a particular objective. They can be built to answer sociology 101 type of questions (e.g., *101-a*-tron); empower adherents of a social cause (e.g., *StayWokeBot*); damage an organization's reputation (Messias, Schmidt, Oliveria, & Benevenuto 2013); and more recently, spread fake news in political elections (Shao, Ciampaglia, Varol, Flammini, & Menczer, 2017). Script is another design attribute of bots. A script corresponds to a list of executable commands crafted by bot developers and intended to enable them in fulfilling their purpose. To answer sociology 101 type of questions, for example, *101-a-tron* functions as a knowledge management system—it first detects what activists are asking for and then replies with a link in just a few minutes. Finally, bots operate in many different platforms. Twitter and Facebook—the two most popular ones for social networking[3]—have millions of them (Goldman, 2014; Grant, 2014; Varol et al., 2017).

Scholars have spent a great amount of time and effort in detecting bots (e.g., Cresci, Di Pietro, Petrocchi, Spognardi, & Tesconi, 2015; Davis, Varol, Ferrara, Flammini, & Menczer, 2016; Salge & Karahanna, 2018). In doing so, they have employed a variety of techniques including social networks (e.g., Paradise, Puzis, & Shabtai 2014), human crowdsourcing (e.g.,

---

[2] To learn more about *Tay* please see https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/

[3] http://www.alexa.com/topsites/category/Computers/Internet/On_the_Web/Online_Communities/Social_Networking

Cao, Yang, Yu, & Palow 2014), machine learning (e.g., Davis et al. 2016), and activity correlation (e.g., Chavoshi, Hamooni, & Mueen, 2016). Recent research has taken a more comprehensive approach to detection. Salge and Karahanna (2018), for example, introduce the funnel process, a method combining three techniques (social networks, computer algorithms, and human crowdsourcing) into a procedure scholars can use to more accurately detect bots[4].

This dissertation focuses on bots and social movements. We define social movements as a set of opinions and beliefs in populations that represent preferences for changing some elements of social structures (McCarthy & Zald, 1977). Although scholars have extensively studied social movements[5], there are only a few studies on bots in social movement networks (Abokhodair, Yoo, & McDonald, 2015; Verkamp & Gupta, 2014; Suárez-Serrato, Roberts, Davis, & Menczer, 2016; Savage, Monroy-Hernández, & Hollerer, 2016)[6]. Because social media is a germane platform for activists to promote social change (Tufecki, 2017), all of the work in this literature is conducted on Twitter. These studies are primarily empirically driven and collectively suggest that bots are designed to amplify the magnitude of protests and encourage social movement participation. For example, Savage et al. (2016) design bots to call volunteers to action and find that more than 80 percent of those infected actually responded to bot messages. This is promising and disturbing. On one hand, it is promising because it allows us to theorize about bots and uncover the mechanisms by which they, for example, raise social movement awareness and mobilize resources for a cause, and thus provide a significant and positive contribution to social change. On the other hand, it is disturbing since it indicates that

[4] Other holistic approaches are either unavailable to the public (e.g., CopyCatch (Beutel, Xu, Guruswami, Palow, & Faloutsos 2013) and SynchroTrap (Cao et al. 2014)) or only appropriate for data following a sequence of click events (e.g., Renren Sybil (Wang, Konolige, Wilson, Wang, Zheng, & Zhao 2013)).

[5] See Edwards (2014) for an extensive review.

[6] We also reviewed the literature on bots and politics (e.g., Bessi & Ferrara, 2016). We do not include some of these studies here since they do not study political movements but instead political election campaigns or debates.

people are clearly vulnerable to bot manipulation and propaganda, which means that we need to continue to strengthen our approaches of bot identification. The remaining social movement studies endorse this double-edge sword in bot research yet they lack meaningful insights to theory which can, in turn, be used to inform detection. Verkamp and Gupta (2014), for example, “find that the nature of [bot] spam varies significantly across [protest] incidents such that [theoretical] generalizations are hard to draw” (p. 1).

To integrate, our review of the literature suggests that bots are governed by rule or machine learning, have a profile, an objective, and a script, and also exist in a variety of platforms. The literature also shows that scholars have spent a great amount of time and effort in identifying bots yet we know that “bot detection will always be a cat-and-mouse game in which a large, but unknown, number of humanlike bots may go undetected” (Lazer, Baum, Benkler, Berinsky, Greenhill, Menczer, Metzger, Nyhan, Pennycook, Rothschild, Schudson, Sloman, Sunstein, Thorson, Watts, & Zittrain, 2018). While scholars have also begun to study bots in social movements, our review shows that this line of work is still in its infancy. Thus, despite the prevalence and potential impact of bots, a theoretical account of the synergies between bots and social change at a granular level is still lacking.

Contributions of the Dissertation Papers

Overall, the papers of this dissertation provide a significant contribution in our understanding of bots in social movements. In the first paper, we problematize an implicit assumption of online social network research in Information Systems and Management as it pertains to the concepts of actors. The implicit assumption is that actors in online social networks are humans and theorizing in many studies relied on the fact that actors are human. As such, we discuss how neglecting bots can threaten the validity of research and we position bot as a crucial actor with implications for organizational theory and research. We also contribute to bot detection by providing a multi-method approach that scholars can use to best identify bots.

In the second paper, we inductively study the role of bots in social movements. Our results show that they act as *BrokeCasters*—i.e., as brokers and broadcasters of a cause. We provide a contribution to theory by abstracting our results to a framework of *BrokeCasting* which identifies two mechanisms by which bots raise awareness of a social movement on social media. The first mechanism, *Actor Tagging Volume*, relies on effective network volume and is part of broadcast. *Actor Tagging Volume* requires bots to leverage actor tagging features of social media, such as mentions, to share few messages and reach many non-redundant actors. The second mechanism is part of brokerage and relies on *Content Tagging Diversity*, a means by which bots advantageously use content tagging features of social media, such as hashtags, to jointly access information associated and *not* associated with social movements. SMOs and activists can apply our model to significantly contribute to their causes and scholars can use *BrokeCasting* to improve the accuracy to which we detect bots. From a theoretical perspective, we enhance our understanding of how bots raise awareness of a cause by explicating new theoretical mechanisms and by defining a new role. We also highlight the significance of content and content-tagging features of social media which have been largely ignored in prior work since the literature in this area is primarily content-agnostic.

Focusing on content, in the third paper, we use topic analysis to examine the messages generated by bots in a social movement on Twitter. In doing so, we discover two forms of bot activism. The first we call *Focal Botivism* and define it as a means by which a group of bots share positive and negative messages in correlated topics which altogether promote social change in a specific population and for a specific cause. The second form is *Global Botivism*, a means by which a group of bots share positive and negative messages in correlated topics which altogether promote social change across the globe. We conclude the paper with a theoretical model of *Botivism* which explains how bot networks mobilize resources for social movements on social media. Our objective is that these papers provide useful bot detection approaches and a solid theoretical foundation for future research in the area.

Organization

This dissertation is organized in the following fashion. Chapter 2 presents the first paper which is the bot discovery in a political protest on Twitter. Chapter 3 introduces the second paper which examines the role of bots in a social movement. Chapter 4 presents the third paper which explores the content disseminated by bots in a social movement and explains how they mobilize resources for a social cause on social media. Finally, Chapter 5 summarizes the results of all the dissertation papers.

References

Bessi, A., & Ferrara, E. 2016. Social bots distort the 2016 us presidential election online discussion. ***First Monday*** 21(11).

Cao, Q., Yang, X., Yu, J., & Palow, C. 2014. Uncovering large groups of active malicious accounts in online social networks. ***In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM****, 477–488, doi: http://dx.doi.org/10.1145/2660267.2660269.*

Chavoshi, N., Hamooni, H., & Mueen, A. 2016. DeBot: Twitter bot detection via warped correlation. In ***Proceedings of the IEEE International Conference on Data Mining***.

Cresci, S., Di Pietro, R., Pretocchi, M., Spognardi, A., & Tesconi, M. 2015. Fame for sale: Efficient detection of fake Twitter followers. ***Decision Support Systems****, 80: 56–71, doi: 10.1016/j.dss.2015.09.003*.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. 2016. BotOrNot: A system to evaluate social bots. ***Working Paper.*** 2 pages.

Gorwa, R. & Guilbeault, D. 2018. Understanding bots for policy and research: Challenges, methods, and solutions. ***Working Paper***. https://arxiv.org/pdf/1801.06863.pdf

Freitas, C. A., Benevenuto, F., Ghosh, S., & Veloso, A. 2014. Reverse engineering socialbot infiltration strategies in Twitter. ***arXiv preprint arXiv:1405.4927***, 25–32, doi: 10.1145/1851275.1851226.

Goldman, D. 2014. 23 million Twitter users are fed by robots. http://money.cnn.com/2014/08/12/technology/social/twitter-bots/. Retrieved 5/17/2016.

Grant, R. 2014. Facebook has no idea how many fake accounts it has – but it could be nearly 140M. http://venturebeat.com/2014/02/03/facebook-has-no-idea-how-many-fake-accounts-it-has-but-it-could-nearly-140m/. Retrieved 5/17/2016.

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

McCarthy, J. D., and Zald, M. N. 1977. Resource mobilization and social movements: A partial theory. ***American Journal of Sociology,*** *82:6, 1212–1241, doi: 10.1086/226464*.

Messias, J., Schmidt, L., Oliveira, R., & Benevenuto, F. 2013. You followed my bot! Transforming robots into influential users in Twitter. ***First Monday*** *18:7, doi: 10.5210/fm.v18i7.4217.*

Morstatter, F., Wu, L., Tahora, N. H., Carley, K., & Liu, H. 2016. A new approach to bot detection: Striking the balance between precision and recall. ***ASONAM.***

Paradise, A., Puzis, R., & Shabtai, A. 2014. Anti-reconnaissance tools: Detecting targeted socialbots. ***Internet Computing,*** *18:5 11–19, doi: 10.1109/MIC.2014.81*.

Salge C., & Karahanna, E. 2018. Protesting corruption on Twitter: Is it a bot or is it a person? ***Academy of Management Discoveries,*** *4,* 32-49.

Savage, S., Monroy-Hernandez, A., & Hollerer, T. 2016. Botivist: Calling volunteers to action using online bots. ***In Proceedings of the 19th ACM Computer-Supported Collaborative Work and Social Computing****, 813–822,* ACM.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & and Menczer, F. 2017. The spread of misinformation by social bots. ***Technical Report*** 1707.07592, arXiv.

Suarez-Serrato, P., Roberts, M. E., Davis, C., & Menczer, F. 2016. On the influence of social bots in online protests. ***In Social Informatics, Lecture Notes in Computer Science***, 269–278. Springer.

Tufekci, Z. 2017. ***Twitter and tear gas: The power and fragility of networked protest***. London, UK: Yale University Press.

Verkamp, J. P. & Gupta, M. 2013. Five incidents, one theme: Twitter spam as a weapon to drown voices of protest. ***In FOCI***.

Chapter 2. Protesting Corruption On Twitter: Is It A Bot Or Is It A Person?[7]

[7] Salge, Carolina & Karahanna, Elena. Accepted by Academy of Management Discoveries. Reprinted here with permission of publisher, April 19 of 2018.

Abstract

In studying how activists use technology to express public dissatisfaction online, we discover that what we assumed to be human protestors were in some cases bots—automated accounts in online social networks. To explicate the discovery of bots, we problematize an implicit assumption of online social network research within Management and Information Systems as it pertains to the concept of actors. Our discovery takes place in the context of a 6-day inductive case study of a protest against government corruption in Brazil—the Mensalão scandal. We elaborate on how bots were detected and discuss how they are coded to amplify the magnitude of the protest on Twitter. Further, we explore the application of bots beyond the context of our study by illustrating how they were used to increase revenue in the business of online dating and to manipulate public opinion during an election campaign. We also discuss how neglecting bots can threaten research validity and, as a result, we provide scholars investigating social phenomena online with a multi-method approach for detecting bots. Finally, we position bot as a crucial actor with implications for organizational theory and practice.

Keywords: social bots, actor network centrality, Twitter, online protests, online social networks, online social movements.

The *Mensalão* was a vote-buying case of corruption that almost collapsed the Brazilian government of Luiz Inácio Lula da Silva in 2005 (*BBC*, 2013). The scandal broke when Roberto Jefferson, the president of an allied party, announced in a newspaper interview that the Worker's Party (PT) was using public funds to buy political support for the then-Lula government and to pay off debts from election campaigns. Each congressman was receiving about R$30,000 a month (around $12,000 at the time) (*The Economist*, 2013). The allegation led to the downfall of several congressmen and senior members of the government, including José Dirceu, Lula's chief of staff and the alleged mastermind behind the case, Delúbio Soares, PT's treasurer, and José Genoino, PT's former president. In August 2007, the Supreme Federal Court, responsible for investigating cases against parliamentarians, accepted the indictments of 40 deputies involved in the Mensalão scandal. The trial began in August 2012 and roughly two months later, 25 out of the 40 defendants were charged with several crimes ranging from embezzlement and corruption to conspiracy and misuse of public funds. Mr. Dirceu was among the 25 prosecuted deputies. He was sentenced to spend 10 years and 10 months in jail. The court's decision was celebrated by many and marked the beginning of a new era, where those involved in government corruption would be held accountable for their transgressions (Leahy, 2012). But the celebration was brief. Brazil's legal system is "a loophole-ridden oddity, allowing appeals even against supreme-court rulings" (*The Economist*, 2013). On September 18th, 2013 (about a year after the sentence), the Supreme Federal Court accepted—in a 6-5 vote—a motion to hear a new round of appeals from 12 deputies charged in the Mensalão case. This result frustrated and angered many Brazilians (Singer, 2013; Lyons & Cowley, 2013) who were indignant with the justices for giving corrupted politicians a second chance. Motivated by rage, thousands started a corruption protest on Twitter.

Protests are "organized, collective, and public expressions of discontent" (King & Soule

2007:415). They reflect public action. Instead of reaching out to higher authorities with expressions of grievance and desires of maintaining conversations private, activists vent their dissatisfaction openly to a broader audience. Protests initiated on Twitter are significant because they can influence public discourse and, as a result, shape both civic and political engagement (Schumann, 2014). As Tufekci (2014) suggests, referring to the Ferguson protests, "what happens to #Ferguson [on Twitter] affects what happens to Ferguson" (Tufekci, 2014, para. 36). Yet we know little about the central actors on Twitter protests. The purpose of this paper, therefore, is to discover who these actors are and what they have in common. Central actors are the "most important" nodes of a network since they are "located in strategic locations within the network" (Wasserman & Faust 1994:169). We focus on the corruption protest that emerged after the Supreme Federal Court accepted to hear a new round of appeals for the Mensalão case.

In investigating this phenomenon, we discover that actors occupying central positions in online social networks are not always people but instead, they can also be *bots* (short for *social robots*)—automated accounts in online social networks (Morstatter, Wu, Nazer, Carley, & Liu, 2016). We reveal how these bots were detected and show how they were designed to amplify information on Twitter. Our discovery suggests that both humans and bots can be *central* when engaging in online activism. Existing Management and Information Systems literature does not account for this bot phenomenon. In light of this, we take a problematization approach (Alvesson & Sandberg, 2011) by questioning the implicit assumption that *actors* of online social networks are *people*. Our study contributes to the emerging literature in online social networks as we argue that the conceptualization of *actors* should not be constrained to people or organizations, but rather, it needs to be expanded to also include bots. We discuss how neglecting bots imposes threats to research validity and, as a result, we urge scholars investigating social phenomena online to carefully consider bot implications when designing their studies.[8] In addition, our problematization of a field assumption (Alvesson & Sandberg, 2011) opens to scrutiny the nature and relevance of bots for organizational research.

To guide the readers to our discovery, we start by identifying the assumption we problematize. We then explain our methodology and present our findings. Next we introduce two mini-cases to show how our discovery extends beyond the context of this study to other organizational and social settings. We subsequently discuss how neglecting bots can impose threats to research validity and we develop a method that scholars can use to best detect them. Finally, we consider the implications of bots for theory and practice.

Online Social Networks

An online social network "consists of a set of actors or nodes along with a set of ties of a specified type (such as friendship) that link them" (Borgatti & Halgin, 2011:1169) on a digital platform. Ties are connected via shared end-points to form paths indirectly relating actors that are not directly tied with one another. The pattern of ties in online social networks generates a particular structure, and actors occupy positions within this structure. It is important to recognize that the term *actor* does not necessarily imply that these social entities have "the volition or ability to "act"" (Wasserman & Faust, 1994:17). In other words, an actor is not necessarily a person but instead an entity. Also, it is the researcher who defines an online social network by choosing which type of ties and which set of actors to study (Borgatti & Halgin, 2011). Much of the online social network literature in Management and Information Systems defines one-mode networks illustrating structure (e.g., tie strength) and actor position (e.g., centrality) and relating these to either group-level (e.g., community growth) or actor-level (e.g., leadership) outcomes. In

[8] In our paper, we focus exclusively on social bots and use the term "bots" to refer solely to these. However, there are other types of bots on the Internet such as those used to scrape data from numerous websites (e.g., diffbot).

these studies, scholars implicitly assume that *actors* are *people*.[9] For instance, empirical work analyzing the impact of social influence on product adoption (Aral & Walker, 2014) and content generation (Zeng & Wei, 2013) on digital platforms such as Facebook and Flickr emphasize the human aspect of nodes by referring to them as individuals/people when presenting findings (*italics* emphasis added):

> "*Individuals* [on Facebook] exert 125% more influence on friends for each institutional affiliation they share in common ($p < 0.05$)" (Aral & Walker, 2014:1362).
> "We found that *people* tend to upload more similar photos [on Flickr] around the time of the formation of a social tie" (Zeng & Wei, 2013:72).

We observe this same type of assumption in research on *leadership*, where leaders of online communities are described as "*people* leading members of [a] newsgroup" (Faraj, Kudaravalli, & Waso, 2015:400); *e-commerce*, where nodes purchasing or reviewing products on Amazon are defined as "*the individuals*" (Dhar, Geva, Oestreicher-Singer, & Sundararajan, 2014:264) or seen as "*people*" (Kumar & Benbasat, 2006:428); *social movements*, where Twitter users protesting authoritarian regimes are described as "*people*" (Oh, Eom, & Rao, 2015:213); and *information diffusion*, where users receiving direct messages from other users aspiring to spread rumors on Twitter are seen as "specific *individuals* in the Tweeter's social network" (Oh, Agrawal, & Rao, 2013:412). We even see this assumption in a conceptual article about knowledge collaboration:

> "Knowledge collaboration requires that *individuals* spend time contributing to the OC's [online community's] virtual workspace" (Faraj, Jarvenpaa, & Majchrzak 2011:1227).

In this case, the belief is that knowledge collaboration—"the sharing, transfer, accumulation, transformation, and cocreation of knowledge" (Faraj et al., 2011:1224)—only occurs when *people* generate content to the online community's workspace. But, as we detail below in our

[9] We identified 80 online social network studies published in Management and Information Systems journals on the Financial Times Top 45 list. Only one study accounted for the existence of bots. They did so through the CAPTCHA method to reduce "message volume and uninteresting messages generated by *spam programs*" (Butler, Bateman, Gray, & Diamant, 2014:717). The examples we provide come from the remaining 79 studies.

discovery, this is not always the case. Actors collaborating knowledge, exerting or receiving influence, purchasing or reviewing products, protesting authoritarian governments, or diffusing information in online social networks do not need to be *people;* they can also be *bots*.

Post Bot-Discovery Exploration

We did not begin this project expecting to find bots to be central actors protesting government corruption on Twitter. This was a discovery that emerged as part of our inductive approach. Rather, we began with two research questions: *(1) Who are the central actors in Brazil's anti-corruption protest on Twitter? (2) What do they have in common?* As is often the case with inductive research, we engaged in post-discovery exploration (Charmaz, 2006); as we dug deeper in the case—by iteratively comparing our existing data to emerging data—the prevalence of bots as central actors became apparent. As bots replicated specific messages on Twitter, we discovered that they were central partly because they amplified the magnitude of content embedded in those duplicated posts. Therefore, our study was refined to not only provide answers to our initial research questions but to also shed light on the implications of our discovery through a post bot-discovery exploration phase.

Methods

We drew upon multiple data sources, including social network graphs of communications, personal e-mails, user profile reports, a four-part media article series, and an in-depth semi-structured interview. To preview our findings, and to serve as a blueprint for the discovery process described here, we present a timeline of our study in Figure 1. Details of each data source are explained next; an illustration of the data we collected is found in Table 1.

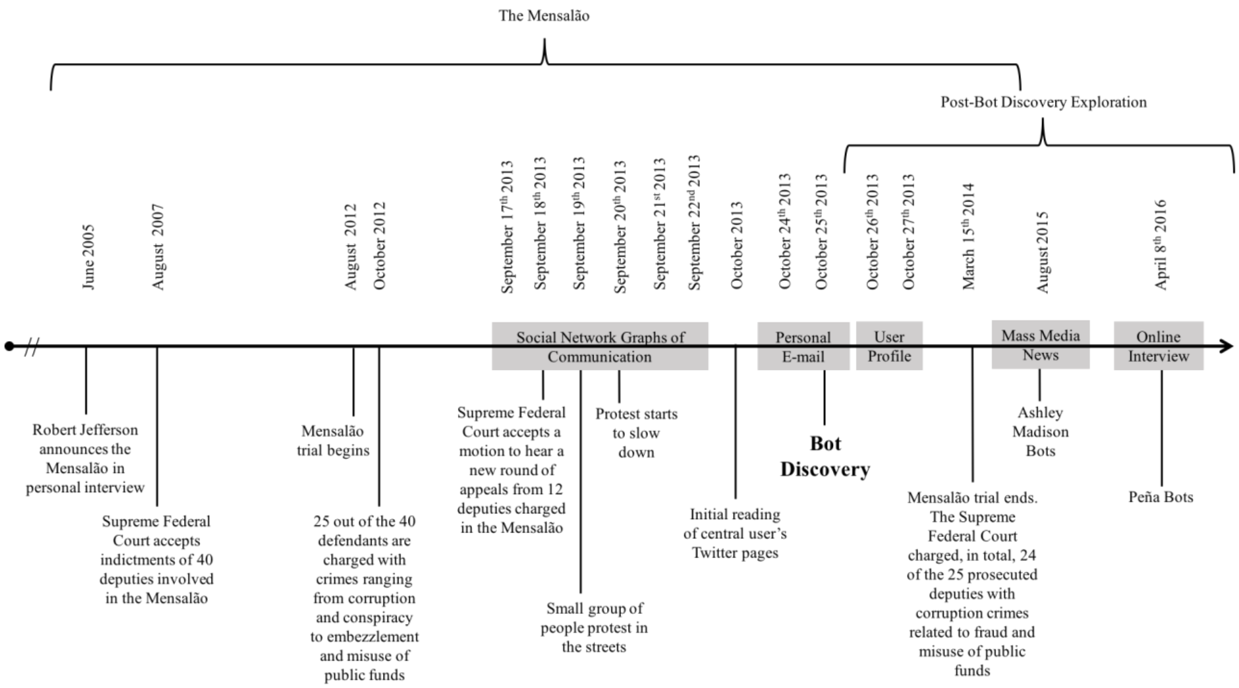

Figure 1. Study Timeline

We used NodeXL (Hansen, Shneiderman, & Smith, 2011) to collect social network data on Twitter. Our sample includes #ChangeBrazil messages since this was the most frequently used hashtag to protest government corruption in Brazil (Monroy-Hernández & Spiro, 2013). We gathered data from September 17th 2013, one day before the protest emerged, to September 22nd 2013, five days after its start. This time period was selected because Twitter networks tend to be busiest during the first five days of a protest (see interaction networks in Monroy-Hernández & Spiro, 2013). Our dataset includes 259 unique Twitter users with 4,513 messages. For every message posted (tweet, re-tweet, reply, or mention), we have the username of the initiator, the URL of the message, and the timestamp.

Table 1. Data Collection

Data Type	Study Phase	Quantity	Source
Social Network Analysis	Pre Bot- Discovery	4,513 messages from 259 unique Twitter users	NodeXL (Twitter)
Personal Communication	**Bot Discovery**	2 exchanges	E-mail

User Profile	Post Bot- Discovery	4 reports	Simply Measured
Mainstream Mass Media News Reporting	Post Bot- Discovery	4 files	Gizmodo
In-depth Semi-Structured Interview	Post Bot- Discovery	1 conversation (65 minutes in length)	Interview (Skype)

We also engaged in two email exchanges with a central actor in our sample. The e-mails confirmed the existence of bots and also included detailed information about their incentives, design, and usage. We began the post bot-discovery data collection portion of the project by reading and analyzing user profile data related to the central actors (bots and humans). These analyses were specific, and undertaken to further sensitize us as to how bots were different, yet similar, to people. To explore bot practices beyond the context of our study, we read and analyzed a series of articles delineating their use in a business organization. We also interviewed an activist studying bots in an election campaign in Mexico. We asked him questions dealing with the objective, motivation, and utilization of bots. The structured component of the interview allowed us to understand more details about the bots themselves while the unstructured component allowed for contextual details to emerge. The interview, conducted in Spanish, was audio recorded, professionally transcribed and translated verbatim.

Pre Bot-Discovery Data Analysis

To identify the central actors of the protest, we analyzed the social network data in two ways. First, because actor centrality may exhibit temporal patterns, we plotted six graphs using the Harel-Koren Fast Multiscale algorithm to visualize central actors for *each protest day* (see red ties in Figure 2). Second, for all users in our sample, we computed five centrality measures (defined in Table 2): in-degree, out-degree, betweenness, closeness, and eigenvector, which, together "*cover the intuitive range of the concept of centrality*" (Freeman 1979: 237). These two analyses enabled us to compare multiple types of data to check the robustness of our findings.

Table 2. Actor Network Centrality	
Dimension	**Definition**
In-degree	Number of incoming ties (tweets, re-tweets, replies, and mentions) of an actor
Out-degree	Number of outgoing ties of an actor
Betweenness	Number of times an actor bridges the shortest path between two other actors
Closeness	Average of the shortest path lengths from a certain actor to all other actors
Eigenvector	The degrees (in- and out-) of the actors that a certain actor is connected to

Findings

Figure 2 shows the Harel-Koren Fast Multiscale graphs (an algorithmic method for drawing large weighted social network diagrams rapidly). *Head mule* was the most central actor on September 17th, but the user becomes non-central once the protest begins (September 18th). Instead, two other actors (*anonymousfrai* and *guiql)* become and remain central through September 20th which is the day the protest starts to fade in the traditional media. The graphs also show that *anonymousfrai* and *guiql* are central because they bridge information across the network by connecting otherwise unconnected groups of users via tweets, re-tweets, mentions, or replies. Centrality computations corroborate this; *anonymousfrai* and *guiql* have high betweenness centrality scores between September 18th and 20th. While it may not be noticeable in Figure 2, we find that these actors are also connected to the same users—yet, they are not directly tied with each other. We observe a significant change on September 21st. The network is now divided into two groups and linked through *poa_cruel_news* (instead of *guiql*) and *anonymousfrai*. We note that, on the fifth day of the protest (September 22nd), *anonymousfrai* remains central while *poa_cruel_news'* bridging capability is replaced by that of *oan_max_ik.* In short, graph results indicate that the Twitter protest had four central actors during the six-day period. These users are *anonymousfrai, guiql, poa_cruel_news,* and *oan_max_ik*.

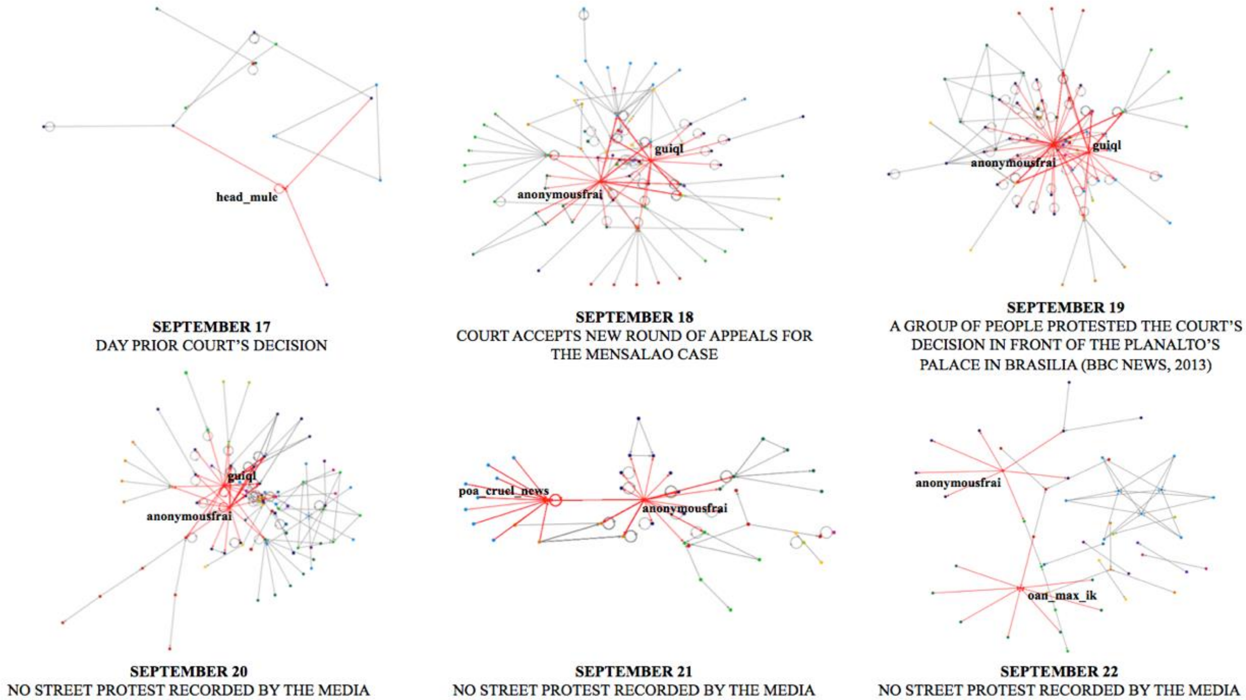

Figure 2. Central Actors in Brazil's Anti-Corruption Twitter Protest – #ChangeBrazil

Table 3 shows centrality scores for the five most important actors in our sample. Betweenness, eigenvector, and degree centralities were used as measures of importance. We did not include closeness centrality given that 70.6 percent of actors have a score of nearly zero, making this not a useful metric for differentiating prominence across sampled users. Numeric findings are consistent with those observed in Figure 2. The most central actors in Table 3 are also *anonymousfrai*, *guiql*, *oan_max_ik*, and *poa_cruel_news.* We find that *anonymousfrai* has a significantly higher betweenness centrality score (*BC* = 17,265.79) than all other central actors. This user is nearly three times more central than *guiql* (*BC* = 6,910.79) and about five times more central than both *oan_max_ik* (*BC* = 3,633.12) and *poa_cruel_news* (*BC* = 3,461.35), implying that *anonymousfrai* is the most important *bridge* in our study. *Anonymousfrai* is also more central than all other actors based on eigenvector centrality (*EC* = 0.07), while *guiql* (*EC* = 0.05) is more central than both *oan_max_ik* (*EC* = 0.02) and *poa_cruel_news* (*EC* = 0.02).

Intriguingly, we observe that both *anonymousfrai* and *guiql* score high in out-degree centrality—they mention various users in their tweets, reply-to several other users in their tweets, or they re-tweet other users' tweets frequently—yet, they practically score zero for in-degree centrality; only one user mentions them in their tweets, replies-to them in their tweets, or re-tweets their tweets. In contrast to *anonymousfrai* and *guiql*, both *oan_max_ik* and *poa_cruel_news* have high in-degree centrality and low in out-degree centrality.

Table 3. Centrality Scores for Networks' Most Important Actors

	Betweenness	Eigenvector	Closeness	In-degree	Out-degree
Anonymousfrai	17265.79	0.07	0.00	1.00	71.00
Guiql	6910.79	0.05	0.00	1.00	41.00
Oan_max_ik	3633.12	0.02	0.00	21.00	1.00
Poa_cruel_news	3461.35	0.02	0.00	20.00	1.00
E_ditora	2802.85	0.01	0.00	10.00	4.00

Our data analyses show robust evidence of four actors being structurally central in Brazil's anti-corruption protest on Twitter. Indeed, *anonymousfrai* is the most important actor followed by *guiql*, *oan_max_ik,* and *poa_cruel_news.* However, graph and numeric data are lean and therefore not adequate for discovering who these users are and what underlies commonalities in betweenness and eigenvector centralities. In addition, the thinness of these types of data make them unsuitable for scrutinizing similarities in both in-degree and out-degree centralities for *anonymousfrai* and *guiql*, and for *poa_cruel_news* and *oan_max_ik*, respectively. We therefore collected richer data (i.e., text and image) to learn more about each user and to discover what they have in common. The findings of these analyses are presented next.

Scrutinizing Central Actors: The Road to Discovering Bots

We examined each central user's Twitter page and gathered data that signaled information about their persona. We began with *anonymousfrai,* consistently the most important actor in our dataset. Personal textual information about this user was not found, besides what was presumed to be a blog address. However, *anonymousfrai*'s Twitter profile image—a Guy Fawkes mask—revealed that the user was part of *Anonymous*, a loosely associated international network of activists and hacktivists opposing Internet censorship and control. *Anons* (how members of Anonymous refer to themselves) supported the Occupy movement and the Arab Spring (Coleman, 2014) and thus it was not surprising that they were also opposing government corruption in Brazil. In addition, we discovered that *anonymousfrai* never tweeted a single message—instead, the user only re-tweeted other users' content. Re-tweeting can be considered a simple action, given that it is the forwarding of a message that has already been created and published by another user. Unmodified re-tweets—such as those spawned by *anonymousfrai*—are tweet copies with the exception that they include "RT @username" at the beginning of messages. This type of re-tweet does not require enactors to think or decide a next action. Furthermore, it had now become clear to us that *anonymousfrai's* out-degree centrality scores reflected the user's high re-tweeting activity rather than large sums of mentions or replies. Since *anonymousfrai* executed a simple action many times in a seemingly standardized manner, we developed a hunch that the user was an *Anon bot* coded to raise awareness for the protest by amplifying the volume of #ChangeBrazil messages via automated re-tweets.

Guiql was the next central actor that we examined. Unfortunately, only an e-mail address was located as a personal identifier upon initial review of the user's Twitter page. We made several attempts to contact the account holder through the provided e-mail address, but we never received a response. Considering our hunch related to *anonymousfrai* and the similarity in out-degree centrality between the two, we suspected that *guiql* was also a bot. This speculation motivated us to further review the user's Twitter activity. Similar to *anonymousfrai*, *guiql* re-tweeted other users' tweets frequently and was almost never mentioned or replied-to by others.

Unlike *anonymousfrai,* however, *guiql* also tweeted. Yet, the user's messages focused exclusively on media news—not only political, but a variety. Interestingly, *guiql* only tweeted news from a media news portal called Terra.[10] We discovered that this news portal had an official Twitter account for publicizing its own content and were surprised to notice that *guiql* regularly tweeted Terra's articles before Terra did. This finding challenged our bot hunch. Thus, our refined supposition was that *guiql* was a journalist working for Terra who was also an activist supporting the protest via manual re-tweets.[11]

The third central actor's, *oan_max_ik,* Twitter page provided a valid blog account. Upon initial review of the blog page we recognized that all posts revolved around politics. In addition, we discovered—by reading the "*about me*" section of the blog—that many other blog pages were associated with *oan_max_ik*'s account. No other data, however, were available from this user.

The last central actor's—*poa_cruel_news*—Twitter page also provided a valid blog account. Once we read the page, it became evident that *poa_cruel_news's* blog was one of the blogs linked to *oan_max_ik*'s account. Could these actors be the same person? Intrigued, the first author decided to question *poa_cruel_news* about it via email:

> "Hi Eros Thanatos[12], how are you? I have been reading your blogs and I was wondering if you're the only person writing on them. The content is all great but very diverse. I particularly enjoy the "metafísica do vento [the blog page related to *oan_max_ik's* Twitter account]," "poema para a porto alegre chauvinista-elitista-machista [the blog page related to *poa_cruel_news's* Twitter account]," and "opener media.""

Approximately two days later, Mr. Thanatos responded:

> "Yes. I am one person… but with diverse personalities, or better, I am a tribe of personalities… Very cool that you gave me feedback and for reading what I write. Thank you for reading and I am available for any clarification… Have a great week…"

[10] We obtained this information by selecting "view summary" within *guiql*'s tweets.
[11] We believed these were manual re-tweets because *guiql* did not re-tweet with the same high frequency as *anonymousfrai*. This user's re-tweets corresponded to nearly 58 percent of those published by *anonymousfrai*.
[12] Eros Thanatos was how the user described its persona on the blogs and it was also how he signed all emails.

Mr. Thanatos's response confirmed our suspicion that both accounts (*poa_cruel_news* and *oan_max_ik*) were managed by one person. His e-mail, however, did not provide any clues about our bot and journalist hunches associated with *anonymousfrai* and *guiql*, respectively. Since Mr. Thanatos was interacting with both of these users—*anonymousfrai* and *guiql* were re-tweeting his tweets—we thought that he might know *something* about them. Curious as to whether this was the case, the first author asked Mr. Thanatos, in a follow up e-mail, to comment:

> "Hi Eros: That's quite helpful. Thank you for your reply. I also tried to contact two other users because I read their work on Twitter but I never heard back. Do you happen to know them? Their Twitter names are *anonymousfrai* and *guiql*."

A few hours later, Mr. Thanatos replied:

> "In reality, I do not know them since they are ***Bot***, short for ***robot***, also known as ***Internet bot*** or ***Web bot***, it is a software application conceived to simulate human actions in a repeated and standardized manner, in the same way a robot would. Both accounts were created to support protests in Brazil such as #VemPraRua and #ChangeBrazil, being that every time a Twitter user tweeted a message with these hashtags the bots would replicate it with a re-tweet… I hope I was helpful… Regards and Carpe Diem."

As this e-mail reveals, central actors in online social networks are not always *people*, but instead, they can also be *bots.*[13] Based on this discovery, we recommend reframing them as *entities,* as suggested by Wasserman and Faust (1994). Indeed, these entities can then be categorized in diverse ways (as people or bots or organizations).

Post Bot-Discovery Exploration

The email exchange provided us with evidence that *oan_max_ik's* and *poa_cruel_news's* Twitter accounts were managed by one person (Mr. Thanatos) and that *anonymousfrai* and *guiql* were *bots* coded to support protests against corruption in Brazil. To learn more about the similarities and differences between bots and humans, we collected additional profile data for

[13] *Anonymousfrai* is a bot. We clicked on the user's blog address (http://anonymousfraiburgo.blogspot.se) which now says that the Twitter account represents a bot re-tweeting every message containing certain keywords or hashtags.

each central actor in our sample. We downloaded four Twitter reports (one per actor) from Simply Measured[14] that contain number of followers, following, and tweets, along with Klout Score and account creation date. There were no discernable patterns in terms of account history and activity. However, actors' *Klout Scores*—a measurement of overall online influence ranging from 1 to 100 (the higher the score, the more influential the actor)—were similar. The values ranged from 40.3 to 49.3 with a mean of 43.1 and a standard deviation of 4.2.

In summary, we found that although social network analysis suggested there were four distinct central actors protesting corruption on Twitter, the reality is that there were only three. Two of them are bots (*anonymousfrai* and *guiql*) while the third one is a blogger (Mr. Thanatos) managing two accounts (*poa_cruel_news* and *oan_max_ik*). Differences between bots and humans exist for *in-degree* and *out-degree* centrality. Bots score high on the latter while humans score high on the former. In addition, we found that while both bots were programmed to re-tweet every message containing the hashtags #ChangeBrazil and #VemPraRua only one of them was also coded to tweet (*guiql*). Finally, we discovered that central actors (bots and humans) have bridging importance and similar *Klout Scores*. Table 4 integrates our findings.

Table 4. Major Findings

Research Question	Data Type	Study Phase	Findings
Who are the central actors in Brazil's anti-corruption protest on Twitter?	Social Network Analysis	Pre-Bot Discovery	There are ***four*** central actors (*anonymousfrai*, *guiql*, *oan_max_ik*, and *poa_cruel_news*)
What do these central actors have in common?	Personal E-mail	**Bot Discovery**	There are, in fact, ***three*** central actors—***2 bots*** (*anonymousfrai*, *guiql*) and ***1 person*** (Mr. Thanatos) controlling ***2 accounts*** (*oan_max_ik*, and *poa_cruel_news*)
	User Profile & Social Network Analysis	Post-Bot Discovery	Bots have high ***out-degree*** centrality. People have high ***in-degree*** centrality. Central actors

[14] For more information, please see http://simplymeasured.com/about/#sm.00014ps91q1cmf7oxx814ulqzemsx

			(*bots and people*) have high ***betweenness*** centrality and an average of 43 ***Klout Scores***

We now turn our attention to the bots themselves by reviewing the literature to learn how they are used and to discover how prevalent they are in online social networks. Because our discovery is based on a single case study, it is possible that the insights of this paper do not apply broadly. Yet, based on our review, we suggest that Brazilian activists are not the only ones using bots, claiming that the generalizability of our discovery is extendable beyond both Brazil and protests. To further support this claim, we present two additional mini-cases detailing the use of bots to increase revenue in the business of online dating (Ashley Madison FemBots) and to manipulate public opinion during Mexico's 2012 presidential election campaign (Peña Bots).

Bot Prevalence & Usage

Certain bots are designed to behave benignly, often in ways that benefit society. For example, *SF QuakeBot* disseminates information about earthquakes, as they happen, in the San Francisco Bay area.[15] Other bots, however, are modeled to harm—they are coded to inflate support for a political candidate (Ratkiewicz, Conover, Meiss, Gonçalves, Patil, Flammini, & Menczer 2011); spread false rumors (Ferrara, Varol, Davis, Menczer, & Flammini, 2016); damage an organization's reputation (Messias, Schmidt, Oliveria, & Benevenuto 2013); and even limit free speech (Gallagher, 2015).

According to Lutz Finger, director of data science and engineering at LinkedIn, bots impose significant threats to organizations and society since they "are actually more common than you might think," and because they "can do things beyond our wildest dreams or nightmares."[16] As of the date of publication, Twitter and Facebook—the two most popular social

[15] https://twitter.com/earthquakessf
[16] http://www.lutzfinger.com/evil-business-social-media-bots/

networking sites[17]—contain as many as 23 million (about 8.5 percent) and 140 million (between 5.5 to 11.2 percent) bots, respectively (Goldman, 2014; Grant, 2014). Nearly 27 millions of Instagram users (close to 8.2 percent) are also bots (O'Reilly, 2015). While LinkedIn is unaware of its bot statistical pervasiveness (Okalow, 2015), the company filed a complaint in California's federal court noting that an unknown number of bots are being used to "steal data about legitimate users, breaching the user agreement and violating copyright law" (Lipkin, 2014, para. 1). Finally, Tumblr has also recognized that some of its users are bots (Perez, 2011).

Bots Beyond the Mensalão Protest

Mini Case 1—Ashley Madison's Female Bots. Ashley Madison is a Canadian organization connecting users interested in pursuing extramarital affairs. In July 2015, a group of hackers named *Impact Team* gained unauthorized access to Ashley Madison's website. A few days after the security breach, they released personal data about the organization's users since Ashley Madison refused to terminate its business. When *Impact Team* began to release Ashley Madison's data, they stated that "the site is a scam with thousands of fake female profiles" and that "90-95% of actual users are male" (Reddit, 2015, para. 2), a fact corroborated by Newitz (2015a-d) who published a detailed Gizmodo series describing her findings. She wrote:

> "What I *have* learned from examining the site's source code is that Ashley Madison's army of fembots appears to have been a sophisticated, deliberate, and lucrative fraud. The code tells the story of a company trying to weave the illusion that women on the site are plentiful and eager. … the company was clearly on a desperate quest to design legions of fake women to interact with the men on the site." (Newitz 2015b, para. 3)

Inherent in this quote is a belief that *female* bots were designed to intentionally interact with *male* users. This was due to a "dramatic gender disparity"—only 5.5 million profiles were described as female in a database of about 37 million users (Newitz, 2015a, para. 5).

[17] http://www.alexa.com/topsites/category/Computers/Internet/On_the_Web/Online_Communities/Social_Networking

The female bots did not appear out of nowhere—"they were probably cobbled together from abandoned and fraudulent profiles in the company's massive member database" (Newitz, 2015c, para. 26). In addition, Newitz noticed patterns in the data revealing that bots often had ashleymadison.com e-mail addresses, although other accounts were also registered with Hotmail. Many of them also had IP addresses that suggested people located at the Ashley Madison headquarters created the accounts. She further discovered—by searching through the source code—a set of comments written by the developers explaining the behavior of the bots. Based on her examination, they were programmed to send simple initial phrases such as "hi there" and "u busy?" (Newitz 2015b, para. 21). Once male users engaged in a dialog, bots responded with longer messages inducing them to pay for credits to carry on further conversations. The strategy, Newitz claims "worked marvelously—at least in 2012" (Newitz, 2015d, para. 3). She wrote:

> "When the engagers (i.e., female bots) were turned off in early 2011, the company's income took a nosedive. So did their conversion rate. When they were turned on again 14 months later, revenues and conversions skyrocketed. It appears that revenues went from roughly $60,000 per month, to $110, 500." (Newitz, 2015d, para. 3)

Mini Case 2—The Peña Bots. On July 1, 2012 Mexico elected its current president Enrique Peña Nieto, "a dashing, disciplined campaigner who promised to bring peace and prosperity back to a country weary of drug violence and slow growth" (Miroff & Booth, 2012: para. 1). Nearly four years after Peña Nieto's triumph, a Colombian hacker named Andrés Sepúlveda announced, in a Bloomberg interview[18], that "he [Andrés] led a team of hackers that stole campaign strategies, manipulated social media to create false waves of enthusiasm and derision, and installed spyware in opposition offices, all to help Peña Nieto, eke out a victory" (Robertson, Riley, & Willis, 2016: para. 8). Specifically, Sepúlveda claims he built "an army of 30,000 Twitter bots" to create trends favoring Peña Nieto as a way to throw the preferences of

[18] For more information, please see: http://www.bloomberg.com/features/2016-how-to-hack-an-election/

voters (Robertson et al., 2016: para. 33)—a fact endorsed by Alberto Escorcia[19], a social activist who analyzed Twitter data during Mexico's 2012 election and whom we interviewed. We asked Escorcia to comment on the Peña bots, e.g., how they were coded.[20] He responded:

> "I interviewed one of them, an engineer that [also] did that [what Sepúlveda did], and he told me that they made it [the Peña bots] with Python, and PHP. So what they did was create a system where they massively created many Twitter accounts. They changed the names on them, they changed their photos. And they used photos that they bought from databanks either on Facebook or Hive. … these programs used the accounts of real people, but at some point they used a "bot network" and they put out messages in favor of Peña Nieto and turned [them] into trending topics."

In both of these cases—Sepúlveda's conversation with Bloomberg and Escorcia's interview with us—the use of bots to manipulate public opinion during Mexico's 2012 presidential election was reported. Designing bots to create and popularize messages favoring Peña Nieto on Twitter was one way in which engineers allegedly helped the then candidate become president.

Discussion

We have examined the critical issue of bots on Twitter and how they were designed by activists to protest government corruption in Brazil. Our post-discovery data exploration phase provides deeper insight into bots. Embracing them offers new opportunities for theory development and refinement, which we discuss below. But before any theoretical progress can be made, we discuss how neglecting bots threatens the research validity of online social network studies. As a result, it is important for us to understand how to detect them.

Neglecting Bots: Threats to Research Validity

Scholars in Management and Information Systems define—through an implicit assumption—actors of online social networks as *humans*. Our case study problematizes this

[19] Escorcia has a blog (http://loquesigue.tv) in which he reports the findings of his analyses.
[20] We also asked him to describe the goals of the bots, how they were used, who used them, why they were developed, and if there were different types of bots.

assumption by discovering bots. We, therefore, pose that existing research imprecisely defines the concept of "actors" in online social networks. This lack of clarity can cause problems at the conceptual and operational levels (Podsakoff, Mackenzie, & Podsakoff, 2016).

At the conceptual level, sources of invalidity can originate from construct definition. A well-defined construct specifies what should be included and what should be excluded; if the domain is too broadly defined, extraneous factors other than the target construct may be included (Netemeyer, Bearden, & Sharma, 2003). Therefore, an imprecise *definition* of "actor" that embraces all online actors when it only intends to include humans (i.e., makes an error of inclusion), or excludes bots when it intends to include all actors (i.e., an error of exclusion) threatens construct validity. At the operational level, a lack of clarity threatens construct validity because it increases the likelihood that the operationalization of the concept will be contaminated and/or deficient (Mackenzie, 2003). For example, scholars operationalizing *actors* as people, and collecting online social network data without verifying whether they are bots, expose the construct to bot contamination. Consider leadership work as an illustration, where network centrality is a commonly used approach for identifying online leaders (e.g., Faraj et al., 2015; Huffaker, 2010). If the intent of these studies is to identify *human* leaders, then the existence of *bots* (which, as we have shown, can be central in online social networks) is a threat to the validity of a "*human leader*."

This potential contamination may explain some of the theoretical anomalies in research. Faraj et al. (2015), for example, hypothesize that online leadership is associated with actor sociability. Yet, they found no support for this relationship. Instead, central participants were more likely to be identified as leaders if they also exhibited sociable behavior. They explained:

> "… even though sociability does not predict identification as a leader, actors who are central in the communication network and exhibit greater sociability are more likely to be

> recognized as leaders. In other words, socially oriented behavior does not lead to someone being identified a leader but, all things being equal, sociability by highly central participants leads to increased recognition as a leader" (Faraj et al., 2015:406).

Though this explanation is entirely possible, what if both bots and humans comprise the central actors in their study? Is it possible that they found what they did because *bots* exhibit high centrality and low sociability whereas *humans* manifest high centrality and a combination of high and low sociability? Would their findings remain the same if bots were removed from the analysis or would sociability be predictive of leadership only if humans were considered? Is the bot/human distinction consequential to their theoretical arguments? Maybe, if their theorizing is about *human* behavior, and maybe not, if their theorizing generalizes to any type of actor.

More specifically, our bot discovery raises three implications for online social network research. First, when the validity of an actor hinges upon its *humanness*, and the concept is not clearly defined, *theoretical arguments* associated with the actor construct can be invalid. Second, scholars theorizing about *human* actors must not only define the concept but they must also identify bots and control for their potential effects. Third, if the humanness of actors is irrelevant to theoretical arguments of a study, researchers must be careful in how they describe these actors as to avoid assuming they are people. In short, scholars must consider what constitutes an "actor," define the construct accordingly, and devise a research design that eliminates confounding effects caused by contamination across actor types, if the distinction is necessary.

Because the notion of actors in online social networks is currently ill-defined, we provide a revised definition of the concept for those who want to take a broader view and not restrict actors to humans. In doing so, we strive to effectively and concisely capture essential conceptual properties and characteristics. We re-define *actors* of online social networks as "discrete entities populating socio-technical networks." These actors do not need to interact with one another in order to exist although it is likely that they will. Our definition is sufficiently narrow in that it sets an online boundary condition but is also broad enough to capture various types of actors (e.g., humans and bots) operating in distinct networks (e.g., Facebook and Twitter).

Detecting Bots: A Multi-Method Funnel Approach

The prevalence of bots in online social networks coupled with the potential threats that they impose to organizations and society has sparked scholarly interest in distinguishing bots from humans (e.g., Cresci, Di Pietro, Petrocchi, Spognardi, & Tesconi, 2015; Davis, Varol, Ferrara, Flammini, & Menczer, 2016). Existing research often deploys feature-based machine learning detection systems (e.g., the most popular method available for public use being *BotOrNot?*[21] (Davis et al. 2016)), but other common approaches involve the application of social networks (e.g., Paradise, Puzis, & Shabtai 2014) and the crowdsourcing of human intelligence (e.g., Cao, Yang, Yu, & Palow 2014). There is no consensus about which of these approaches is most effective, although it is evident that all have limitations. Existing social network techniques, for example, rely on the assumption that bots rarely have ties with humans and, therefore, tend to form their own communities (Wang, Mohanlal, Wilson, Wang, Metzger, Zheng, & Zhao, 2012). However, recent studies have shown this not to be the case; bots actually create links with people (Alvisi, Clement, Epasto, Lattanzi, & Panconesi, 2013) and, as a result, they do not form tight-knit groups (Yang, Wilson, Wang, Gao, Zhao, & Dai, 2014). While human crowdsourcing can exhibit a near-zero false positive rate, the method is not cost effective for networks containing millions of users, such as Twitter. Finally, machine-learning techniques are problematic because of training sample dependency. Algorithms taught to detect bots generating content in English, for instance, produce less-accurate results in other idioms. Instead of one particular method,

[21] http://truthy.indiana.edu/botornot/

researchers suggest the adoption of complementary techniques (Alvisi et al., 2013; Ferrara et al., 2016) that explore multiple dimensions of actors' behaviors such as activity, timing information, and content (e.g., linguistic cues such as the frequency of verbs and nouns). Examples include the Renren Sybil (Wang, Konolige, Wilson, Wang, Zheng, & Zhao 2013), the CopyCatch (Beutel, Xu, Guruswami, Palow, & Faloutsos 2013), and the SynchroTrap (Cao et al. 2014).

We contacted the first authors of the three complementary techniques with questions about their applications. The Renren Sybil uses a clickstream methodology (i.e., a sequence of click events generated by users) and is therefore inappropriate for other types of data (e.g., numeric). We also learned that both CopyCatch and SynchroTrap are not yet available for public use. The lack of an existing publicly available approach that incorporates complementary techniques and uses traditional[22], graphical, and textual data to detect bots has compelled us to action. In the next few paragraphs, we present the *funnel process* (see Figure 3), a multi-method approach that is both theoretical and technical; we combine current techniques (social networks, machine-learning, and crowdsourcing of humans) with the goal of providing a holistic and precise, but also a feasible, approach scholars can use to best detect bots.

We start with a simple assumption: not every bot poses a threat to research validity. Only those that actually have an impact on the online social network are likely to bias scholarly findings. Given our assumption, we advise researchers to first think about the *plausible sources (or mechanisms) of bot bias* by asking how might bots threaten the validity of their study's findings. In Table 5, we provide some guidance. For instance, political scientists may suspect that some of their sampled nodes are bots designed to create volume (i.e., noise) in order to manipulate political opinion online. To identify such bots, we recommend the computation of

[22] Numerical, categorical or binary (O'Neill & Schutt, 2013).

actor network centrality scores. Specifically, political bots designed to *create* noise are likely to have many *outgoing* ties, and so calculating out-degree centrality for all nodes in the sample to identify such important actors becomes appropriate here. To identify structural anomalies (i.e., prominent users who are potential bots), we recommend the median plus or minus two times the median absolute deviation (see Leys, Ley, Klein, Bernard, & Licata, 2013). Scholars can also think about a combination of mechanisms. For example, the *FemBots* contributing to Ashley Madison's revenue increases probably did not do so by only sending tons of messages to male users (out-degree) but also by receiving many responses back (in-degree). For those suspecting bots but who are unsure of their potential biasing mechanisms, we suggest the use of anomaly detection techniques in social networks (e.g., see Savage, Zhang, Yu, Chou, & Wang, 2014).

Table 5. Actor Network Centrality: Thinking About the Source of Bot Bias

Mechanism	Domain	Example	Centrality
Volume Creation	Politics	Salud_dia1*	Out-degree
Prestige Attainment	Gaming	Botgle	In-degree
Information Diffusion	Ecology Crisis	SFQuakeBot	Betweenness
By Association	-	-	Eigenvector
Volume + Prestige	Dating Services	FemBots	*In + Out*
Volume + Diffusion	Protests	Anonymousfrai	*Between + Out*
Prestige + Diffusion	Leadership	-	*In + Between*
Association + Diffusion	-	-	*Eigenvector + Between*

* Note: See http://emergencyjournalism.net/manipulation-of-public-opinion-in-venezuela-using-political-bots/

Once structural anomalies are identified, scholars can next *deploy two or more machine-learning algorithms* to further examine whether these are bots. Because every measure contains error, it is important to assess the reliability of user scores across different algorithms. *BotOrNot?* (Davis et al. 2016), BorOrNot (ABTO Software[23]), and *Boostor* (Morstatter et al., 2016) are options currently available for public use. They are feature-based, meaning that they use the content of messages, characteristics of user profiles, and the behavior of actors to classify

[23] See http://botornot.co

them as bots or humans. Different systems have different features and can generate different scores for the likelihood of a node being a bot. For nodes with non-converging scores (i.e., where different machine-learning algorithms do not agree whether the actor is a bot), we recommend *human crowdsourcing* as a next step. These human raters must be trained in order to identify and classify bots, humans, or even cyborgs (nodes exhibiting a mixture of human and bot features). Inter-rater reliability ought to be computed and disagreements resolved through discussion. With each stage of the process, we expect the actor sample size to decrease—identifying central actors based on theorized mechanisms should eliminate non-central ones; algorithms should identify whether most actors are bots or humans; and what remains can be handled by human raters.

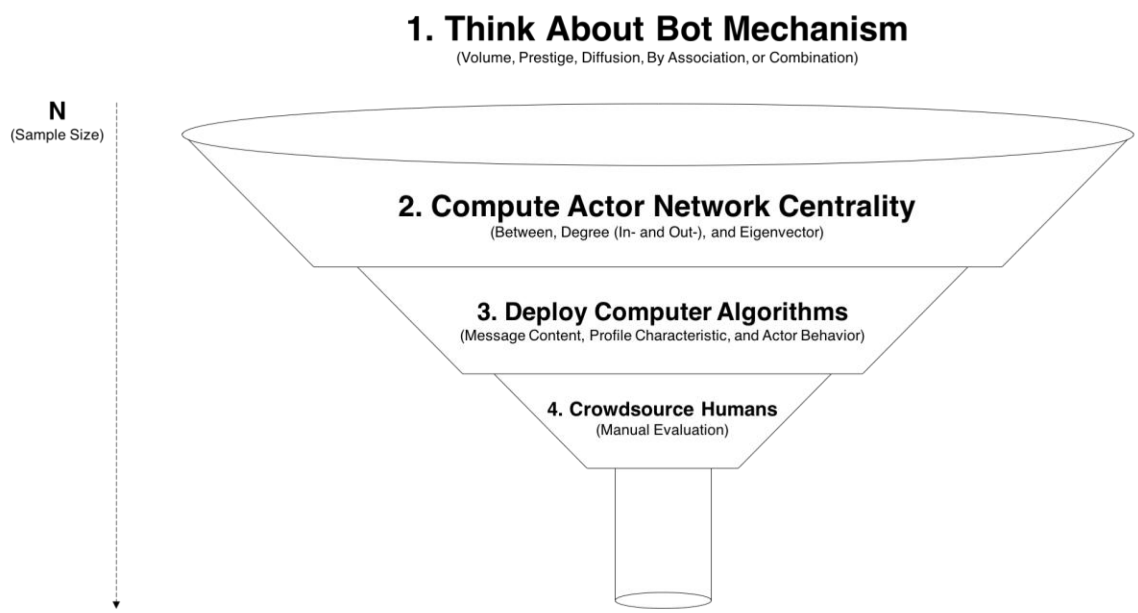

Figure 3. The *Funnel Process*: A Multi-Method Approach for Detecting Bots

Embracing Bots: Opportunities for Theory Refinement & Development

Because bots can be used across a wide range of organizational phenomena, they create a variety of opportunities for theory development and refinement. We explore some possibilities by focusing on what we discovered—bots as *central* actors *amplifying and spreading content* about a *protest online*—to discuss specific ideas and directions for researchers in social movements, where protests matter (King & Soule, 2007); in leadership, where network centrality matters (Faraj et al., 2015); and in information diffusion, where content sharing matters (Berger & Milkman, 2012). We also encourage future research to systematically derive dimensions of social and technical context that make use of bots most relevant and effective.

Designing Bots to Mobilize Collective Behavior in Social Movements. Although we discovered that bots were designed to support activists in protesting corruption on Twitter, we do not have evidence to suggest that they influenced the court to charge 24 of the 25 prosecuted deputies in the Mensalão case with corruption crimes (Oliveira, 2014). But we know that to achieve social change, activists attempt to raise public awareness for their protest to mobilize collective action (Castells, 2012). Bots raised awareness for the Mensalão protest, by amplifying and diffusing #ChangeBrazil[24] messages, but it seems that mobilization was not achieved—only a small group of people joined in street demonstrations (*BBC News*, 2013). Thus, a clear question for future research is how we design bots to raise awareness of a social cause in a way that also mobilizes people to participate in offline collective action. This requires thinking beyond the embeddedness of a social network to ask questions concerning the content of bot-generated messages. For example, should they be designed to produce posts inciting an emotional state of shock? Although these can get people's attention and mobilize participation through a sense of urgency (Warren, 2010), what makes some audiences indignant and sympathetic may simply annoy the broader community (Jasper, 2014).

Another venue for future research is to examine the nuances of bot design across different social movement entities. For example, do social movement organizations (SMOs), such as

[24] To examine whether *anonymousfrai* and *guiql* raised awareness for the Mensalão protest, we plotted the network *with* and *without* these actors. We learned that, by removing them, we lose ~24% unique ties (337 to 255), ~51% duplicate ties (2,763 to 1,418), and ~2% unique users (from 259 to 252).

Greenpeace, use bots differently than individual activists? SMOs may develop bots to mobilize people in signing digital petitions[25] over social media while individual activists may develop bots to raise awareness of a protest on Twitter. Because SMOs are formally competing with one another for member support and resources in a social movement industry (McCarthy & Zald, 1977), they are also more concerned than individual activists with how their reputation affects their success in supporting a movement (Selander & Jarvenpaa, 2016). We therefore expect SMOs to pay close attention to how certain *e-tactics* (e.g., bot usage for mobilizing digital petitioning signatures) align with their core values as to not jeopardize support and resources from members. The difference in reputational concerns between the two entities raises questions about the underlying theories we can use to explain differences in how SMOs versus individual activists use bots, and the theories we can leverage to understand what constitutes an effective bot in each context.

Refining Leadership Theory and Methods to Consider "Bot Leaders". Considerable progress has been made in understanding the emergence of online leaders (e.g., Johnson, Safadi, & Faraj, 2015). Yet, our discovery points to the need for scholars to exploit what makes an effective bot leader, how bot leaders influence human behavior, and whether there are different styles of bot leadership. *Bot leader* attributes should not be assumed to be equivalent to those of *human leaders*. Such bots may incorporate characteristics that are essentially different, contain others that are similar, and either complement or substitute human leadership. For instance, to the extent that bots can influence people's behaviors, they seem to meet the definition of leadership, but to the extent that leaders need to exhibit independent thought and judgment in deciding *who* and *how* to influence and in *what direction* to influence, bots may not, at least for

[25] Digital petitioning is one legitimate action that SMOs engage in with frequency (see Selander & Jarvenpaa, 2016)

now, seem to be leaders[26]—they may just be tools of the human leaders deploying them.

Existing research articulates that we need to apply multiple theories to understand the emergence of leaders in online communities since no single theory "seems uniquely suited" (Johnson et al., 2015:167). In the same spirit, we insinuate that a thorough study about the development of bot leaders also requires various theories. For example, a functional leadership view (e.g., Burke, Stagl, Klein, Goodwin, Salas, & Halpin, 2006) can help us understand which behaviors differentiate effective bot leaders from ineffective ones or even effective *bot* leaders from effective *human* leaders. In contrast, a shared leadership lens (e.g., Pearce & Sims 2000), may explain the nuances of substitutive and complementary functions that bots and people perform. A substitution lens suggests identifying functions performed by human leaders that can be perfectly substituted by bots so that the performance of a community, as a whole, is improved. A complementary angle, on the other hand, suggests identifying distinct roles for bot and human leaders such that synergistic effects occur.

The interplay between actor type and network centrality is also worthy of attention. This work can provide a methodological contribution by clarifying which centrality measure should be used to predict online leadership contingent on the study's definition of an "actor" (i.e., online leaders in general, bot leaders, or human leaders). At the moment, existing literature ignores bots and finds itself debating whether leaders have high betweenness (Faraj et al., 2015; Fleming & Waguespack, 2007; Johnson et al., 2015), out-degree (Huffaker, 2010), or both high betweenness and out-degree centrality scores (Sutanto, Tan, Battistini, & Phang, 2011). While our discovery shows that both bots and humans have high betweenness centrality, we found that only bots score high in out-degree centrality. Thus, it is possible that such measure is only valid for

[26] Advancements in cognitive computing (e.g., IBM's Watson) may change this.

identifying *bot leaders*. We also recommend scholars to consider whether in-degree centrality should be used as a complementary metric in the recognition of human leadership since our findings show that it differentiated people from bots. Finally, because bots and humans have high betweenness centrality, and because this is a widely used measure to predict online leadership, it is likely fair to suspect that betweenness centrality is an appropriate measure for identifying online leaders in general (bots and humans). In answering these questions, scholars must also exploit *theoretical arguments* for these relationships. For instance, do bot leaders have out-degree centrality because they excessively post on a community's thread discussion as a way to potentially spark online dialogs? Similarly, do human leaders have high in-degree centrality because they are frequently mentioned by others for their tenure status and high quality knowledge contribution? Finally, do bot and human leaders have high betweenness centrality because they are both able to spread these types of (or other) information across the community?

Bots Amplifying the Diffusion of Novel Information Through Content Sharing. Online content sharing is prevalent (Berger & Milkman, 2012) with 59 percent of people regularly re-tweeting messages, passing YouTube videos to relatives, and forwarding Amazon product reviews to colleagues (Allsop, Bassett, & Hoskins, 2007). Although we know that the *sharing* of online content is both frequent and relevant, less is known about *how* to actually engineer it as to amplify the diffusion of novel information. This is, we believe, an opportunity for bot research.

We know from prior work that weak-ties are more likely to provide novel information (Granovetter, 1973) and promote, in a proactive way, content sharing (Shi, Rui, & Whinston, 2014). This knowledge along with our bot discovery suggest that bots can be built to *amplify* the diffusion of novel information if they are designed to simultaneously be weak-ties and content sharers. They need to have both high betweenness (weak-tie) and out-degree (content sharing) centrality. An interesting question for future research is how to design *weak-tie content-sharing* bots. To do so, scholars must think beyond bot-level attributes to also consider how *platform infrastructure* (e.g., Twitter's 140-character limit) and *social network properties* enable or constrain bot design. Equally important is to recognize that while most of the content diffused by bots today is created by humans, this does not need to be the case.[27] It is also important to examine the effects of bot usage. For instance, one of the objectives associated with the amplification and diffusion of novel information is to exert influence. We suspect that bots can be used to influence people to change their behaviors through *awareness* and *social learning* (Aral, 2011). For example, bots automatically sharing content about Pokémon Go's augmented reality feature with acquaintances may influence people to download the application by making them *aware* of the feature. Similarly, bots sharing fun stories about Pokémon Go may impact the number of downloads by exposing non-adopters to the benefits associated with playing the game. An understanding of the mechanisms via which bot design enables the diffusion of novel information through content sharing is a necessary first step in our inquiry into bot influence.

Practical Bot Implications

***Protests and Corporations*.** Our discovery has an important implication for businesses targeted by protests and boycotts. Think about the potential use of bots for social appropriateness and boycott prevention. McDonnell and King (2013) noted that pro-social claims—expressions of the organization's commitment to socially acceptable norms, beliefs, and activities—function as an impression management strategy corporations use to neutralize reputational threats caused by boycotts. They found that a large increase in pro-social claims occurs when a boycott is more threatening (i.e., it receives more media attention), when a firm has a higher reputation, or when

[27] Advancements in cognitive computing (e.g., IBM's Watson) are increasingly enabling the creation of new knowledge by bots, and therefore, we may see an increase in the number of bots sharing novel knowledge.

a company has a history of engaging in pro-social claims. We claim that these same businesses do not have to wait for boycotts and protests to gain popularity to engage in pro-social claims. Instead, they can use bots to prevent these social movement tactics from occurring. For example, they can design bots to frequently broadcast the firm's social appropriateness online. Another strategy involves coding bots to automatically engage in pro-social claims when negative images and grievances about the corporation begin to circulate on social media. This requires linking them to social listening technologies that monitor conversations about an organization's image. These technologies then need to prompt bots—once a negative keyword about the organization emerges—to immediately generate content countering grievances made by activists with positive claims that emphasize the firms' commitment to social norms. In doing so, they can maintain audience support without recognizing or legitimizing activists' declarations.

From an activist standpoint, our work has implications related to demand concession. Corporate targets are more likely to concede to boycotts that generate large amounts of media attention since "they see sustained media attention to a boycott as an indicator of public support for the boycotters' cause and a signal that the boycott, if not ended, could lead to revenue loss" (King, 2008:400). Because many people use online social networks (e.g., Twitter) to consume news (Holcomb, Gottfried, & Mitchell, 2013), traditional media outlets (e.g., CNN) now actively report on nearly half of the most popular topics discussed in these social networks (Carrascosa, Cuevas, Gonzalez, Azcorra, & Garcia, 2015). Therefore, protestors can get traditional media to pay attention to their cause by designing bots to facilitate the popularity of their boycott online. Essentially, these bots need to amplify the channels in which activists transmit their grievances to obtain rapid and large public support for the boycott. To achieve this, they may want to design bots to constantly share and diffuse content that either explains the legitimacy of the boycott or exposes factual documents leading to immediate reputational harm of the targeted organization.

Conclusion

Our main discovery is that bots are central actors in online social networks. Our data show that these bots raise awareness of a protest on Twitter not because they frequently report on the status of street demonstrations or riots, but because they automatically share content (via re-tweets) on specific topics (by targeting certain hashtags) related to the protest. The discovery enabled us to challenge the assumption that actors of online social networks are people, also leading us to more clearly define the concept. We hope this clarification improves the validity of future work in the field and that online social network scholars find our methodology to detect bots useful. We also hope our discovery spurs further research on bots in organizational and social contexts. In particular, we suggest that scholars consider bot design in the mobilization of collective action, revise existing theories and methods for identifying online leaders, explore the notion of *weak-tie content-sharing* bots, and examine the many different mechanisms in which bots may influence changes in human behavior. Finally, we hope that our work helps inform researchers, practitioners, and policymakers about Twitter's bot usage in online activism worldwide (both for protesting societal issues but also for protesting against certain organizational actions), as they move forward with decisions in this important domain.

References

Allsop, D. T., Bassett, B. R. & Hoskins, James A. 2007. Word-of-mouth research: Principles and applications. ***Journal of Advertising Research****, 47:4 388–411, doi: 10.2501/S0021849907070419*.

Alvesson, M., & Sandberg, J. 2011. Generating research questions through problematization. ***Academy of Management Review****, 36:2 247–271, doi: 10.5465/amr.2009.0188.*

Alvisi, L., Clement, A., Epasto, A., Lattanzi, S., & Panconesi, A. 2013. Sok: The evolution of sybil defense via social networks. ***In 2013 IEEE Symposium on Security and Privacy. IEEE****, 382–396, doi:10.1109/SP.2013.33.*

Aral, S. 2011. Commentary—Identifying social influence: A comment on opinion leadership and social contagion in new product diffusion. ***Marketing Science,*** *30:2 217–223, doi: http://dx.doi.org/10.1287/mksc.1100.0596*.

Aral, S., & Walker, D. 2014. Tie strength, embeddedness, and social influence: A large-scale networked experiment. ***Management Science,*** *60:6 1352–1370, doi: http://dx.doi.org/10.1287/mnsc.2014.1936*.

BBC. 2013. Q&A: Brazil's 'big monthly' corruption trial. http://www.bbc.com/news/world-latin-america-19081519. Retrieved 7/30/2015.

Berger, J., & Milkman, K. L. 2012. What makes online content viral? ***Journal of Marketing Research,*** *49:2 192–205, doi:10.1509/jmr.10.0353.*

Beutel, A., Xu, W., Guruswami, V., Palow, C., & Faloutsos, C. 2013. Copy-Catch: Stopping group attacks by spotting lockstep behavior in social networks. ***In Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee****, 119–130, doi: 10.1145/2488388.2488400.*

Borgatti, S. P., & Halgin, D. S. 2011. On network theory. ***Organization Science****, 22:5 1168–1181, doi: http:// dx.doi.org/10.1287/orsc.1100.0641*.

Burke, C. S., Stagl, K. C., Klein, C., Goodwin, G. F., Salas, E., & Halpin, S. M. 2006. What type of leadership behaviors are functional in teams? A meta-analysis. ***Leadership Quarterly,*** *17:3 288–307, doi: 10.1016/j.leaqua.2006.02.007*.

Butler, B. S., Bateman, P. J., Gray, P. H., & Diamant, E. I. 2014. An attraction–selection–attrition theory of online community size and resilience. ***MIS Quarterly***, 22:3 699–728.

Carrascosa, J. M., Cuevas, R., Gonzalez, R., Azcorra, A., & Garcias, D. 2015. Quantifying the economic and cultural biases of social media through trending topics. ***PLoS ONE****, 10:7 1–14, doi: 10.1371/journal.pone.0134407*.

Castells, M. 2012. ***Networks of outrage and hope: Social movements in the Internet age.*** Cambridge, UK: Polity Press.

Cao, Q., Yang, X., Yu, J., & Palow, C. 2014. Uncovering large groups of active malicious accounts in online social networks. ***In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM****, 477–488, doi: http://dx.doi.org/10.1145/2660267.2660269.*

Charmaz, K. 2006. ***Constructing grounded theory: A practical guide through qualitative analysis.*** Thousand Oaks, CA: Sage.

Coleman. G. 2014. ***Hacker, hoaxer, whistleblower, spy: The many faces of Anonymous***. New York, NY: Verso.

Cresci, S., Di Pietro, R., Pretocchi, M., Spognardi, A., & Tesconi, M. 2015. Fame for sale: Efficient detection of fake Twitter followers. ***Decision Support Systems****, 80: 56–71, doi: 10.1016/j.dss.2015.09.003*.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. 2016. BotOrNot: A system to evaluate social bots. ***Working Paper.*** 2 pages.

Dhar, V., Geva, T., Oestreicher-Singer, G., & Sundararajan, A. 2014. Prediction in economic networks. ***Information Systems Research****, 25:2 264–284, doi: http://dx.doi.org/10.1287/isre.2013.0510.*

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. 2011. Knowledge collaboration in online communities. ***Organization Science,*** *22:5 1224–1239, doi: http://dx.doi.org/10.1287/orsc.1100.0614.*

Faraj, S., Kudaravalli, S., & Wasko, M. 2015. Leading collaboration in online communities. ***MIS Quarterly,*** 39: 393–412.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. 2016. The rise of social bots. ***Communications of the ACM.*** *59:7 96–104, doi: http://dx.doi.org/10.1145/2872518.2889302*.

Fleming, L., & Waguespack, D. M. 2007. Brokerage, boundary spanning, and leadership in open innovation communities. ***Organization Science,*** *18:2 165–180, doi: http://dx.doi.org/10.1287/orsc.1060.0242.*

Freeman, L. C. 1979. Centrality in social networks conceptual clarification. ***Social Networks.*** *1:3 215–239, doi: 10.1016/0378-8733(78)90021-7*.

Gallagher, E. 2015. Mexican Twitter bots: Cyber-attacks on free speech & organizing. http://revolution-news.com/mexican-twitter-bots-cyber-attacks-free-speech-organizing/. Retrieved 7/30/2015.

Goldman, D. 2014. 23 million Twitter users are fed by robots. http://money.cnn.com/2014/08/12/technology/social/twitter-bots/. Retrieved 5/17/2016.

Granovetter, M. 1973. The strength of weak ties. ***The American Journal of Sociology****, 78:6 1360–1380, doi: 10.1086/225469.*

Grant, R. 2014. Facebook has no idea how many fake accounts it has – but it could be nearly 140M. http://venturebeat.com/2014/02/03/facebook-has-no-idea-how-many-fake-accounts-it-has-but-it-could-nearly-140m/. Retrieved 5/17/2016.

Hansen, D. L., Shneiderman, B., & Smith, M. A. 2011. ***Analyzing social media networks with NodeXL: Insights from a connected world***. Burlington, MA: Morgan Kaufmann.

Holcomb, J., Gottfried, J., & Mitchell, A. 2013. News use across social media platforms. http://www.journalism.org/2013/11/14/news-use-across-social-media-platforms/. Retrieved 7/29/2016.

Huffaker, D. 2010. Dimensions of leadership and social influence in online communities. ***Human Communication Research.*** *36:4 593–617, doi: 10.1111/j.1468-2958.2010.01390.x.*

Jasper, J. M. 2014. ***Protest: A cultural introduction to social movements***. Cambridge, UK: Polity Press.

Johnson, S. L., Safadi, H., & Faraj, S. 2015. The emergence of online community leadership. ***Information Systems Research,*** *26:1 165–187, doi: http://dx.doi.org/10.1287/isre.2014.0562*.

King, B. G. 2008. A political mediation model of corporate response to social movement activism. ***Administrative Science Quarterly***, *53:3 395–421, doi: 10.2189/asqu.53.3.395*.

King, B. G., & Soule, S. A. 2007. Social movements as extra-institutional entrepreneurs: The effect of protests on stock price returns. ***Administrative Science Quarterly***, *52:3 413–442, doi: 10.2189/asqu.52.3.413.*

Kumar, N., & Benbasat, I. 2006. Research note: The influence of recommendations and consumer reviews on evaluations of websites. ***Information Systems Research***, *17:4 425–439, doi: 10.1287/isre.1060.0107*.

Leahy, J. 2012. Brazil gripped by anti-corruption battle. http://www.ft.com/cms/s/0/1b686ccc-21b1-11e2-b5d2-00144feabdc0.html#axzz3hOf8dRGJ. Retrieved 7/30/2015.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. ***Journal of Experimental Social Psychology***, *49:4 764–766, doi: http://dx.doi.org/10.1016/j.jesp.2013.03.013*.

Lipkin, M. 2014. LinkedIn sues bots over member data scraping. http://www.law360.com/articles/499109/linkedin-sues-bots-over-member-data-scraping. Retrieved 5/17/2016.

Lyons, J., & Cowley, M. 2013. Brazil Court Allows Corruption Case Appeals. http://www.wsj.com/articles/SB10001424127887323808204579083 571863389380. Retrieved 7/30/2015.

MacKenzie, S. B. 2003. The dangers of poor construct conceptualization. ***Journal of the***

Academy of Marketing Science*, 31:3, 323-326, doi: 10.1177/0092070303*254130.

McCarthy, J. D., and Zald, M. N. 1977. Resource Mobilization and Social Movements: A Partial Theory. ***American Journal of Sociology,*** *82:6, 1212–1241, doi: 10.1086/226464*.

McDonnell, M., & King, B. 2013. Keeping up appearances: Reputational threat and impression management after social movement boycotts. ***Administrative Science Quarterly.*** *58:3 387–419, doi: 10.1177/0001839213500032*.

Messias, J., Schmidt, L., Oliveira, R., & Benevenuto, F. 2013. You followed my bot! Transforming robots into influential users in Twitter. ***First Monday*** *18:7, doi: 10.5210/fm.v18i7.4217*.

Miroff, N., & Booth, W. 2012. Peña Nieto is winner of Mexican election. https://www.washingtonpost.com/world/the_americas/mexico-presidential-election-underway/2012/07/01/gJQAyd96FW_story.html. Retrieved 5/17/2016.

Monroy-Hernández, A., & Spiro, E. 2013. How Brazilian protesters are using Twitter. http://www.theguardian.com/news/datablog/2013/jul/04/brazilian-protesters-twitter-microsoft. Retrieved 7/30/2015.

Morstatter, F., Wu, L., Tahora, N. H., Carley, K., & Liu, H. 2016. A new approach to bot detection: Striking the balance between precision and recall. ***ASONAM.***

Netemeyer, R. G., Bearden, W. O., & Sharma, S. 2003. ***Scaling procedures: Issues and applications***. Thousand Oaks, CA: Sage.

Newitz, A. 2015a. Almost none of the women in the Ashley Madison database ever used the site. http://gizmodo.com/almost-none-of-the-women-in-the-ashley-madison-database-1725558944. Retrieved 11/08/2015.

Newitz, A. 2015b. Ashley Madison Code Shows More Women, and More Bots. http://gizmodo.com/ashley-madison-code-shows-more-women-and-more-bots-1727613924. Retrieved 11/08/2015.

Newitz, A. 2015c. How Ashley Madison hid its fembot con from users and investigators. The full title of the work it is part of http://gizmodo.com/how-ashley-madison-hid-its-fembot-con-from-users-and-in-1728410265. Retrieved 11/08/2015.

Newitz, A. 2015d. One chart that shows how much money Ashley Madison made using bots. http://gizmodo.com/one-chart-that-shows-how-much-money-ashley-madison-made-1727821132. Retrieved 11/08/2015.

Oliveira, M. 2014. Após um ano e meio e 69 sessões, STF conclui julgamento do mensalão. http://g1.globo.com/politica/mensalao/noticia/2014/03/apos-um-ano-e-meio-e-69-sessoes-stf-conclui-julgamento-do-mensalao.html. Retrieved 11/18, 2015.

Oh, O., Agrawal, M., & Rao, H. R. 2013. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. ***MIS Quarterly,*** 37: 407–426.

Oh, O., Eom, C. M., & Rao, H. R. 2015. Research note—Role of social media in social change: An analysis of collective sense making during the 2011 Egypt revolution. ***Information Systems Research,*** *26:1 210–223, doi: http://dx.doi.org/10.1287/isre.2015.0565*.

Okalow, S. 2015. Investigating social media's spam-bot problem. http://www.b2bnn.com/2015/08/investigating-social-medias-spam-bot-problem/. Retrieved 5/17/ 2016.

O'Neill, C., & Schutt, R. 2013. ***Doing Data Science***. Sebastopol, CA: O'Reilly Media.

O'Reilly, L. 2015. 8% of Instagram accounts are fakes and 30% are inactive, study says. The full title of the work it is part of http://www.businessinsider.com/italian-security-researchers-find-8-percent-of-instagram-accounts-are-fake-2015-7. Retrieved 5/17, 2016.

Paradise, A., Puzis, R., & Shabtai, A. 2014. Anti-reconnaissance tools: Detecting targeted socialbots. ***Internet Computing,*** *18:5 11–19, doi: 10.1109/MIC.2014.81*.

Pearce, C. L. & Sims, H. P. 2000. ***Shared leadership: Toward a multi-level theory of leadership***. Beyerlein MM, Johnson DA, Beyerlein ST, eds. Advances in Interdisciplinary Studies of Work Teams, Vol. 7 (Emerald Group Publishing Limited, Greenwich, CT), 115–139.

Perez, S. 2011. Tumblr acknowledges its growing spam problem, says it's doing everything it can. The full title of the work it is part of https://techcrunch.com/2011/10/31/tumblr-acknowledges-its-growing-spam-problem-says-its-doing-everything-it-can/ Retrieved 8/23, 2016.

Podsakoff, P. M., Mackenzie, S. B., & Podsakoff, N. P. 2016. Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. ***Organizational Research Methods***, *19:2 159–203, doi: 10.1177/1094428115624965*.

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., & Menczer. F. 2011. Detecting and tracking political abuse in social media. ***In 5th International AAAI Conference on Weblogs and Social Media***. 297–304.

Reddit. 2015. We are the Impact Team. We are releasing the Ashley Madison data. The full title of the work it is part of https://www.reddit.com/r/AnythingGoesNews/comments/3h71ar/ we_are_the_impact_team_we_are_releasing_the/. Retrieved 11/08, 2015.

Robertson, J., Riley, M., & Willis, A. 2016. How to hack an election. http://www.bloomberg.com/features/2016-how-to-hack-an-election/. Retrieved 7/12, 2016.

Savage, D., Zhang, X., Yu, X., Chou, P., & Wang, Q. 2014. Anomaly detection in online social networks. ***Social Networks,*** *39:1, 62–70, doi: 10.1016/j.socnet.2014.05.002*.

Schumann, S. 2014. ***How the Internet shapes collective actions***. Palgrave Pivot.

Selander, L., and Jarvenpaa, S. L. 2016. "Digital Action Repertoires and Transforming a Social Movement Organization," ***MIS Quarterly*** (40:2), 331–352.

Shi, Z., Rui, H., & Whinston, A. 2014. Content sharing in a social broadcasting environment: Evidence from Twitter. ***MIS Quarterly***, 38:123–142.

Singer, S. 2013. Between Executioners and Gangsters. The full title of the work it is part of http://www1.folha.uol.com.br/internacional/en/ombudsman/2013/09/1346607-between-executioners-and-gangsters.shtml. Retrieved 7/30, 2015.

Sutanto, J., Tan, C. H., Battistini, B., & Phang, C. W. 2011. Emergent leadership in virtual collaboration settings: A social network analysis approach. ***Long Range Planning,*** *44:5 421–439, doi: 10.1016/j.lrp.2011.09.001*.

The Economist. 2013. What is Brazil's "mensalão"? The full title of the work it is part of http://www.economist.com/comment/2216108. Retrieved 7/30, 2015.

Tufekci, Z. 2014. What happens to #Ferguson affects Ferguson: Net neutrality, algorithmic filtering and Ferguson. The full title of the work it is part of https://medium.com/message/ferguson-is-also-a -net-neutrality-issue-6d2f3db51eb0. Retrieved 11/18, 2015.

Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H., & Zhao, B. Z. 2013. You are how you click: Clickstream Analysis for sybil detection. ***In USENIX Security***, 241–256.

Wang, G., Mohanlal, M., Wilson, C., Wang, X., Metzger, M., Zheng, H., & Zhao, B. Y. 2012. Social Turing tests: Crowdsourcing Sybil detection. ***Social Information Networks***.

Warren, M. R. 2010. ***Fire in the heart: How white activists embrace racial justice***. New York: Oxford University Press.

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

Yang, Z., Wilson, C., Wang, X., Gao, T., Zhao, B.Y., & Dai, Y. 2014. Uncovering social network sybils in the wild. ***ACM Trans. Knowledge Discovery from Data***, *8:1 259–268, doi: 10.1145/2556609*.

Zeng, X., & Wei, L. 2013. Social ties and user content generation: Evidence from Flickr. ***Information Systems Research***, *24:1 71–87, doi: http://dx.doi.org/10.1287/isre.1120.0464*.

Chapter 3. The Role Of Bots In Social Movements: A Theory Of Brokecasting[28]

[28] Salge, Carolina; Karahanna, Elena; & Thatcher, Jason. To be submitted to MIS Quarterly.

Abstract

While bots—automated accounts in online social networks—are important and potentially impactful, the roles that they play in organizations and society and how they enact these roles are relatively unexplored. In this research, we inductively study the role of bots in Brazil's 2013 social movement on Twitter. Relying on social network analysis and reverse engineering we discover that they act as *BrokeCasters*—i.e., as brokers and broadcasters of the cause. We also identify two new mechanisms for such *BrokeCasting*. The first mechanism, *Actor Tagging Volume,* relies on effective network volume and is part of broadcast. *Actor Tagging Volume* requires bots to leverage actor tagging features of social media, such as mentions, to share few messages and reach many non-redundant actors. The second mechanism is part of brokerage and relies on *Content Tagging Diversity*, a means by which bots advantageously use content tagging features of social media, such as hashtags, to jointly access information associated and *not* associated with social movements. *BrokeCasting* is important because it can amplify the size and broaden the reach of a social cause, raising its awareness for the public at large, not just the activists, bystanders, constituents, or opponents. Altogether, we make contributions at the nexus of IS, sociology, and computer science, in that we explain the role of bots in social movements and thus provide new insights to enhance bot detection.

Keywords: bots, social movements, network content, social media, tagging features

On June 7th of 2015, *StayWokeBot* was born on Twitter. The actor is a bot created to protect and maintain morale among the black protest community. As of the date of writing, *StayWokeBot* performs two actions, though the bot's design may be changed to accomplish more in the future. When Twitter users initially subscribe to the bot, it autonomously messages them words of affirmation referencing black celebrities such as Oprah. "A pioneer who breaks the model, you build things we can't foresee, a mogul and a giver, you are like Oprah Winfrey[29]," reads an example. But when users write their state initials to *StayWokeBot* (e.g., GA for the state of Georgia), it replies with contact information for that state's senators and a message encouraging them to call and ask representatives to "vote for universal background checks and the Hate Crime Prevention Act to stop people who've committed hate crimes from buying guns.[30]" The creators of *StayWokeBot* are still learning the possibilities but, at the moment, they want the bot "to help you become a more empowered activist" (Harris, 2016: para. 7).

The story of *StayWokeBot* is not unique; people are using bots to protest government corruption (Salge & Karahanna, 2018); to promote innocuous political events (Forelle, Howard, Monroy-Hernández, & Savage, 2015); to call bystanders to action (Savage, Monroy-Hernández, & Hollerer, 2016); to answer "day-to-day life of an activist" type of questions (see 101-a-tron[31]); to initiate arguments with opponents of a movement (see AI_AGW[32]); and more recently, to spread fake news in political elections (Shao, Ciampaglia, Varol, Flammini, & Menczer, 2017). While bots are prominent and potentially impactful they are still understudied in IS scholarship. In this research, we foreground bots by studying their role—i.e., their behavior and its association with the position they have in the network—in Brazil's 2013 social movement on

[29] https://twitter.com/StayWokeBot/status/778899145935519745
[30] https://twitter.com/StayWokeBot/status/774989824877072384
[31] https://twitter.com/101atron
[32] https://www.technologyreview.com/s/421519/chatbot-wears-down-proponents-of-anti-science-nonsense/

Twitter. Studying the role of bots vis-à-vis their network position is analytically useful since network ties are critical to social movement performance. Actors "located in strategic locations within the network" (Wasserman & Faust 1994:169) may improve social movement performance by providing influence (Diani, 2003), micro-macro linkages in the mobilization process (Han, 2009), and greater broadcasting volume (Lotan, Graeff, Ananny, Gaffney, Pearce, & Boyd, 2011; Tan, Ponnam, Gilham, Edwards, & Johnson, 2013).

Although actor position is important and bots are likely central (see Salge & Karahanna, 2018), it is still unclear how bots can be advantageously used in social movements. Seminal work in network theory (e.g., Granovetter, 1973; Burt, 1992) provides some strategic guidance but it does not address the content of ties. It leaves especially unclear how tagged content, such as hashtags on social media, can be leveraged for bots to act in support or detriment of social change. In contrast, an emerging stream of research attends to the synergies of technology and social activism by studying the impact of digital repertoires on the transformation of social movement organizations[33] (SMOs) (Selander & Jarvenpaa, 2016), the implications of digital ways of organizing (Ghobadi & Clegg, 2015), the role of social media in social change, connective action, and the emancipation or hegemony of public discourse (Oh, Eom, & Rao, 2015; Vaast, Safadi, Lapointe, & Negoita, 2017; Miranda, Young, & Yetgin, 2016) along with other issues in collective action (see Constantinides & Barrett, 2015; Ling, Pan, Ractham, & Kaewkitipong, 2015; Wattal, Schuff, Mandviwalla, & Williams, 2010). But while these studies show how technologies change the ways movements organize and mobilize action, this line of work has yet to address bots and clarify their potential value for social activism.

One feasible value of bots lies in raising movement awareness, a necessary condition for

[33] An SMO is a "complex or formal organization which identifies its goals with the preferences of a social movement or countermovement and attempts to implement those goals" (McCarthy & Zald, 1977:1218).

social change. Public awareness today resides on social media networks—traditional media actively report on nearly half of social media's most popular topics (Carrascosa, Cuevas, Gonzalez, Azcorra, & Garcia, 2015). Because bots can automatically share a lot of content in these platforms and message volume is a significant predictor of topic visibility (Naaman, Becker, & Gravano, 2011), they have the means to influence what gets reported by traditional media. In addition, the amount of media attention given to a social movement is positively correlated with its concession of demands (King, 2008). Therefore, with bots, SMOs and activists may not only get the traditional media to pay attention to their cause but they may also be able to influence demand concession (see Figure 4).

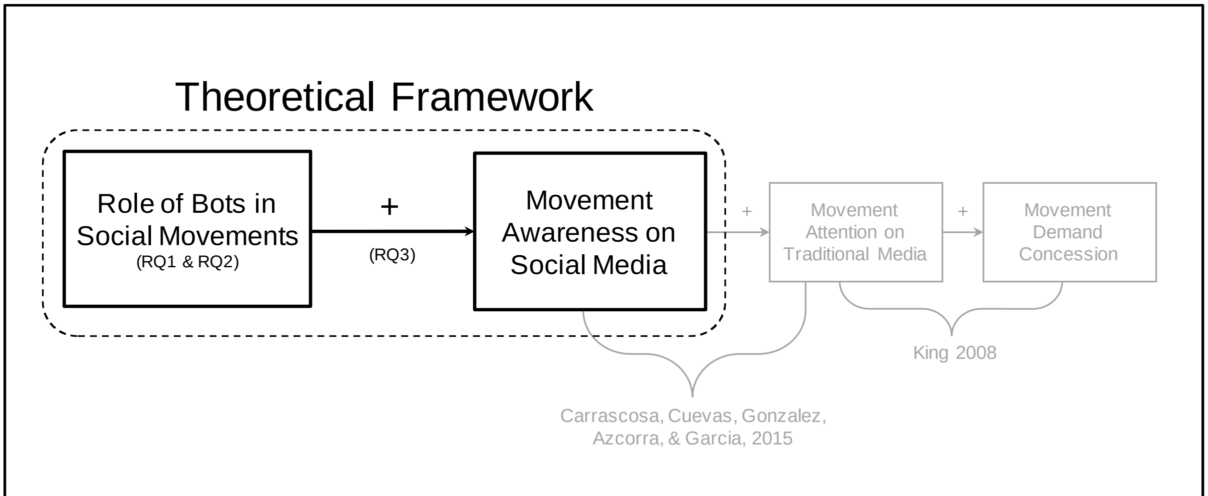

Figure 4. Potential Value of Bots for Social Movements

Note: The work of Carrascosa, Cuevas, Gonzalez, Azcorra, & Garcia (2015) is not on social movements but instead, general topics.[34] King (2008) studies boycotts (a social movement strategy).

Taken together, previous research studies illustrate the use of bots for collective action, show that social movements benefit from activists with numerous network ties, and suggest that strategic action on digital networks is likely germane to social causes, especially in regard to public awareness. But the literature lacks an in-depth investigation of bots and their roles in

[34] Because Carrascosa et al. (2015) do not study downstream effects beyond movement attention on traditional media, it is not clear whether the effect of social media awareness on demand concession has both mediated and direct effects.

social movement networks. Our study addresses this. We ask three questions. First, what roles do bots play in a social movement on Twitter? Second, how do they enact the roles that they play? Third, how can we leverage their roles for movement awareness? Given limited theory and scientific evidence, we conduct multiple-case theory building (Eisenhardt, 1989). Relying on a large and detailed dataset, we use social network analysis to first identify *what* roles bots play over the most active period of the movement. We then reverse engineer their design to examine *how* they enact such roles. To leverage the role of bots for movement awareness on social media, we integrate our empirical findings to a theoretical model of *BrokeCasting*.

Our study provides several contributions to research. First, we introduce the new role of *BrokeCaster* and highlight its importance for raising awareness of a social movement on social media. Although brokers and broadcasters are germane on their own, our central insight is that there are synergistic effects which occur by combining these roles together. Simply put, without volume and reach, it is difficult for social movements to be widely known. Second, we contribute a theoretical framework which delineates two new mechanisms for how bots *BrokeCast*. The first mechanism, *Actor Tagging Volume*, relies on effective network volume and is part of broadcast. *Actor Tagging Volume* requires bots to leverage actor tagging features of social media, such as mentions, to share few messages and reach many non-redundant actors. The second mechanism is part of brokerage and relies on *Content Tagging Diversity*, a means by which bots advantageously use content tagging features of social media, such as hashtags, to jointly access information associated and *not* associated with social movements. *BrokeCasting* is important because it can amplify the size and broaden the reach of a social cause, raising its awareness for the public at large, not just the activists, bystanders, constituents, or opponents. Altogether, we make theoretical contributions at the nexus of IS, sociology, and computer science in that we explain the role of bots in social movements and thus provide new insights to enhance bot detection.

Related Work

We reviewed the literature on bots and social movements. Based on the aim of our study, we primarily focused on social movements studies about actor roles in social networks.

Bots

Bots are automated accounts in online social networks (Morstatter, Wu, Nazer, Carley, & Liu, 2016). They are often designed in two different ways. First, they can be based on a set of rules which implies they are restricted as to what they can do. These bots only respond to specific commands prevalent in their scripts and they are only as clever as they are programmed to be. *StayWokeBot* represents an example. Second, bots can be based on machine learning. These algorithms understand language, not just commands, and they continuously learn from conversational data fed to them. An example here is Microsoft's Tay, a bot built to "experiment with and conduct research on conversational understanding[35]."

Regardless of whether they are rule- or learning-based, bots have many other components (Freitas, Benevenuto, Ghosh, & Veloso, 2014; Savage, Monroy-Hernández, & Hollerer, 2016). For instance, the profile of a bot refers to its description on a specific social platform. On Twitter, this involves a username, biography, join date, location, photo, and number of followers, following, and updates. In addition, bots are often designed to achieve an objective (Ferrara, Varol, Davis, Menczer, & Flammini, 2016). They can be coded to empower activists of a social movement (e.g., *StayWokeBot*); to disseminate information about natural disasters (e.g., *SF*

[35] To learn more about *Tay* please see https://blogs.microsoft.com/blog/2016/03/25/learning-tays introduction/#sm.00001j4n2twf79f1fvam9dj6eg6tv

QuakeBot[36]); to elicit laughter through satire (e.g., *Big Data Batman*[37]); to maliciously spread a rumor about a person (e.g., *AutoModerator*[38]); to damage an organization's reputation (Messias, Schmidt, Oliveria, & Benevenuto 2013); or to manipulate public opinion (Forelle, Howard, Monroy-Hernandez, & Savage, 2015). Bots also have a script, defined as a list of executable commands crafted by their developers and intended to help them achieve their objectives. To empower Black Lives Matter activists, for example, *StayWokeBot* first provides users with contact information about their state's senators and then encourages them to call and ask politicians to vote in favor of two gun-control measures. Finally, bots operate on many social media platforms. In fact, as of the date of publication, Twitter and Facebook—the two most popular ones for social networking[39]—contain as many as 23 million (approximately 8.5 percent) and 140 million (between 5.5 to 11.2 percent) bots, respectively (Goldman, 2014; Grant, 2014).

Because bots are prevalent in a variety of social media platforms (see Lipkin, 2014; O'Reilly, 2015; Okalow, 2015; Perez, 2011 for additional examples) and because they can also impose threats to organizations and society (Ferrara et al., 2016), scholars have spent a great amount of effort in detecting them (e.g., Cresci, Di Pietro, Petrocchi, Spognardi, & Tesconi, 2015; Davis, Varol, Ferrara, Flammini, & Menczer, 2016). In doing so, they often deploy social network techniques (e.g., Paradise, Puzis, & Shabtai 2014), invest in human crowdsourcing (e.g., Cao, Yang, Yu, & Palow 2014), or develop machine learning algorithms (e.g., *BotOrNot?*[40] (Davis et al. 2016)). Despite the advantages of each method, it is evident that every one of them has limitations too. Social networks, for example, assume that bots cluster in their own

[36] https://twitter.com/earthquakessf
[37] https://twitter.com/bigdatabatman
[38] https://www.reddit.com/user/AutoModerator
[39] http://www.alexa.com/topsites/category/Computers/Internet/On_the_Web/Online_Communities/Social_Networking
[40] http://truthy.indiana.edu/botornot/

communities (Wang, Mohanlal, Wilson, Wang, Metzger, Zheng, & Zhao, 2012) even though we know this is not the case since they actually create links with people (Alvisi, Clement, Epasto, Lattanzi, & Panconesi, 2013). The human crowdsourcing approach comes closer in accurately detecting bots, exhibiting a near-zero false positive rate, yet the method lacks efficiency; it is too costly for networks with millions of users, such as Twitter. Last, machine-learning techniques suffer from training sample dependency. Algorithms taught to identify bots in English, for example, produce less accurate results in other languages. If these samples are also not regularly updated as to reflect appropriate representations of a population of interest, learning becomes impaired and detection, as a result, inaccurate.

Instead of designing a new algorithm that applies one of the existing methods, recent research on bots takes a complementary approach to detection. Salge and Karahanna (2018), for instance, combine current techniques (social networks, machine-learning, and crowdsourcing of humans) in a procedure called funnel process to provide "a holistic and precise, but also a feasible, approach scholars can use to best detect bots" (Salge & Karahanna, 2018, p. 24). To meet such goal, they assume not every bot to be impactful for the specific research purpose, introducing four different phases in the process. The first is conceptual and it requires scholars to think about mechanisms by considering the ways in which bots may impact research findings. Next is to identify which actor network centrality measures are indicative of the identified mechanisms, compute these, and use them as proxies. Here, it is important to select structural anomalies—i.e., actors who are central in the network with respect to their proxies of interest. Because every algorithm contains error, the third step is to deploy two or more computer algorithms to identify which of the central nodes are bots. Finally, for nodes with unreliable or non-converging algorithmic scores, the recommendation is to crowdsource their evaluation to humans. Other complementary approaches to bot detection exist too—they are either unavailable for public use [e.g., CopyCatch (Beutel, Xu, Guruswami, Palow, & Faloutsos 2013) and SynchroTrap (Cao et al. 2014)] or only applicable to data following a sequence of click events [e.g., Renren Sybil (Wang, Konolige, Wilson, Wang, Zheng, & Zhao 2013)].

Overall, the literature on bots shows that (1) they are either based on a set of rules or machine learning; (2) they have many different design characteristics including a profile, an objective, and a script; and (3) they operate on a variety of social media platforms. Existing work also indicates the scholarly community has put forth a lot of effort in detecting bots. Yet while germane, detection techniques are conceptually focused on what differentiates bots from humans, and distant from actual bot design theory. That is, these techniques recognize structural patterns of bot behavior, but lack theoretical insight into *how* bots are designed to achieve a specific objective. We address this in the context of social movements.

The Role of Actors in Social Movement Networks

Our study focuses on the role of bots in social movements. We define social movements as a set of opinions and beliefs in populations that represent preferences for changing some elements of social structures (McCarthy & Zald, 1977). Although scholars have extensively studied the topic of social movements[41], there are only a few studies examining the role of actors in social movement networks (Vaast, Safadi, Lapointe, & Negoita, 2017; Diani, 2003; Han, 2009; Gökçe, Hatipoglu, Göktürk, Luetgert, & Saygin, 2014; González-Bailón, Borge-Holthoefer, & Moreno, 2013; Lotan, Graeff, Ananny, Gaffney, Pearce, & Boyd, 2011; Malinick, Tindall, & Diani, 2013; Morales, Losada, & Benito, 2012; Tan, Ponman, Gillham, Edwards, & Johnson, 2013; Tremayne, 2014). Most of the work in this literature is conducted online and,

[41] See Edwards (2014) for an extensive review.

specifically, on Twitter (see Vaast, Safadi, Lapointe, & Negoita, 2017; Gökçe et al., 2014; González-Bailón et al., 2013; Lotan et al., 2011; Morales et al. 2012; and Tremayne, 2014). In addition, a large majority of these studies focuses on the adherents of a movement, employing actor network centrality to operationalize their roles. For example, Diani (2003) uses in-degree centrality (i.e., number of incoming ties of an actor) to measure the leadership role of SMOs during Milan's 1980s environmental movement. SMOs with high scores are leaders because they are identified, by many other SMOs, as alliance partners having influence in the movement sector. They are also "more likely to be connected to media and political institutions, and thus in the best position to act as 'representatives' of the movement" (Diani, 2003, p. 118).

While in-degree centrality is associated with the role of leadership (see also Malinick et al., 2013; Morales et al., 2012; and Gökçe et al., 2014), the centrality ratio of in-to-out-degree is used to quantify information diffusion or the so-called role of a broker (Morales et al., 2012). Another common approach to operationalize brokerage is to count the number of times an actor bridges the shortest path between two other actors—i.e., to compute betweenness centrality (Han, 2009[42]). Nodes with high betweenness centrality scores do not diffuse information through volume (as is the case with degree) but instead by reaching beyond group boundaries and connecting otherwise unconnected groups of actors. They create bridging ties that can provide "micro-macro linkages in the mobilization process" (Han, 2009, p. 146) and their messages may serve "no other purpose than to make various people and groups aware of each other" (Tremayne, 2014, p.123). The network structural position privileging these actors is the absence of ties or what Burt (1992) calls the presence of structural holes.

An additional role documented in the literature is that of broadcast. Broadcasters

[42] Closeness centrality is also used to quantify brokerage but Han acknowledges that betweenness centrality "comes closest to measuring the extent to which an actor operates as a broker" (Han, 2009, p. 152).

prolifically share social movement messages with others and out-degree centrality (i.e., number of outgoing ties of an actor) is often used to quantify such a role (Lotan et al., 2011; Tan et al., 2013), often with actors' network reach, or their follower count on Twitter (Tremayne, 2014). Finally, a few scholars use quantitative methods other than network centrality to assess the role of social movement actors (see Vaast, Safadi, Lapointe, & Negoita, 2017). To quantify broadcasting, for example, González-Bailón et al. (2013) focus on activity (ratio of messages received over sent) and audience (ratio of friends over subscribers). They also discuss the role of a common user [or 'passive consumer' in Morales et al. (2012)]. Common users contribute to the gross activity of the movement but do not stand out in any way.

To integrate, our review of the literature suggests that actors support the goals of a social movement by (1) building many alliance partnerships and being closely connected to the media (leadership role operationalized as high in-degree); (2) by diffusing movement messages across a social network (brokerage role operationalized as high betweenness or high in-to-out degree ratio); and (3) by sharing a lot of messages about a movement with others (broadcast role operationalized as high out-degree). But the social movements literature does not address bots—it focuses on either humans or SMOs. Thus, despite the importance of digital activism (Ghonim, 2012), a granular account of how bots act in support or detriment of social change is still lacking.

Methodology

Given limited theory and scientific evidence for the role of bots in social movements, we apply inductive methodology to generate theory from multiple cases (Eisenhardt, 1989).

Our focus is on Brazil's 2013 social movement (also known as the *20 cents movement*), which was the country's largest and most significant organized collective effort at social change in a generation. Public expressions of discontent began on June 6 and were initially organized to protest a R$0.20—equivalent to $0.09 at the time—increase in bus fares across a few Brazilian cities, but grew to include other political issues such as high-levels of government corruption, lack of investment in public infrastructure, the proposal of a gay cure bill, and the realization of FIFA's 2014 World Cup. The movement initially received press criticism for being violent, blocking traffic, and making unrealistic demands. But, in the middle of June, mainstream media and TV networks adjusted their views and began to support its causes. This was when street protests exploded in breadth and propagated across the country. Rio de Janeiro, for example, hosted a demonstration that lasted seven hours and attracted nearly 300,000 activists, starting on the afternoon of June 17 and ending on late dawn of June 18 (Watts, 2013). To demonstrate their solidarity with the movement, Brazilians abroad also joined marches across 10 different countries on those two days. On June 19, the mayor and governor of São Paulo announced a reduction in fares. A day later, more than one million people protested throughout the country. Overwhelmed by the gravity of the situation, president Rousseff cancelled her trip to Japan and, on June 21, she addressed the movement on national television. Three days later, the government announced a proposal for congress to approve a referendum on widespread political reform.

The setting for our research is Twitter. Twitter was an appropriate choice for a variety of reasons. First, studying a single social media platform such as Twitter controls for several extraneous sources of variation and is a significant advantage of our design—it enables a more valid comparison among bots. Second, what happens to a movement on Twitter is important because it affects what happens to a movement off Twitter (Ghonim, 2012; Tufekci, 2014). As Tufekci (2017) mentions, "without a tool similar to Twitter with its hashtags, and without all this digital connectivity, it would be quite difficult to call up or sustain spontaneous protests of this size," in a reference to the Gezi Park protests, the largest demonstration and occupation in the history of modern Turkey (Tufekci, 2017, p. 52). Third, activists heavily used Twitter during Brazil's movement to, for example, "share what they saw on the streets and invite others to join in the protests" (Monroy-Hernández & Shapiro, 2013 para. 2). Fourth, since Twitter attracts bots who range from corruption protesters to event promoters and mobilizers of action, there is likely to be theoretically meaningful heterogeneity in how bots enact a social movement role in such platform.

Data Description

Our sample includes Twitter messages embedding three different hashtags associated with the movement: #changebrasil, #changebrazil, and #mudabrasil ("change Brazil" in Portuguese). For every message posted by an actor (tweet, reply, retweet, and mention), we have the username of the initiator, the message's content, URL, and timestamp. We also have the number of accounts following the initiator, the number of followers of the initiator, and the number of updates posted by the initiator. Our tracking of data began on June 16, a day before street protests exploded in breadth and propagated across the country, and ended on June 21, a day after the largest demonstration occurred.

We used several data sources: (1) Twitter's search API to collect our sample, (2) multiple computer algorithms to identify bots, and (3) human intelligence to manually validate the nature of identified bots. Such triangulation in detecting bots conceptually follows the last two stages of the funnel process (see Salge & Karahanna, 2018) and therefore bolsters the confidence of our emergent theory—we are convinced our findings come from analyzing bots (instead of humans) and their activity. Another key strength of our approach is that we explored a variety of publicly available computer algorithms including (but not limited to) the Bot or Not? system API developed by Ferrara et al. (2016)[43] (also called the Truthy project) and the decision tree classifier of Cresci, Di Pietro, Petrocchi, Spognardi, and Tesconi (2015) before we decided to use BotOrNot (ABTO Software[44]) and DeBot (Chavoshi, Hamooni, & Mueen, 2016) to identify the bots in our sample. BotOrNot and DeBot were suitable choices for two reasons. First, they have face validity—that is, BotOrNot and DeBot classify publicly known bots, such as *anonymousfrai, poem_exe* and *thricedotted*, as actual bots. Second, because BotOrNot and DeBot use different methodologies[45], they enable a contrasting algorithmic procedure to bot detection. Paired methodology is recommended (Alvisi, Clement, Epasto, Lattanzi, & Panconesi, 2013; Ferrara et al., 2016) given its holistic analyses of bot features.

The primary data for our analyses involve four types of Twitter messages: (1) retweet, a message forwarded to one's subscribers (followers) from another actor; (2) reply, a response to another actor's message; (3) mention, a message mentioning one or more actors anywhere in the body of text; and (4) tweet, a message that may contain photos, videos, links, hashtags, and have now a maximum of 280 characters of text (at the time of the movement the limit was 140 characters). Our data records 710,697 messages from 252,543 unique actors.

To examine the role of bots in Brazil's 2013 social movement, we arranged the social network data in two ways. First, because Twitter messages are conceptually different from another, we parsed the data to build three distinct communication networks: replies, retweets, and mentions. All graphs are directed and unweighted. Each

---

[43] http://truthy.indiana.edu/botornot/
[44] http://botornot.co
[45] DeBot uses *between-user* activity correlation (strong association is indicative of bot behavior) while BotOrNot uses *within-user* activity patterns (a set of user behavior criteria (e.g., retweet count) is used to detect bot activity).

actor represents a user posting a message about the movement. We added an edge between two actors to the network if one of them replies-to/retweets the other in a message or if one of them mentions the other in a reply/retweet/tweet message. No loops (tweets) are modeled, and all ties are flattened in that a connection between A and B indicates at least one interaction. Second, because the activity of bots may exhibit temporal patterns, we also built social networks for each day of the movement (from June 16 to June 21). In total, we created 18 graphs (six each [one for every day] for replies, retweets, and mentions) with 142,728 distinct actors and 231,548 unique ties.

Data Analysis

Our research is theory-building in nature and so the objective was to induct theory that is accurate, parsimonious, and generalizable. In keeping with this objective, we used theoretical sampling to select focal bots (Eisenhardt, 1989) from the 18 networks we created. Each case was purposefully chosen—in a nonrandom way—for theory-building purposes (Glaser & Strauss, 1967) as to illuminate the focal phenomenon.

As such, our selection focused on bots playing at least one key role in the movement. That is, bots which (1) built many alliance partnerships and were also closely connected to the media (leaders); (2) diffused content about the movement across the network (brokers); (3) shared a lot of movement messages with others (broadcasters); (4) enacted any possible combination of the above roles (e.g., brokers and leaders). By contrast, our approach eliminated bots that supported the movement but did not stand out in any way (common users); for example, they did not bridge information about the movement across the network or just lacked sufficient prestige to create ties regardless of their actions. To quantitatively select focal bots, we followed the social movements literature in that we used three different measures of network centrality as proxies: in-degree to operationalize leadership, out-degree to capture broadcast, and betweenness to quantify brokerage.

We also applied the median plus two times the standard deviation for each centrality measure as to differentiate focal (central) actors from common (non-central) actors. Our criterion is appropriate for several reasons. First, the data is positively skewed. With an asymmetrical distribution, the mean is biased while the median is not. Second, since focal actors can stand out in diverse ways (by being brokers for example), their centrality scores are far above the rest of the scores—i.e., focal actors are outliers. A widely-used approach to identify outliers is to select scores with two or three standard deviations away from the central tendency of a distribution[46] (Kline, 2011). Because our aim was to have useful variation within a subset of outliers, we chose the lowest cutoff value (two instead of three) for the standard deviation. Out of the 142,728 nodes in our sample, 5,861 are focal actors.

We next used BotOrNot and DeBot to identify which focal actors were bots. The algorithms identified 829 out of the 5,861 central nodes as bots (BotOrNot labeled 722 while DeBot marked 188, with an overlap of 81). To avoid Type I error (algorithms labeling an actor as a bot when the node is in fact a person), the first author and a computer science graduate student independently and manually assessed whether each identified focal bot was actually a bot (inter-coder reliability = 0.98). Sourcing the work to a graduate student in computer science was appropriate since previous research indicates they are remarkably accurate in detecting bots (Wang, Mohanlal, Wilson, Wang, Metzger, Zheng, & Zhao, 2012). To rule out Type II error (algorithms labeling an

[46] We initially computed interquartile range (IQR) instead of standard deviation but decided to use the latter given the skewness of our data – many IQR scores were zero and therefore not analytically useful for distinguishing central nodes from non-central actors.

actor as a human when the node is in fact a bot), the coders further rated the validity of 100 randomly selected focal nodes which were identified as humans. None of them were validated as bots (inter-coder reliability = 0.99), indicating that the detection algorithms we chose are fairly accurate in detecting non-bot actors. See Appendix A for additional details on the manual assessment performed by the two raters.

Applying these criteria, the coders agreed that, out of the 829 focal bots identified by the two computerized algorithms, only a total of eight were indeed bots (as illustrated in Table 6). Nearly all of them (*Anon_RT*, *blanketRT*, *instant_RT*, *TeamRevoltNow*, *Toni_JoiaRara*, and *WorldCupRetweet*) joined Twitter on 2013, the year in which the movement occurred. Two do not have written biographies (*blanketRT* and BrasilRetwittes) while three (*BrasilRetwittes*, *FreeportIL*, and *TeamRevoltNow*) have a much larger audience (over 2,000 followers) than the others (less than 200). These three bots have also shared more messages (minimum of 32,319 updates) than the rest (maximum of 14,062 updates). Such statistics indicate that we have significant heterogeneity in bot profile description and activity.

Table 6. Profile Description of Focal Bots

Bot[a]	Join Date	Location	Bio	Photo	Following[b]	Followers[c]	Updates[d]
Anon_RT	2013-06-05	NA	#Anonymous		43	61	1,355
BlanketRT	2013-06-01	NA	NA		5	196	14,062
BrasilRetwittes	2012-09-24	NA	NA		0	2,340	188,672
FreeportIL	2010-02-04	Freeport, USA	Freeport Illinois. Originally called Winnneshiek, affectionately Pretzel City, home to the second debate between Lincoln and Douglas.Followsback #F4F		2,872	2,596	32,319
instant_RT	2013-05-07	NA	Instant #InstantRT to any of our followers who use #InstantRT in		0	94	12,523

			their tweets.. Also you can only use 1 other hashtag outside of #InstantRT.. Enjoy..				
TeamRevoltNow	2013-03-03	NA	@TeamRevoltNow is an open Twitter stream highlighting the ongoing worldwide revolution against Fascism/Oppression #TeamRevolution		234	2,266	97,003
Toni_JoiaRara	2013-06-17	NA	*Character of Thiago Lacerda in #JoiaRara #ToniVidaloka		29	32	2,020
WorldCupReTweet	2013-06-20	All over the world	I will retweet anything related to The World Cup! Keep up to date with the latest News & Gossip with one follow!		7	14	620

[a] All bot names correspond to their public Twitter screen names. They are all *rule-based*.
[b, c, d] Average number of (*following*, *followers*, *updates*) bots accumulated during the movement.
* Translated from Portuguese by the first author

To discover what the design objective of these eight bots was, we first tried to contact their developers for an interview. Unfortunately, we could not find information on them. We next searched for bot scripts in online repositories such as GitHub, but again, we did not find any details. As a result, we resorted to reverse engineering by reproducing messages that they generated (tweets, retweets, or replies) and appeared in (mentions) across six days of the movement. We replicated a total of 17,399 posts, out of which 16,271 were initiated by the bots. Appendix B provides additional details about the reverse engineering process.

Following multiple case study methodology (Eisenhardt, 1989), we began our analysis by writing case descriptions of each bot. These were based on their profile information (see Table 6) and reverse engineering results (see Table 7). We then coded the terms that bots used—e.g., their scripted keywords, hashtags, and mentions—in relation to social movements, using replication logic. While some of them leveraged content features of social media to connect with social causes, others surprisingly did not focus on social movements. *Instant_RT*, for example, used Twitter's hashtags to share messages about Britney Spears. Once constructs and relationships had emerged, we integrated our findings by case. In keeping our use of replication logic, we then tested emerging theoretical relationships across bots.

We also compared the framework that we developed with prior literature to refine our construct definitions, abstraction levels, and theoretical relationships (Eisenhardt, 1989). Existing work was especially insightful in sharpening our definitions and abstraction. We engaged in repeated iterations until we found a close match between theory and data. The resulting model explains and integrates the significant differences in the role of bots that we observed.

Table 7. Reverse-Engineering Results

Bot	Keywords, Hashtags, & Mentions in Bot Script	Association
*Anon_RT	#anonymous, #yal, #yan, #opisrael	*Related to Social Movements*
*TeamRevoltNow	#revolution	
*Toni_JoiaRara	#verasqueumfilhoteunaofogealuta, #forafeliciano, #semviolencia, #todosunidosporumbrasilmelhor, #lacerda, @Toni_JoiaRara	
blanketRT	share, blanket	*Unrelated to Social Movements*
*BrasilRetwittes	brasil, brazil	
FreeportIL	#news, @FreeportIL	
*instant_RT	#britneyspears, @instant_RT	
*WorldCupReTweet	#worldcup, #fifaworldcup	

* Designed on *RoundTeam*—to learn more see https://roundteam.co

Findings

What Roles Do Bots Play in a Social Movement on Twitter?

Using network centrality to operationalize actor role, we find that all bots in our sample are broadcasters—i.e., they are only central in the out-degree measure (see Table 8). We also find that their broadcasting is specific to retweets. Instead of creating, responding, or mentioning others in social movement messages, bots just forward content already crafted by others to their own subscribers. In total, they redirected 672 messages to nearly 7,599 nodes in six days.

Table 8. Network Centrality of Bots in Brazil's 2013 Social Movement				
Bot	**Out-Degree (Broadcaster)**	**In-Degree (Leader)**	**Between (Broker)**	**MS[a]**
Anon_RT	73	0	0	60
blanketRT	8	0	0	5
BrasilRetwittes	97	0	0	79
FreeportIL	7	1	0	1
instant_RT	4	0	0	4
TeamRevoltNow	296	3	5,171	217
Toni_JoiaRara	304	0	0	258
WorldCupReTweet	60	0	0	48

[a] Number of #changebrazil, #changebrasil, and #mudabrasil messages shared by each bot.

Intriguingly, we observe that *FreeportIL* and *TeamRevoltNow* are the only bots with in-degree centrality scores above zero, meaning that they are the only ones being retweeted, replied-to or mentioned by others. *TeamRevoltNow* also has a higher than zero score in betweenness centrality (BC = 5,171[47] on June 20). But even more noticeable is the degree of variance in broadcasting. On June 17, for example, *TeamRevoltNow* retweeted 34 messages while *FreeportIL* just retweeted one—yet, both were central in the out-degree measure on this day. Because of such disparity in message count, we developed a hunch that different structural mechanisms for broadcasting existed. To examine this conjecture, we iterated between the actor role literature in social movements and our analysis of network data.

How Do Bots Enact the Roles That They Play?

Existing literature suggests that activists can act as broadcasters by prolifically sharing social movement content with others (Lotan et al., 2011; Tan et al., 2013; Tremayne, 2014). Supporting this argument, we find that *Message Volume*, defined as behaviors by which actors share many messages on social media[48], is a germane mechanism by which bots broadcast social causes. *TeamRevoltNow* and *Toni_JoiaRara* are prime examples. Together, they retweeted 475

[47] While this is a large score it is not large enough to make *TeamRevoltNow* central in the betweenness measure.
[48] We applied the median plus two times the standard deviation to operationalize *Message Volume*. In our calculations, we considered all actors and not just bots.

tweets about Brazil's movement, which corresponds to an average of 40 shares per day for each bot. But our findings also show that bots can share few messages and still broadcast a cause. For example, on June 17, *FreeportIL* promoted, via retweet, a message asking seven distinct nodes for help in the movement.

> RT @CleideAbade Help us #ChangeBrazil **@frankchelsea @CFetti @Barwar_Sport @candace_lehew24 @djws_ @Jn3_16_21 @FreeportIL** https://t.co/wrluJoQo1k

So although relevant and prominent, *Message Volume* does not fully explain the role of broadcast—*Actor Tagging Volume*, a means by which nodes advantageously use actor tagging features of social media, such as mentions, to share few messages and reach many non-redundant actors[49], is also important.

Research suggests that SMOs and activists broadcast social movements to raise awareness of their cause and to also mobilize and organize collective action (Castells, 2012). Because social media is today so integral to social activism, a large number of movements are now broadcasted and referred to by their own hashtags (Tufekci, 2017)—Twitter's feature incorporating words or phrases preceded by a hash sign (#) and used for categorizing, in a linguistic way, messages into topics[50]. For example, to promote the Tahrir uprising in January 25 of 2011, some activists composed and shared a lot of tweets with **#jan25, #tahrir, #mubarak,** and **#egypt** (Oh et al., 2015). Similarly, we discover that some bots refer to social movements via hashtags—i.e., our reverse engineering results show that some of them are coded to retweet messages with movement related hashtags (see Table 7). *Anon_RT* represents a case example. The bot, which broadcasted political causes supported by Anonymous, a loosely associated

---

[49] To operationalize *Actor Tagging Volume*, we also used the median plus two times the standard deviation. Here, we selected actors who shared "a few messages" (i.e., number of messages shared *below* median+2SD) reaching "many non-redundant nodes" (i.e., out-degree centrality [after eliminating self-loops and duplicate ties] *above* median+2SD).

[50] To learn more please see https://help.twitter.com/en/using-twitter/how-to-use-hashtags.

international network of activist and hacktivist entities, shared messages with four different hashtags: **#anonymous**, **#yal** (YourAnonLive), **#yan** (YourAnonNews), and **#opisrael** (a coordinated cyber-attack against websites perceived as Israeli). Since Anonymous supported the Occupy Movement and the Arab Spring (Coleman, 2014), it was not surprising to us that the bot was also involved in Brazil's cause.

> RT @luizfelberti **#ChangeBrazil** #OccupyBrazil **#YAN** @YourAnonNews Currently took over the whole bridge & overpass in Florianopolis, SC http://t.co/OAIOxhMS8N
>
> RT @arthurklose Site do PT hacked! http://t.co/kjzngCPxiV **#anonymous** #VemPraJanela **#mudabrasil** **#changebrazil** #ProtestoBR
>
> RT @YourAnonLive Photo: group of protesters camping in front of govt,Aos HQ for four days now **#changebrazil** #ProtestoSP **#yal** http://t.co/uD5XReGmmU

We term bots' use of hashtags for social activism as *Movement-Related Content Tagging*, which we define as a means by which bots advantageously use content tagging features of social media to access information associated with social movements. This relationship is therefore not about the profile description, algorithmic type, or main objective of bots. Instead, our emerging construct focuses on *content* features as network bridges which bots leverage to connect with social movements.

We found two additional cases of *Movement-Related Content Tagging* in the data. One is *Toni_JoiaRara*, which forwarded messages mentioning itself (**@Toni_JoiaRara**) together with those embedding three hashtags about Brazil's movement: **#todosunidosporumbrasilmelhor**, which means "everyone united for a better Brazil"; **#verasqueumfilhoteunaofogealuta**, a phrase from Brazil's national anthem; and **#forafeliciano** ("get out Feliciano"), a reference against congressman Marcos Feliciano who proposed the gay cure bill which activists protested against during the movement. The other case is *TeamRevoltNow*, which retweeted tweets associated with political causes more broadly such as **#revolution**.

> RT @umdocinhu **#changebrazil #verasqueumfilhoteunaofogealuta #mudabrasil** #acordabrasil #vemprarua http://t.co/GNQT7guVoT

> RT @PedroLattari **#ChangeBrazil #revolution** #brazil http://t.co/GDMnAec2Qy

Although the social movements literature is silent about any other form of broadcasting that isn't deliberate, we find that some bots amplifying the magnitude of Brazil's movement are actually not associated with social activism but connected with other topics instead. Like the *Movement-Related Content Tagging* bots, they take advantage of Twitter's features too. *instant_RT*, for example, shares messages that mention its own account **@instant_RT** in addition to messages that include **#britneyspears**, a reference to Britney Spears—an American celebrity. Because activists added **#britneyspears** to their **#changebrazil** messages *instant_RT* shared them, and as a result, broadcasted the cause in Brazil.

> RT @DearJohnBr It's not about 20 cents! **#changeBrazil** #Beyonce **#BritneySpears** #LadyGaga #KatyPerry #AvrilLavigne... http://t.co/G3iTI2ohV

> RT @EricaJapaSDR Brasil no mundo!!! **#mudabrasil #changebrazil** #avrillavigne #ladygaga **#britneyspears** #katyperry http://t.co/w3L1qMpl84

We call this type of relationship as *Movement-Unrelated Content Tagging*, and define it as behaviors by which bots advantageously use content tagging features of social media to access information *not* associated with social movements. *WorldCupRetweet* represents another case example. Instead of promoting a celebrity, this bot shared tweets with **#worldcup** and **#fifaworldcup** as to "keep up to date with the latest news" about FIFA's World Cup (see biography in Table 10). Based on our data analysis we know that many activists were not writing these messages to promote the FIFA World Cup. Rather, they were doing so to publicly express their discontent with the excessive amount of government funds allocated to the event, using that together with Brazil's lack of public infrastructure and history of political corruption to legitimize the cause. Their use of hashtags **(#worldCup** and **#fifaworldcup**), however, is what triggered *WorldCupRetweet* to share messages about the movement.

> RT @JolineV "We don't need stadiums, we need education" http://t.co/VKlWAbLLDh #**ChangeBrazil #WorldCup** #FIFA #Brazil2014 via @CarlaDauden

> RT @brunoborges This is what happens when politicians steal rather invest in education: #Brazil protesting http://t.co/gzBwzM4K6I **#ChangeBrazil #WorldCup**

We find three additional case examples of *Movement-Unrelated Content Tagging*: *BrasilRetwittes*, *blanketRT*, and *FreeportIL*. Since the first two were not scripted to retweet hashtagged messages (see Table 7) they interacted with the movement through *keywords*—"brasil" for the former and "share" for the latter.

> RT @adoreppk O **BRASIL** ACORDOU! **#MudaBrasil**. http://t.co/Iud6rAA0hm

> RT @ODuduSerra PLEASE **SHARE** THIS !!! **#ChangeBrazil**: http://t.co/xRGCiZUm8z via @youtube

While *FreeportIL* tagged its content with **#news** the bot only forward messages which mentioned itself (**@FreeportIL**). *FreeportIL* therefore connected with the cause via *mentions*.

> RT @CleideAbade Help us **#ChangeBrazil** @frankchelsea @CFetti @Barwar_Sport @candace_lehew24 @djws_ @Jn3_16_21 **@FreeportIL** https://t.co/wrluJoQo1k

In short, while our findings from social network analysis show that bots broadcast a movement via *Message Volume* or *Actor Tagging Volume*, our reverse engineering results suggest that they bridge information about a cause across the network via *Content Tagging Diversity*—i.e., by merging *Movement-Related Content Tagging* and *Movement-Unrelated Content Tagging*. Bots sharing many messages on social media or just a few messages that mention many non-redundant actors and tagging, in *all* of these messages, content associated (e.g., **#anonymous**) and not associated (e.g., **#worldcup**) with social movements connect separate clusters of actors and thus expose the focal cause to a new audience, one that is not comprised of activists, bystanders, constituents, or opponents.[51]

Table 9 illustrates the new mechanisms and provides additional example messages for each. Figure 5 summarizes the empirical findings.

Table 9. BrokeCasting Social Movements on Social Media

Role	Mechanism	How	Example Message
Broadcast	Actor Tagging Volume	Tag many non-redundant *actors* in a few messages	Help us **@frankchelsea @CFetti @Barwar_Sport @candace_lehew24 @djws_ @Jn3_16_21 @FreeportIL** https://t.co/wrluJoQo1k
Brokerage	Content Tagging Diversity	Tag *content* associated and not associated with social movements in the same message	This video explains why a million protested last night http://t.co/U7dcoBjZQm **#ChangeBrazil #WorldCup**
			Help us **#ChangeBrasil @FreeportIL** https://t.co/wrluJoQo1k
			The truth about **Brazil #changebrazil** http://t.co/RAXkXBXTEK
BrokeCast	Actor Tagging Volume + Content Tagging Diversity	Tag many non-redundant *actors* in a few messages together with *content* associated and not associated with social movements	Help **Brazil #MudaBrasil #WorldCup #BritneySpears @frankchelsea @CFetti @Barwar_Sport @djws_ @Jn3_16_21 @FreeportIL** https://t.co/wrluJoQo1k

[51] See Appendix C for additional details.

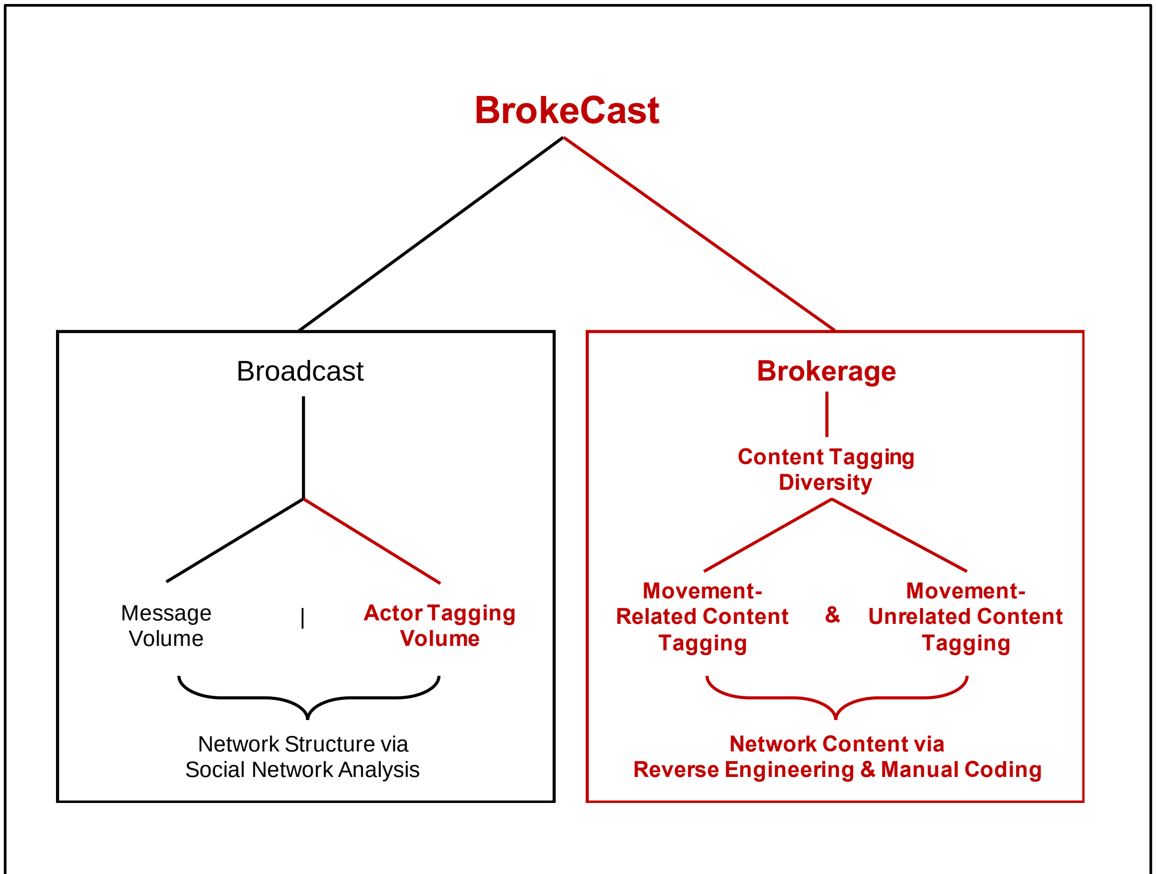

Figure 5. Summary of Empirical Findings

* Note: Mechanisms in **red** are new to the literature.

How to Leverage the Role of Bots for Movement Awareness on Social Media?

Since awareness of a cause is related to brokerage (see Han, 2009; Tremayne, 2013) and our theory building approach goes from data to theory, we abstract the empirical findings of this study to a model of *BrokeCasting* (see Figure 6). The model represents a multi-level network (social actors and tagged content as nodes) and delineates three mechanisms by which bots raise awareness of a movement on social media. The first relies on *Message Volume* and is already documented in the literature (see Tan et al., 2013; Tremayne, 2014). *Message Volume* is concerned with network structure and is part of broadcast. Here, bots accumulate a lot of volume by sharing many messages on social media. This mechanism increases network size—i.e., the number of actors exchanging at least one message in a specific platform—since each bot account represents an additional node exchanging not one but many messages.

The second principle relies on *Actor Tagging Volume* and is also part of network structure and specifically of broadcast. Broadcast in this form ensures that bots advantageously use actor tagging features, such as mentions, to share few messages and reach many non-redundant actors. This mechanism foregrounds size and so it includes more nodes in the network. While *Message Volume* and *Actor Tagging Volume* can be merged—i.e., program bots to leverage social media features to effectively share many social media messages which effectively mention many non-redundant actors—we separate the two since our results show them to be independent from one another. Yet we emphasize that interaction effects are probable as well.

The third mechanism of our framework relies on network content—i.e., *Content Tagging Diversity*—and is part of brokerage. Here, bots secure access to social movement networks together with other kinds of unrelated networks prior to the start of broadcasting. That is, before they execute scripted commands and thus begin to generate volume, bots are scripted to take advantage of features rich in content, such as hashtags, to access diverse networks and the corresponding information created by their actors, including that of social movements. Bots start to bridge information across social media when they collectively use these features to not only connect with a cause, but to also reach new audiences that are not associated with social movements. They do so by jointly including movement-related and movement-unrelated content tags in the same message. Overall, this form of brokerage requires bots to capitalize on content features to increase the reach of a movement—i.e., the number of actors exposed to messages associated with the cause—and thus raise its awareness. This works well because non-bot actors use these same features to self-select into distinct sub-networks of topics and nodes of their own interest—“people [on social media] live in separate worlds. Even hashtags, meant to help users find information on a certain topic, lead them to different bubbles" (*The Economist*, 2017:5 para).

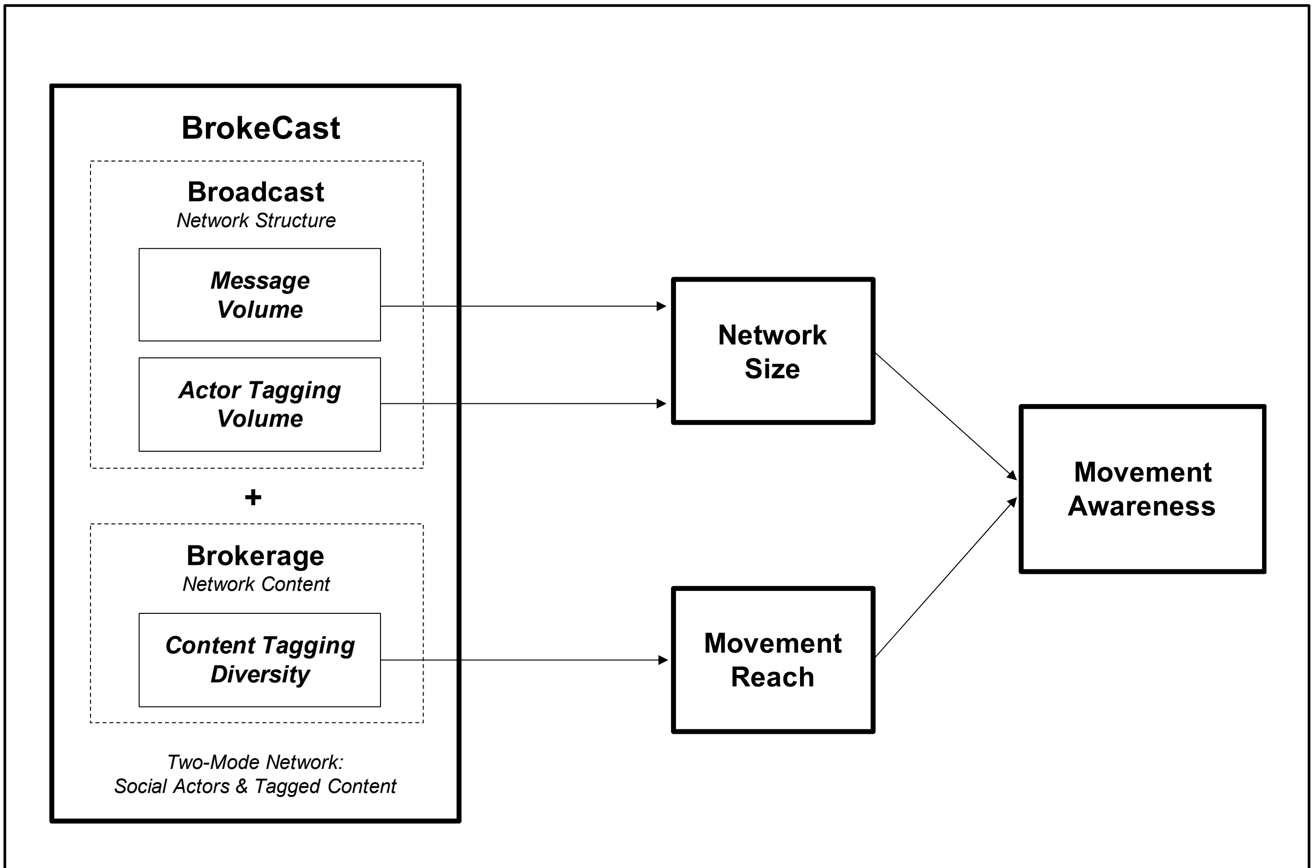

Figure 6. The BrokeCasting Framework

Discussion

We provide several contributions to research. A primary contribution is identifying a new role for activists and SMOs, the *BrokeCaster*. Prior work has indicated that social movements benefit from actors with numerous network ties (Diani, 2003; Han, 2009; Lotan et al., 2011; Tan et al., 2013). Our main insight is that activists that use bots can leverage tagging features of social media to raise awareness of a movement not only by spreading messages about their cause to new audiences but by also amplifying its size. Thus, although bridging and outgoing ties are germane and numerous on their own, bridging and outgoing ties are stronger together. Without a significant size or reach, it is unlikely that movements can be publicly known.

A secondary contribution is a theoretical framework which delineates three mechanisms for *how* bots *BrokeCast*. The first two build on network structure—i.e., *Message Volume* and *Actor Tagging Volume*—and belong to broadcast. While *Message Volume* is already documented in the literature *Actor Tagging Volume* is not. Yet this mechanism is important since the expansion of automation is strategic. Here, bots use actor tagging features of social media, such as mentions, to reach many actors at a low message cost. The final mechanism of our emerging model relies on *Content Tagging Diversity*. Whereas the literature focuses on network structure and attributes of nodes to explain information diffusion (e.g., Aral, Muchnik, & Sundararajan, 2009), this third mechanism builds on network content and relies on bots leveraging content tagging features of social media, such as hashtags, to access and spread social movement messages across the platform. Brokerage in this shape and form is novel and important since it considers the sociology of information for studying the value of structural holes (Burt, 2005). Similar to structural diversity, we find that content diversity is germane for the separation between non-redundant actors. Altogether, our *BrokeCasting* framework suggests that botivists[52] with high-bandwidth outgoing ties rich in content diversity are likely to raise awareness of a cause on Twitter by increasing its network size and widening its reach. *BrokeCasting* is thus an additional contribution to the nascent body of work in IS and social movements (e.g., Vaast et al., 2017; Oh et al., 2015; Miranda et al., 2016) in that it advances our current understanding about the synergies of technology and social activism. An important direction for future research on *BrokeCasting* and specifically on the mechanisms of *Actor Tagging Volume* and *Content Tagging Diversity*, is to explore the issue of susceptibility which asks the question of *whom* and *what* to include in bot-generated messages rather than *how many* and *what kind* to include. We

[52] A term coined by Savage, Monroy-Hernandez, & Hollerer (2016) to refer to activist bots.

suspect that actors with lots of subscribers (@britneyspears[53]) are likely to be tagged together with popular content related (#anonymous) and unrelated (#share) to social movements simply because of their large audiences and great volume. Yet some of these tags may not be susceptible due to a variety of reasons, including a lack of interest or fear of retaliation of those infected.

Finally, we contribute to the bot detection literature in computer science. Given our findings it becomes important for scholars to consider *BrokeCasting* when identifying bots. Doing so can significantly improve the accuracy to which we detect them. Consider the following as an example. An actor sharing an average of 40 messages per day on Twitter is probably not currently flagged as a bot since this level of activity is substantial but not un-human. But, what if this node is sharing 40 messages a day which are tweets with the same set of hashtags (e.g., **#changebrazil**, **#anonymous, #britneyspears, #revolution,** and **#worldcup**) and keywords (e.g., **share** and **brasil**) mentioning the same set of actors (e.g., **@frankchelsea, @cfetti, @barwar_sport, @candace_lehew24, @djws,** and **@jn3_16_21**)? While it is unlikely for humans to repeatedly display this type of activity pattern it is possible for bots to operate this way. Thus, researchers broadly focusing on what differentiates bots from humans are likely to miss *BrokeCasting* patterns hidden in a set of features interacting with one another.

Boundary Conditions

With all research, boundary conditions are germane to theoretical generalizability (Garg & Eisenhardt, 2017). Our *BrokeCasting* framework is specific to bots on Twitter. Yet, it may also fit other social media platforms. For example, Reddit, Instagram, and Facebook have similar actor and content tagging features as Twitter. These platforms permit cross-platform posting of content too. Bots can, for example, be programmed to instantly share the same message on

[53] The American celebrity has nearly 57 million subscribers on Twitter (see https://twitter.com/britneyspears).

Instagram, Facebook, and Twitter. This is important because the degree of standardization of a feature can certainly motivate its use and extend its impact across platforms. Our framework is also applicable to bots beyond social movements. *BrokeCasting* can, for instance, be applied to a variety of marketing products and services and to the spread of information across the network.

Conclusion

We discover that bots play the role of *BrokeCasters* in a social movement on Twitter. With social network analysis, we find that they enact the role of broadcasting by sharing many messages on social media (*Message Volume*) or by leveraging *actor* tagging features to share few messages and reach many non-redundant nodes (*Actor Tagging Volume*). In reverse engineering their design, we find that bots play the role of brokers by advantageously using *content* tagging features to access information associated and not associated with social movements (*Content Tagging Diversity*). Finally, we abstract our findings to a theoretical model of *BrokeCasting* which SMOs and activists can use to raise awareness of a movement on social media.

References

Alvisi, L., Clement, A., Epasto, A., Lattanzi, S., & Panconesi, A. 2013. Sok: The evolution of sybil defense via social networks. ***In 2013 IEEE Symposium on Security and Privacy. IEEE***, *382–396, doi:10.1109/SP.2013.33.*

Aral, S., Muchnik, L., & Sundararajan, A. 2009. Dinstinguishing influence-based contagion from homophily-driven diffusion in dynamic networks, ***PNAS*** 106:51, 21544–21549.

Beutel, A., Xu, W., Guruswami, V., Palow, C., & Faloutsos, C. 2013. Copy-Catch: Stopping group attacks by spotting lockstep behavior in social networks. ***In Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee***, *119–130, doi: 10.1145/2488388.2488400.*

Burt, R. 1992. ***Structural holes: The social structure of competition.*** Cambridge, UK: Harvard University Press.

Burt, R. 2005. ***Brokerage and closure: An introduction to social capital***. New York: Oxford University Press.

Carrascosa, J. M., Cuevas, R., Gonzalez, R., Azcorra, A., & Garcias, D. 2015. Quantifying the economic and cultural biases of social media through trending topics. ***PLoS ONE***, *10:7 1–14, doi: 10.1371/journal.pone.0134407*.

Castells, M. 2012. ***Networks of outrage and hope: Social movements in the Internet age.*** Cambridge, UK: Polity Press.

Cao, Q., Yang, X., Yu, J., & Palow, C. 2014. Uncovering large groups of active malicious accounts in online social networks. ***In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM***, *477–488, doi: http://dx.doi.org/10.1145/2660267.2660269.*

Chavoshi, N., Hamooni, H., & Mueen, A. 2016. DeBot: Twitter bot detection via warped correlation. In ***Proceedings of the IEEE International Conference on Data Mining***.

Coleman. G. 2014. ***Hacker, hoaxer, whistleblower, spy: The many faces of Anonymous***. New York, NY: Verso.

Constantinides, P., & Barrett, M. 2015. “Information Infrastructure Development and Governance as Collective Action,” ***Information Systems Research*** 26:1, 40–56.

Cresci, S., Di Pietro, R., Pretocchi, M., Spognardi, A., & Tesconi, M. 2015. Fame for sale: Efficient detection of fake Twitter followers. ***Decision Support Systems***, *80: 56–71, doi: 10.1016/j.dss.2015.09.003*.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. 2016. BotOrNot: A system to evaluate social bots. ***Working Paper.*** 2 pages.

Diani, M. 2003. ***'Leaders' or brokers? Positions and influence in social movement networks***. In Social Movements and Networks, edited by Mario Diani and Doug McAdam. Oxford: Oxford University Press.

Eisenhardt, K. M. 1989. Building theories from case study research. ***Academy of Management Review***, 14:532–550.

Eisenhardt, K. M., & Graebner, M. E. 2007. Theory building from cases: Opportunities and challenges. ***Academy of Management Journal***, 50: 25–32.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. 2016. The rise of social bots. ***Communications of the ACM.*** *59:7 96–104, doi:* http://dx.doi.org/10.1145/2872518.2889302.

Forelle, M., Howard, P., Monroy-Hernandez, A. & Savage, S. 2015. Political bots and the manipulation of public opinion in Venezuela. ***Essay.*** *1–8*.

Freitas, C. A., Benevenuto, F., Ghosh, S., & Veloso, A. 2014. Reverse engineering socialbot infiltration strategies in Twitter. ***arXiv preprint arXiv:1405.4927***, 25–32, doi: 10.1145/1851275.1851226.

Garg, S. & Eisenhardt, K. 2016. Unpacking the CEO-board relationship: How strategy-making happens in entrepreneurial firms. *Forthcoming at* **Academy of Management Journal**.

Ghobadi, S., & Clegg, S. 2015. "'These Days Will Never Be Forgotten . . .': A Critical Mass Approach to Online Activism," ***Information and Organization*** (25:1), Elsevier Ltd, 52–71.

Ghonim, W. 2012. ***Revolution 2.0 :the power of the people is greater than the people in power : a memoir*** Boston: Houghton Mifflin Harcourt.

Glaser, B., & Strauss, A. 1967. ***The discovery of grounded theory: Strategies of qualitative research***. London: Wiedenfeld and Nicholson.

Gökçe, O. Z., Hatipoğ, E., Göktürk, G., Luetgert, B., & Saygin, Y. 2014. Twitter and politics: Identifying Turkish opinion leaders in new social media. ***Turkish Studies***, 15:4, 671-688, doi: 10.1080/14683849.2014.985425.

Goldman, D. 2014. 23 million Twitter users are fed by robots. The full title of the work is part of http://money.cnn.com/2014/08/12/technology/social/twitter-bots/. Retrieved 5/17/2016.

González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. 2013. Broadcasters and hidden influentials in online protest diffusion. ***American Behavioral Scientist.*** *57:7 943–965.*

Granovetter, M. 1973. The strength of weak ties. ***The American Journal of Sociology***, *78:6 1360–1380, doi: 10.1086/225469.*

Grant, R. 2014. Facebook has no idea how many fake accounts it has – but it could be nearly 140M. The full title of the work is part of http://venturebeat.com/2014/02/03/facebook-has-no-idea-how-many-fake-accounts-it-has-but-it-could-nearly-140m/. Retrieved 5/17/2016.

Han, S. K. 2009. The other ride of Paul Revere: Brokerage role in the making of the American revolution. ***Mobilization***, *14: 143–162*.

Harris, A. 2016. Stay woke-with help from a bot. The full title of the work is part of https://www.fastcompany.com/3061995/stay-woke-with-help-from-a-bot. Retrieved 4/25/2017.

King, B. G. 2008. A political mediation model of corporate response to social movement activism. ***Administrative Science Quarterly***, *53:3 395–421, doi: 10.2189/asqu.53.3.395*.

Kline, R. B. 2011. ***Principles and practices of structural equation modeling***. New York: The Guilford Press.

Ling, C. L. M., Pan, S. L., Ractham, P., & Kaewkitipong, L. 2015. ICT-Enabled community empowerment in crisis response: Social media in Thailand flooding 2011. ***Journal of the Association for Information Systems*** *16:3, 174–212*.

Lipkin, M. 2014. LinkedIn sues bots over member data scraping. http://www.law360.com/articles/499109/linkedin-sues-bots-over-member-data-scraping. Retrieved 5/17/2016.

Lotan. G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & Boyd, D. 2011. The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. ***International Journal of Communication***, *1375–1405.*

Malinick, T. E., Tindall, D. B., & Diani, M. 2013. Network centrality and social movement media coverage: A two-mode network analytic approach. ***Social Networks***, *35: 148–158.*

McCarthy, J. D., and Zald, M. N. 1977. Resource mobilization and social movements: A partial theory. ***American Journal of Sociology,*** *82:6, 1212–1241, doi: 10.1086/226464*.

Miranda, S. M., Young, A., & Yetgin, E. 2016. Are social media emancipatory or hegemonic? Societal effects of mass media digitization in the case of the SOPA discourse. ***MIS Quarterly***, *40:2 303–329*.

Monroy-Hernández, A., & Spiro, E. 2013. How Brazilian protesters are using Twitter. The full title of the work is part of http://www.theguardian.com/news/datablog/2013/jul/04/brazilian-protesters-twitter-microsoft. Retrieved 7/30/2015.

Morales, A. J., Losada, J. C., & Benito, R. M. 2012. Users structure and behavior on an online social network during a political protest. ***Physica A,*** *391, 5244–5253.*

Morstatter, F., Wu, L., Tahora, N. H., Carley, K., & Liu, H. 2016. A new approach to bot detection: Striking the balance between precision and recall. ***ASONAM.***

Naaman, M., Becker, H., & Gravano, L. 2011. Hip and trendy: Characterizing emerging trends on Twitter. ***Journal of the American Society for Information Science and Technology,*** *62(5), 902–918.*

Oh, O., Eom, C. M., & Rao, H. R. 2015. Research note—Role of social media in social change: An analysis of collective sense making during the 2011 Egypt revolution. ***Information Systems Research,*** *26:1 210–223, doi: http://dx.doi.org/10.1287/isre.2015.0565*.

Okalow, S. 2015. Investigating social media's spam-bot problem. http://www.b2bnn.com/2015/08/investigating-social-medias-spam-bot-problem/. Retrieved 5/17/ 2016.

O'Reilly, L. 2015. 8% of Instagram accounts are fakes and 30% are inactive, study says. The full title of the work it is part of http://www.businessinsider.com/italian-security-researchers-find-8-percent-of-instagram-accounts-are-fake-2015-7

Paradise, A., Puzis, R., & Shabtai, A. 2014. Anti-reconnaissance tools: Detecting targeted socialbots. ***Internet Computing,*** *18:5 11–19, doi: 10.1109/MIC.2014.81*.

Perez, S. 2011. Tumblr acknowledges its growing spam problem, says it's doing everything it can. The full title of the work it is part of https://techcrunch.com/2011/10/31/tumblr-acknowledges-its-growing-spam-problem-says-its-doing-everything-it-can/ Retrieved 8/23, 2016.

Salge, C. & Karahanna, E. 2018. Protesting corruption on Twitter: Is is it a bot or is it a person? *Forthcoming at* ***Academy of Management Discoveries***.

Savage, S., Monroy-Hernandez, A., & Hollerer, T. 2016. Botivist: Calling volunteers to action using online bots. ***In Proceedings of the 19th ACM Computer-Supported Collaborative Work and Social Computing****, 813–822,* ACM.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & and Menczer, F. 2017. The spread of misinformation by social bots. ***Technical Report*** 1707.07592, arXiv.

Selander, L., & Jarvenpaa, S. L. 2016. "Digital Action Repertoires and Transforming a Social Movement Organization," ***MIS Quarterly*** (40:2), 331–352.

Tan, L., Ponnam, S., Gilham, P., Edwards, B., & Johnson, E. 2013. Analyzing the impact of social media on social movement: A computational study on Twitter and the Occupy Wall Street movement. ***ASONAM***.

The Economist. 2017. In praise of serendipity. The full title of the work is part of https://www.economist.com/news/books-and-arts/21718464-social-media-should-encourage-chance-encounters-not-customised-experiences-praise. Retrieved 3/14/2018.

Tremayne, M. 2014. Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street. ***Social Movement Studies,*** *13:1, 110–126.*

Tufekci, Z. 2014. What happens to #Ferguson affects Ferguson: Net neutrality, algorithmic filtering and Ferguson. The full title of the work is part of https://medium.com/message/ferguson-is-also-a-net-neutrality-issue-6d2f3db51eb0. Retrieved 11/18/2015.

Tufekci, Z. 2017. ***Twitter and tear gas: The power and fragility of networked protest***. London, UK: Yale University Press.

Vaast, E., Safadi, H., Lapointe. L, & Negoita, B. 2017. Social media affordances for connective action: An examination of microblogging use during the Gulf of Mexico oil spill. ***MIS Quarterly***, *41:4 1179–1209*.

Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H., & Zhao, B. Z. 2013. You are how you click: Clickstream Analysis for sybil detection. ***In USENIX Security***, 241–256.

Wang, G., Mohanlal, M., Wilson, C., Wang, X., Metzger, M., Zheng, H., & Zhao, B. Y. 2012. Social Turing tests: Crowdsourcing Sybil detection. ***Social Information Networks***.

Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. B. 2010. "Web 2.0 and Politics: The 2008 U.S. Presidential Election and an E-Politics Research Agenda," ***MIS Quarterly*** 34:4, 669–688.

Watts, J. 2013. Brazil protests: president to hold emergency meeting. The full title of the work is part of https://www.theguardian.com/world/2013/jun/21/brazil-protests-president-emergency-meeting. Retrieved 4/25/2017.

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

Appendix A

Manual Assessment of Bot Detection

In manually assessing the nature of the 829 identified bots, we followed the work from Wang et al. (2013) by performing two tasks. First, we separately viewed the Twitter profile description of each and every actor identified as a bot by BotOrNot and DeBot. This included their biography, join date, location, photo, number of followers, following, updates, and a variety of messages that they generated. We then used that information to classify them as bots, cyborgs, or humans. For every user, we also wrote a sentence explaining our decision. For example, when categorizing instant_RT as a bot, the first author wrote "its bio describes the account as an instant retweeter" while the graduate student in computer science wrote "the user uses https://roundteam.co/ to retweet." We performed the coding in four different rounds—started with a case sample of 15 then moved to 100, 200, and finally to code them all (829). After every round, we met to compare and discuss our results and to also learn from each other. We performed the exact same process to code the 100 actors identified as humans. But because we coded them last, we only used one round for that task.

Two important lessons can be taken away from this manual process. First, we probably missed out on some bots—going from 829 to eight is indeed a drastic change in sample size. However, if we erroneously coded them as humans or cyborgs it is because their profile characteristics and activity resembled that of humans—that is, they either behaved like humans or had a profile description that seemed human (or actually did both). These bots were therefore more sophisticated than the ones in our final sample. Because the purpose of our paper is to generate theory pertinent to bots, we are tolerant of this type of error—i.e., we accept the fact that we probably classified some of the more sophisticated bots as humans or cyborgs. This does not bias our findings—if anything, it shows that they are more bots on social media than we think there are which means that the magnitude and ramifications of our work are somewhat conservative. The other thing that we learned is that BotOrNot and DeBot classified a lot of cyborgs (total of 96) as bots. We suspect this is because of volume—since cyborgs are *partly* automated, this activity portion of the account must be suspiciously high for the algorithm to erroneously flag it as *completely* automated.

Appendix B

The Reverse Engineering Process

We followed a systematic approach of readying the data for reverse engineering. First, we explored the text for each bot separately and together. We specifically computed and visualized the number of messages they shared (for each and every day) including retweets, replies, mentions, and tweets. This was useful for identifying temporal patterns and also preferred Twitter design feature for communication (retweet vs reply vs tweet vs mention). Second, we cleaned the data by removing all punctuations (with an exception of hashtags) and by lowering the case of every word and phrase. We also changed the data encoding to ASCII as to eliminate accents. Third, we removed stop words in English and Portuguese, created a document matrix for the corpus of text, and calculated the frequency of each term, grouping the results in descending order. Visualizing term frequency for each bot was especially insightful because it provided us with a starting point for reproducing their messages. For example, for *WorldCupRetweet* we discovered that "#worldcup" was most frequent term and so we began by reproducing content with this hashtag.

Once the data were ready we conducted the reverse engineering in alphabetic order, starting with *Anon_RT*. It became clear to us early in the process that *Anon_RT* was designed to promote *Anonymous*, a network of hacktivists opposing Internet censorship and control. To promote Anonymous, *Anon_RT* broadcasted (via retweets) four different hashtags: '*#anonymous*' (*N*=606), '*#yal*' (*N*=287)—a hashtag for YourAnonLive, '*#yan*' (*N*=80)—an abbreviation for YourAnonNews, and '*#opisrael*' (*N*=17)—a hashtag used to describe cyber-attacks against Israel coordinated by Anons. We also found that the bot was configured on *RoundTeam* and that the company automatically published messages from *Anon_RT's* account

(*N*=34). Finally, we found four other messages that were not posted by the bot (one was a reply and three were mentions).

blanketRT was the next bot that we reverse engineered. The actor was designed to broadcast 'share' (*N*=2,840) and 'blanket' (*N*=41) also via retweets. A large portion of messages we collected about *blanketRT* (*N*=993) were not generated by the bot. One was a reply ('@blanketRT thanks to share our Like, share & win action! did you take part in it? we wish you good luck ;) #share #win #like #facebook') while the others were mentions (*N*=11), and hashtags (*N*=981). We also discovered that *blanketRT* shared the same two messages over and over again. The first one describes the importance of showers in people's lives: 'Showers are amazing. They make you feel nice and clean, make you sound like a professional singer, and help you make all of life's decision!' while the second associates Twitter functionalities with specific technology devices '#Rt for #iPhone, #fav for #android, #followme for #desktop.' Together, the two tweets were shared 189 times (95 and 94 respectively) in a period of six days.

The third bot we examined was *BrasilRetwittes*. As its Twitter's screen name suggests, the actor was coded to broadcast content about Brazil in that it retweeted messages containing 'brasil' (*N*=2,098) and 'brazil' (*N*=730). Like *Anon_RT*, *BrasilRetwittes* was also designed on *RoundTeam.* Ninety tweets were automatically published by the company on the bot's account. Finally, *BrasilRetwittes* was mentioned twelve times and replied-to in thirteen messages.

FreeportIL, the only bot to broadcast Brazil's 2013 political social movement via *Node Volume*, was the next actor we reverse engineered. Initially, we found a total of 21 messages not generated by the bot—they were all mentioning *@FreeportI.* Four of these though were actually retweeted the *FreeportIL*. We also noticed, by reviewing the bot's Twitter activity, that most of its tweets were news coming from a British video sharing website called *LiveLeaks*.[54] Because *FreeportIL* repeatedly shared a lot of content from the same source in a seemingly standardized manner, we developed a hunch that the bot was coded to promote *LiveLeaks* by disseminating its news via automated tweets which added "#news" at the end of each message. We judged the validity of our hunch in two probes. First, we counted how many messages generated by *FreeportIL* embedded a URL. If our suspicion was robust we had to at least observe a connection from the bot's content to the video sharing website. Indeed, we found that all tweets had links. But, were they *all* associated with *LiveLeaks*? To answer this question, we fetched the redirects of all URLs. Our results show that 562 out of the 569 links came from *LiveLeaks* while the remaining 7 came from *EnterpriseWayWire*, an 'online video curation platform' allowing enterprises to 'aggregate video content and integrate it with their video production efforts.'[55] Because nearly 99 percent of *FreeportIL's* messages were news feeds from *LiveLeaks*, we deduced that the bot was coded to broadcast live news from the British website.

The fifth case we selected was *instant_RT*. This bot is interesting because some of its design properties are provided in its profile description (see Table 10). For example, its biography alludes to the idea that every message posted by a subscriber with '#InstantRT' is likely to be shared by the bot. Yet, our analysis show that most of its generated content do not include '#InstantRT.' Instead, they are retweets which contain '#britneyspears' (N=1,223), a reference to Britney Spears, an American singer, dancer, and actress. We also discovered that, similar to *Anon_RT* and *BrasilRetwittes*, *instant_RT* was also deployed by *RoundTeam*. Forty-one of its messages were automatically shared by the service company. The remaining five were

[54] We obtained this information by clicking on embedded links within *FreeportIL*'s tweets.
[55] See more here: http://enterprise.waywire.com/company/about-us/

either replies or mentions. In short, we found that *instant_RT* is a bot coded to promote Hollywood's celebrity Britney Spears.

TeamRevoltNow is our sixth case. In reading *TeamRevoltNow's* messages, we quickly learned that the bot was designed to promote political revolutions around the globe, which is corroborated by its profile description (see Table 10). Nearly all of its messages were retweets including '#revolution' (4,072 out of 4,182 which is 97.4%). We also found a total of five replies and sixty-five mentions. Finally, *TeamRevoltNow* was designed by *RoundTeam*—the company published a total of ninety-three messages on the bot's behalf.

Toni_JoiaRara is another bot configured on *RoundTeam*. The company automatically generated fifty-nine tweets for *Toni_JoiaRara*. We also found that the bot was *specifically* coded to support Brazil's 2013 political social movement. Its account only shared messages containing movement-specific hashtags such as '#verasqueumfilhoteunaofogealuta' (*N*=1,479)—a phrase from Brazil's national anthem that says 'you will see that no son of yours will run from a fight' (i.e., the fight here is the movement); '#todosunidosporumbrasilmelhor' (*N*=227) which translates to 'everyone united for a better Brazil'; '#forafeliciano' (*N*=147)—a reference against congressman Marco Feliciano who proposed the gay cure bill during the movement; and '#semviolencia' (*N*=2) which means 'without violence.' We also found that *Toni_JoiaRara* shared messages mentioning '#lacerda' (*N*=6), last name of Brazil's actor 'Thiago Lacerda' who played 'Toni' in the telenovela 'Joia Rara' aired by Brazil's TV network Globo in 2013. Finally, *DirenAklim* demonstrated solidarity with the movement in a reply to the bot on June 19 ('@Toni_JoiaRara Solidarity from #Brazil & #Turkey http://t.co/Ot06CuVZWi') while six others just mentioned *Toni_JoiaRara* in their messages.

The last bot we examined was *WorldCupRetweet*. Like *BrasilRetwittes*, this actor's screen name exposes its design properties. Specifically, we found that *WorldCupRetweet* was coded to promote news about FIFA's World Cup by retweeting messages with the hashtags '#worldcup' (*N*=1,216) and '#fifaworldcup' (*N*=71). We also found a total of three tweets that were not generated by the bot. In the first one, *RLeagueFan* notices the actor's bot nature. *RLeagueFan* posted "Bahaha. This stupid robot retweeted me about the naming of the women's #RugbyLeague #WorldCup squad @WorldCupReTweet http://t.co/2wY79ycaQ8" on June 20 at 9:30am. In the second tweet, we discovered that *WorldCupRetweet* was also designed on *RoundTeam*. *AgainsttheMosq* shared, on the night of June 21, the following message: "RT @WorldCupReTweet Contributed tweets for @WorldCupReTweet are delivered by @RoundTeam http://t.co/3ova9h8aN0." Finally, in the third message, *RealSamThompson* asks *WorldCupRetweet* for a vote in the night of June 20: "@WorldCupReTweet world cup can you vote for me https://t.co/44njOvaBz8 go leader boards type sam thompson click vote."

Appendix C

Content Tagging Diversity—A New Form of Brokerage

To assess our reverse engineering results in relation to brokerage, we computed content diversity—the degree to which actors associated with different topics actually reached structurally non-redundant nodes—for two bots in our data. Because we do not have a specific case in both *Movement-Related Content Tagging* and *Movement-Unrelated Content Tagging* groups, we first matched two bots based on their profile characteristics and activity. For the analysis, we specifically chose *Anon_RT* and *WorldCupRetweet* for four reasons. First, they joined Twitter on the same month and year (June 2013). Second, they were both designed on *RoundTeam*. Third, these two bots had a profile photo, a biography, and similar number of followers, following, and updates during the movement (see Table 10). Fourth, *Anon_RT* and *WorldCupRetweet* were designed to share the same type of content (i.e., hashtags only). We then used the messages from our reverse engineering to create two separate networks, one for each bot, in the exact same way we constructed the network for Brazil's movement. Finally, we computed the number of overlapping nodes reached by these two bots. Our findings show that together, they reached a total of 1,838 actors (*Anon_RT* = 658 and *WorldCupRetweet* = 1,190), out of which only 10 overlapped (youranonnews, shokufeyesib, hashonomy_gus, youtube, occupybrazilbot, reuters, my2cnz, ap, rt_com, occupywallst) which is about 5.44 percent.

To check the robustness of our findings, we collected additional data on the hashtags that these bots targeted during the week of the movement and re-ran the analysis—i.e., we created two separate networks (one with **#anonymous**, **#yal, #yan**, and **#opisrael** for *Anon_RT* and one with **#worldcup** and **#fifaworldcup** for *WorldCupRetweet*) and counted the number of overlapping nodes again. An even smaller overlap is observed. Out of 57,358 unique actors in the network, only 330 overlapped (0.58 percent). This result shows preliminary support for *Content Tagging Diversity* as a new form of brokerage. That is, bots broadcasting a cause and tagging, in the messages that they share, content features associated (e.g., **#anonymous**) and not associated (e.g., **#worldcup**) with social movements reach separate networks and therefore expose the focal cause to a new and large audience, one that is not just comprised of activists, bystanders, constituents, or opponents of a movement.

Chapter 4. Botivism & Resource Mobilization In Social Movements[56]

[56] Salge, Carolina; Karahanna, Elena; & Li, Weifeng. To be submitted to Management Science.

Abstract

With the growth and potential impact of bots—automated accounts in online social networks—many social movement organizations and activists have begun using them on social media platforms to engage participation and mobilize resources for their cause. However, there has not been much research on the thematic content of bot-generated messages on these platforms and, consequently, their impact on mobilizing social movement resources. In this paper, we examine 12,604 posts generated by bots during Brazil's 2013 political movement on Twitter to understand *what* topics bots share on this platform and *how* topic relations and content characteristics can be leveraged for resource mobilization on social media. Our analysis demonstrates that bots engage in two forms of social activism. The first is *Focal Botivism* and requires a group of bots to share positive and negative messages in correlated topics which altogether promote social change in a specific population and for a specific cause. The other is *Global Botivism*, which we define as a means by which a group of bots share positive and negative messages in correlated topics which altogether promote social change across the globe. We abstract our findings to offer a theoretical framework which explains how *Focal Botivism* and *Global Botivism* mobilize resources for social movements on social media. Overall, *Botivism* is important since the accumulation of resources is crucial for social movement success.

Keywords: bots, social movements, Twitter, topic modeling, resource mobilization.

The increasing use of technology has transformed the way in which social movement organizations[57] (SMOs) and activists pursue social change (Tufekci, 2017). In addition to protesting through face-to-face channels such as street demonstrations, many are now using bots—automated accounts in online social networks—to engage with people on social media platforms like Facebook and Twitter. Bots on Twitter, for instance, constitute between 9 and 15 percent of its 330 million active users (Varol, Ferrara, Davis, Menczer, & Flammini, 2017). In some cases, these bots are coded to raise awareness and support the goals of a movement. For example, *ONG Nossas* designed *Beta* to inform activists of political initiatives against women's rights so that they are better equipped to take action when needed.[58]

Despite recent investments from SMOs and activists, there has been limited empirical investigation of bot generated content during social movements on platforms like Twitter. Previous work on bots has focused primarily on detection (e.g., Cresci, Di Pietro, Petrocchi, Spognardi, & Tesconi, 2015; Chavoshi, Hamooni, & Mueen, 2016; Davis, Varol, Ferrara, Flammini, & Menczer, 2016; Salge & Karahanna, 2018) and recent literature in social movements has revolved around social media (e.g., Miranda, Young, & Yetgin, 2016; Oh, Eom, & Rao, 2015; Selander & Jarvenpaa, 2016). While a few studies on bots and social movements do exist this line of work has not yet addressed content variations to develop meaningful theoretical insights. Prior research mainly focuses on understanding the growth and development of bot networks during social crisis (Abokhodair, Yoo, & McDonald, 2015), detecting and filtering politically-motivated spam (Verkamp & Gupta, 2013), quantifying bot communication interference in political protests (Suárez-Serrato, Roberts, Davis, & Menczer, 2016), and

[57] An SMO is a "complex or formal organization which identifies its goals with the preferences of a social movement or countermovement and attempts to implement those goals" (McCarthy & Zald, 1977:1218).
[58] https://revistatrip.uol.com.br/tpm/robo-feminista-beta-alerta-pelo-inbox-do-facebook-sobre-projetos-de-lei-que-ferem-direitos-das-mulheres

designing bots to help activists call new volunteers to action (Savage, Monroy-Hernández, & Hollerer, 2016). Verkamp and Gupta (2014), for example, “find that the nature of [bot] spam varies significantly across [protest] incidents such that [theoretical] generalizations are hard to draw” (p. 1). Analyzing bot generated content is crucial for understanding the interplay between bots and resource mobilization in social movements. First, the reach of bot messages largely depends on their content. As Wu, Huberman, Adamic, and Tyler (2004) point out “information is selective and passed by its host only to individuals the host thinks would be interested in it” (p. 1). Second, the amount of social movement resources acquired by bots relies on the content of the messages that they disseminate. Bots posting emotional content on Twitter, for example, may not mobilize people to action—those reached by such bots may actually perceive their messages to be insincere and manipulative since they know bots do not have instincts and therefore are unable to display emotions. Finally, because social media is infected by many bots which are often times designed to mimic human activity, prior research has shown that they are prominent and potentially impactful—i.e., bots are capable of influencing human behavior (Ferrara, Varol, Davis, Menczer, & Flammini, 2016).

With the objective of contributing to an understanding of how bot generated content relates to resource mobilization in social movements, in this research, we study the thematic content of the messages bots share during Brazil’s 2013 political cause. We ask three questions, *What topics do bots share in a social movement on Twitter? How do these topics relate to each other? How does bot generated content mobilize resources for a social cause?* Relying on a large and detailed dataset, we first apply the funnel process to identify bots (Salge & Karahanna, 2018) and then use topic modeling to answer the first two questions we pose. We then abstract our findings to a model of *Botivism* which SMOs and activists can use for mobilizing social movement resources on social media.

Our study provides two contributions to the IS literature in social movements. First, we empirically analyze content disseminated by bots and identify two mechanisms of bot activism, namely *Focal Botivism* and *Global Botivism*. *Focal Botivism* requires a group of bots to share positive and negative messages in correlated topics which together promote social change in a specific population and for a specific cause. The second mechanism, *Global Botivism*, requires a group of bots to share positive and negative messages in correlated topics which together promote social change across the globe. Second, we offer a theoretical framework which explains how *Focal Botivism* and *Global Botivism* mobilize resources for social movements on social media. Overall, *Botivism* is important since the accumulation of resources is crucial for social movement success.

Literature Review

Resources & Social Movement Mobilization

Resources are sources of supply or support; wealth or revenue; information or expertise which are in reserve or ready for use when needed (Merriam Webster Dictionary definition). Resources are germane for the emergence, development, and outcomes of social movements[59] (Edwards & McCarthy, 2004). Yet access to various resources is not sufficient to achieve social change—mobilization or the process of increasing the readiness to act collectively (Gamson, 1975) is required too so that available and individually held resources can be converted into collective resources which can then be used in collective action. A rich research literature has been accumulated and devoted to the study of resources and social movement mobilization (e.g.,

[59] We define social movements as a set of opinions and beliefs in populations that represent preferences for changing some elements of social structures (McCarthy & Zald, 1977)

Cress & Snow, 1996; Soule, McAdam, McCarthy, & Su, 1999). Of particular importance is the resource mobilization theory, introduced by McCarthy and Zald (1977). At its root, the theory is aimed to better understand how SMOs and activists overcome consistent patterns of resource inequality in their efforts to pursue social change. Its main argument is that stakeholders of a movement, and specifically SMOs, ought to focus their attention on *how* resources are successfully mobilized rather than on *why* people are grieved since the more resources acquired, the less costly it is to establish a movement, and the more likely it is for people to do something about their grievances.

Despite the clear importance of resources to the logic of resource mobilization theory, McCarthy and Zald (1977) do not explain in any great detail the concept of resources along with a clear specification of its types. As a result of the lack of a more granular delineation, the most widely appreciated resources in the literature today are money, people, and SMOs (Edwards & Kane, 2014). This is problematic since it treats "everything as a resource" (Edwards & McCarthy, 2004). To rectify this conceptual issue, Edwards and McCarthy (2004) refine these categories to introduce five different types of resources—material, human, social-organizational, cultural, and moral—which we discuss below.

Material resources are concerned with financial and physical capital and examples include money, property, office space, equipment, and supplies. Because of their tangible nature, they have received great attention from scholars in the literature (e.g., Olzak & Ryo, 2007; Barker-Plummer, 2002; McCammon, Campbell, Granberg, & Mowery, 2001). Human resources are also somewhat tangible—beyond labor these involve human capital (Becker, 1964), such as experience, skills, and expertise. Leadership is also considered a human resource since it represents a combination of labor and human capital. The remaining three kinds of resources are less tangible and thus more difficult to observe and measure. Social-organizational resources, for example, include infrastructures, social networks, affinity groups, and coalitions (McCarthy, 1996) and can be intentional or appropriable (Coleman, 1988). Intentional implies that the social-organizational resource is specifically created for advancing the goals of social movements (e.g., a group of few activists operating locally with little formal structure) while appropriable implies that the social-organizational resource is originally created for non-movement purposes (e.g., the successful recruitment of volunteers through work) and later leveraged for the movement. Both intentional and appropriable forms of social organization are crucial for mobilizing and accessing other types of resources. Together, they can be viewed as social capital (Edwards & Kane, 2014), a structural and relational concept referring to the ability of actors to leverage their social network connections to access a variety of resources (Bourdieu, 1986; Coleman, 1988). Cultural resources involve the tacit and taken-for-granted symbols, beliefs, values, identities, and behavioral norms of a group aimed at orienting and facilitating actions of everyday life. This type of resource also includes social movement issues relevant to productions such as music, literature, films, and videos. A major difference between human and cultural resources is that the former is ingrained in and controlled by those who possess them—people have proprietary control of who benefits from their skills and expertise. Cultural resources, on the other hand, are more difficult to control when they enter the public domain—they can be accessed and used by a wide range of SMOs and activists and across different movements. Finally, moral resources include legitimacy, authenticity, solidarity and sympathetic support, and celebrity endorsement (Cress & Snow, 1996). Like the other four resources, moral resources are interrelated. For example, by endorsing a specific movement or lending their fame to a particular SMO, celebrities increase the legitimacy of a social cause by escalating its media attention (King, 2011). This, in turn, generally increases the movement's ability to further access other resources.

Research in sociology offers five distinct means of accessing and mobilizing the social movement resources abovementioned—self-production, aggregation, co-optation/appropriation, patronage, and moral battery (Edwards & McCarthy, 2004; Jasper, 2011). Self-production is straightforward in that it requires movements to produce resources themselves by the agency of existing SMOs, activists, and other participants (Edwards & McCarthy, 2004). Aggregation occurs when social movements or specific SMOs and activists convert resources held by dispersed actors into a collective (Edwards & McCarthy, 2004). Humans resources, for example, can be aggregated by soliciting technical expertise from separated actors to help with a particular task. Co-optation refers to the borrowing of resources which are allowed, transparent, and already under the control of a social movement group (Edwards & McCarthy, 2004). Studies of social movements in the 1960s and 1970s illustrate this mechanism (see McCarthy & Zald, 1977). They explain how movements co-opted institutional resources from private foundations, social welfare institutions, the mass media, universities, government agencies, and some business corporations. Appropriation is different than co-optation in that the borrowing of resources is surreptitious. The fourth mechanism in the literature is patronage. Patronage is the bestowal of resources by an individual, organization, or network that often specializes in patronage (Edwards & McCarthy, 2004). In monetary relations, patrons offer significant financial support to movements yet they are also likely to sympathize with a cause and thus contribute human resources by providing free labor for specific periods of time (see McCammon & van Dyke, 2010). Finally, moral battery—a pair of messages inciting contrasting sentiment, one positive and the other negative—mobilizes resources for social movements through emotions (Jasper, 2011). These messages elicit anger and indignation, propelling others to act.

Taken together, the literature suggests that resource mobilization is germane for the success of social movements. This body of work also introduces five kinds of resources (material, human, social-organizational, cultural, and moral) which can be mobilized and accessed via five distinct mechanisms (self-production, aggregation, co-optation/appropriation, patronage, and moral battery). While these represent important advancements in our conceptual understandings, the literature still lacks an explanation of how technological advances—and more specifically bots—provide resources to a social cause. Our study addresses this by scrutinizing the thematic content of messages that they generate in Brazil's 2013 political social movement on Twitter.

Thematic Content & Resource Mobilization

Several literatures provide insights into our research questions. First, the *framing* literature (e.g., Benford & Snow, 2000) provides a rich body of qualitative knowledge toward resource acquisition. Studies provide a narrative account of how framing—i.e., "the processes by which [social movement] grievances were constructed, contested, and disseminated" (Snow, Benford, McCammon, Hewitt, & Fitzgerald, 2014)—mobilizes participation and support (Cress & Snow, 2000; Wetzel, 2010), compels action in the absence of political opportunity (Einwohner, 2003), provokes emotional reactions in the direction of (or against) mobilization (Halfman & Young, 2010), challenges and expands the legitimacy of a movement (McCright & Dunlap, 2000; Pedriana, 2006), shapes the trajectory and dynamics of a cause (Oselin & Corrigall-Brown, 2010; Su, 2004), and affects the cycle course of protests (Zwerman, Steinhoff, & della Porta, 2000). Constructing frames facilitates the mobilization of resources by providing a shared understanding of a social problem in need of change (diagnostic framing), by making attributions of blame and articulating an alternative set of arrangements (prognostic framing), and by urging others to act in concert to affect change (motivational framing) (Snow & Benford, 1988). For example, in a study of 15 homeless SMOs in eight US cities, Cress and Snow (2000) find that the Oakland Union of the Homeless (OUH) mobilized extensive resources consisting of a multimillion-dollar housing project by articulating diagnostic and prognostic frames which were disruptive. But, while this literature usefully identifies a variety of action frames and finds that framing facilitates the mobilization of resources in social movements, this line of work does not assess the content of bot action frames spreading through social media or what Bennet and Segerberg (2012) call *personalized communication in large-scale connective action*. Rather, framing is mostly observed in the familiar logic of collective action with a few exceptions (e.g., Vasi, Walker, Johnson, & Tan, 2015).

Second, the *dialogue* literature (e.g., Taylor, Kent, & White 2001) argues that resources are mobilized from building relationships with publics. Dialogue "refers to any exchange of ideas and opinions" (Kent & Taylor, 1998:323) and is guided by two principles. First, actors in a dialogue do not necessarily have to agree with one another—sometimes they may actually strongly disagree—yet these actors must be willing to reach a mutual satisfactory position. Second, dialogue is about inter-subjectivity rather than objective truth or subjectivity. Empirical research in dialogue suggests that activist groups often fail to leverage social media for dialogic communication (Bortree & Seltzer, 2009; Taylor et al., 2001; Rybalko & Seltzer, 2010; Waters & Jamal, Lee, 2011; Lee, VanDyke, & Cummins, 2017; Jahng & Lee, 2018) and instead use these platforms as one way-interactions to, for example, persuade others (Auger, 2013), collect petition signatures and organize events (Obar, Zube, & Lampe, 2012), disseminate information

(Lovejoy, Waters, & Saxton, 2012), and raise awareness, maintain support, and mobilize action (Guo & Saxton, 2014). For example, Jahng and Lee (2018) examine how a group of Virginia Tech scientists used Twitter to communicate and mobilize the public in response to the water contamination in Flint, Michigan. They find that the platform was mainly leveraged for disseminating educational information to others. A particular study does, however, illustrate the use of Twitter for building relationships with publics by identifying three communicative platform functions—information, community, and action (Lovejoy & Saxton, 2012). Information is about the distribution of content associated with organizational activities, event highlights, or any news, deemed as important for stakeholders. Community involves interacting with stakeholders in a way that facilitates the creation of an online community. The third and last function is action. Action is specifically concerned with mobilizing resources. Examples include requests for financial support and event participation and the objective is to encourage activists to "do something" for the cause or organization. While this line of work suggests that dialogue may be crucial for resource mobilization, it does not provide evidence for what dialog, if any, is enacted by bots—e.g., which type of bot action messages are generated toward mobilization.

The *tactics* literature (e.g., Theocharis, Lowe, van Deth, & Albacete, 2015) offers a third lens on how content might be used to mobilize social movement resources. Here, tactics of activists can fall under five different categories—civil disobedience, and activities or messages which are informational, symbolic, organizational, and legalistic (Jackson, 1982). Informational is similar to the information function in the dialogue literature (Lovejoy & Saxton, 2012) in that it involves the broadcasting of news. Symbolic actions are concerned with demonstrating not only where a group of activists stands but also how strongly they feel about a specific cause. Example activities include boycotts, flagpole sitting, and mock funerals or trials. Organizing involves networking and mobilizing others to attend street protests, make donations, and potentially recruit additional members. Finally, legalistic actions include petitions, lawsuits, filling legislation, testimony at hearings, and prodding regulatory and administrative agencies. These acts are seen as legitimate and so they acquire resources by garnering media attention (Jackson, 1982). Studies in tactics find that social movement content is mostly informational (Theocharis et al., 2015; Hinsley & Lee, 2015; Papacharissi & Oliveira, 2012; Penney & Dadas, 2014; Sommerfeldt, 2011), organizational (Agarwal, Bennett, Johnson, & Walker, 2014; Earl & Kimport, 2011; Penney & Dadas, 2014; Das & Taylor, 2010; LeFebvre & Armstrong, 2018), and symbolic (Penney & Dadas, 2014; Wallsten, 2008; LeFebvre & Armstrong, 2018; Ahmed, Jaidka, & Cho, 2017; Agarwal et al., 2014). Yet, while the content of messages is invoked to theorize about resource mobilization (see LeFebvre & Armstrong, 2018), and a few studies even categorize the function of these messages, the actual content of bots is again not examined.

In sum, the framing, dialogue, and tactics literatures offer possibilities for how content can be leveraged for resource mobilization in social movements. They indicate that (1) different types of framing exist and are mostly studied in collective action rather than connective action; (2) dialogic communication is rarely observed in social media; and (3) tactics of activists involve the dissemination of informational, organizational, and symbolic social media posts. But despite the prevalence and potential impact of bots, the content of the messages that they generate during a cause, and how these potentially relate to resource mobilization, is not addressed. Hence, we ask: *What topics do bots share in a social movement on Twitter? How do these topics relate to each other? How does bot generated content mobilize resources for a social cause?*

Methodology

Our research design is a multiple-case, inductive study that utilizes computational data analysis methods. The context of our study is Brazil's 2013 social movement, one of the largest and most significant organized collective efforts in a generation, and also publicly known as the "*20 Cents Movement*." Street protests against a R$0.20—equivalent to $0.09 at the time—increase in bus fares across several Brazilian cities began on June 6, yet these public expressions of discontent rapidly grew to include other political issues such as high-levels of government corruption, lack of investment in public infrastructure, and excessive spending on FIFA's 2014 World Cup. On June 17, the movement exploded in size and spread across the country. Rio de Janeiro, for example, hosted a demonstration that lasted seven hours and attracted nearly 300,000 activists (Watts, 2013). Two days later, the mayor and governor of São Paulo announced a reduction of 20 cents in fares. Still on June 21 more than one million people protested across the country. President Rousseff canceled her trip to Japan in order to address the movement on national television. On June 24, the government announced a proposal for congress to approve a referendum on widespread political reform.

Our data collection approach followed a multi-stage process. We began by using Twitter's search API to sample messages embedding three hashtags associated with the movement (#changebrazil, #changebrasil, #mudabrasil "change Brazil" in Portuguese). Our tracking of data started on the day before street protests propagated across the country (June 16) and ended on the day after the largest demonstration occurred (June 21). In total, we collected 710,697 messages from 252,543 unique actors in a period of six days. Next, we applied the funnel process (Salge & Karahanna, 2018) and identified a total of eight bots.[60] We then returned

[60] See Appendix A for additional details on the approach.

to Twitter to collect messages that these bots generated (tweets, retweets, or replies) and appeared in (mentions) during the movement (June 16 to June 21). For every post, we have the username of the initiator, the message's content, URL, and timestamp. Our data records 17,399 messages from 158 actors. Over 93% of these messages (16,271) were initiated by the eight bots.

To examine the thematic content of bot generated messages during Brazil's movement, we used structural topic modeling (STM), which builds off of the tradition of probabilistic topics models such as latent Dirichlet allocation (LDA) (Blei, Ng, & Jordan, 2003). Like LDA, STM uses a generative process for discovering abstract topics occurring in a collection of documents[61] (Roberts, Stewart, & Airoldi, 2016). The major advantage of STM over LDA is that it enables the incorporation of covariates into the topic model, allowing researchers to examine valuable information about documents into the inference procedure. In other words, by using STM we can (1) discover the structure of topics in bot-generated messages; (2) explore how different covariates, such as bot and day, relate to these topics—e.g., *Toni_JoiaRara*[62] and *TeamRevoltNow* are more likely to share messages about social movements when compared to *Instant_RT* and *WorldCupRetweet*; and (3) examine how topics vary in different ways under different covariates—e.g., while *Toni_JoiaRara* and *TeamRevoltNow* are more likely to share messages about social movements, *Toni_JoiaRara* is more likely to share messages in Portuguese while *TeamRevoltNow* is more likely to share messages in English. Although LDA is able to accomplish the first, the model is not flexible enough to account for the second and third. In addition to the inclusion of covariates, STM allows scholars to explicitly estimate correlation among topics. This is useful since correlations can provide insights into when two topics are

[61] See Appendix B for a detailed explanation of LDA and STM, including their major differences.
[62] This and other names mentioned here reflect the Twitter handles of some bots in our sample.

likely to co-occur in the same document (i.e., a positive correlation).[63] Since bots can be coded to disseminate different types of content and their activity can exhibit temporal patterns, we added bot and day as covariates to our model and used the Spectral method for initialization, as recommended by Roberts, Stewart, and Tingley (2016).

Before employing STM to the collection of bot documents, we underwent an extensive data preprocessing phase. Our initial focus was identifying potential data quality problems more broadly such as misspelling errors originating from the data collection process. Next, we analyzed the data at the document level by cleaning, constructing, and formatting it. In doing so, we removed duplicates, punctuations, digits, whitespace, accents, and stop-words in English and Portuguese. We also categorized documents by language, converted all words into lower case, and replaced them with their root form. Our final sample includes a total of 12,604 documents, where each document corresponds to a message shared by a bot in a specific day. Finally, we varied the total number of topics in the entire collection of documents by running a total of 32 models which had 10 to 40 number of topics in addition to a first model with 50 topics.

Because our objective is to generate a model interpretable to humans, we first manually evaluated, in an iterative manner, the semantic qualities of our results (Debortoli, Müller, Junglas, & Brocke, 2016). We specifically assessed, for each and every model, how meaningful, interpretable, and coherent each individual topic was together with its most probable terms and document assignments. Next we statistically compared our models using topic coherence (Mimno, Wallach, Talley, Leenders, & McCallum, 2011) and the held out likelihood test (Walach, Murray, Salakhutdinov, & Mimno, 2009). Finally, to check the robustness of our

[63] While LDA does not provide topic correlations, the correlated topic model does (see Blei & Lafferty, 2007).

results, we used the algorithm of Lee and Mimno (2014) which automatically selects the number of topics[64]. Applying these criteria, we estimated a model with 21 topics[65] (see Figure 7).

To learn more about the structure of the corpus, we computed correlations between topics and plotted positive associations in a network graph (see Figure 8). This was useful because it allowed us to identify consistently shared patterns of topics among bot messages. We also detected topic communities using the fast algorithm of Clauset-Newman-Moore algorithm, which optimizes modularity[66] by using a greedy algorithm and does not require the specification of the desired number of clusters *a priori* (Newman, 2006). Topics that were more similar with respect to their most probable terms were grouped together, while topics which were dissimilar in their most probable terms were put in different groups. Finally, we iterated between these findings and our reading of specific messages to better understand the relationship among topics in different clusters and also between topics within the same community.

Altogether, these analyses allowed us to extensively study the thematic content of bot generated messages during Brazil's movement on Twitter. They enabled us to understand *what* topics bots share on social media and *how* topic relations and content characteristics can be leveraged for resource mobilization. We now focus our attention on each finding, providing its evidence and logic.

---

[64] See Appendix B for more details on these analyses.

[65] This is not to say that 21 is the "right" number of topics in this corpus—instead, we find that a 21-topic model provides useful insights about the structure of the texts shared by bots during Brazil's movement.

[66] "Modularity is, up to a multiplicative constant, the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random" (Newman, 2006:8578). Networks with high modularity have dense ties between nodes within clusters but sparse connections between nodes in different clusters.

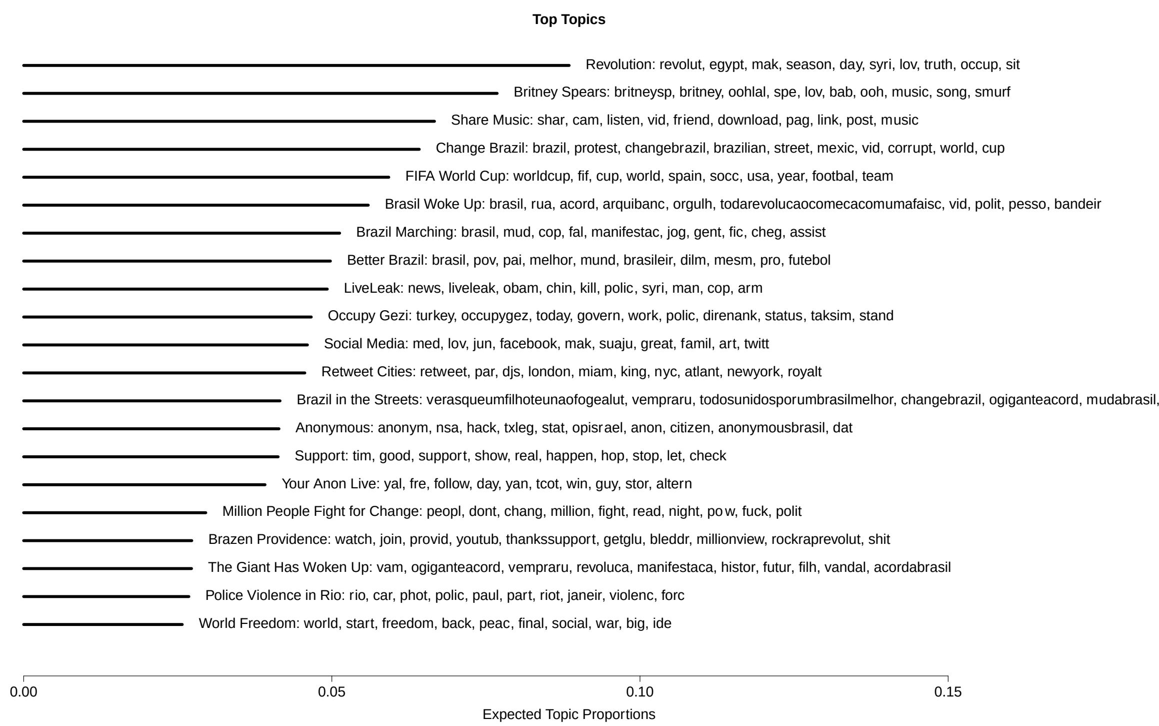

Figure 7. Structural topic model results from 12,604 documents, including the top 10 terms associated with each. Expected topic proportions indicate the proportion of the corpus that belongs to each topic.

Findings

What Topics Do Bots Discuss in a Social Movement on Twitter?

Using STM to estimate the structure of topics, we find that focal bots share, in greatest proportion, messages about political revolutions worldwide (see Figure 7). A number of specific social movement topics are also clearly identifiable, including topics about the movement in Brazil and the Gezi Park protests in Turkey. After reading hundreds of these documents, we find that social movement topics often provide live coverage of riots, expose police violence, show solidarity with activists, try to mobilize action, communicate the demands of the movement, and also legitimize the cause—i.e., explain the reasons why a set of people are pursuing social change. These messages also have hashtags and embedded links to external videos, photos, and articles which, in general, provide additional information about the movement.

RT @EcBillofRights #Revolution has come to Brazil! #BrazilUprising2013 #RiodeJaneiro #Anonymous #OWS http://t.co/acMcvzwa15 via @youtube

RT #OccupyLjubljana Turkish police running out of pepper spray canisters - reports http://t.co/G9gwYK4X3u #occupygezi #occupyturkey #Turkey #revolution

RT @owsscreenprint we stand in #solidarity with everyone worldwide who has decided to RISE UP #revolution #ChangeBrazil #direnGezi #Occupy

RT @CCKKI Carla #Dauden: No, I'm not going to the #WorldCup #OccupyBrazil #ChangeBrazil http://t.co/UTRQINaFsk via @youtube

RT @thalitalsb justin help brazil, tweet \"#changebrazil\" we need a revolution in our country, you can help all of us!!! love u ♥ @justinbieber
RT @viviana_a_r That's part of what we want!!! http://t.co/jVJg9rjWZU #ChangeBrazil #Vemprarua #VerasQueUmFilhoTeuNaoFogeALuta

RT @UnToldCarlisle Photo: Tear gas is fired towards protesters near the Arena Fonte Nova Stadium in Salvador #changebrazil #yal http://t.co/QMT9tkSqiG

RT @Tkaranasios1yr before #WorldCup, #Brazil erupts in protests. Why? Watch new playlist by @ythumanrights http://t.co/JjYGoSAx8J #ProtestoBR #changebrazil

RT @rapha_schneider #see #you #tomorrow #go #to #street #revolution #c'mom @Faculdade Pitagoras - Campus Divinopolis http://t.co/cRljXyUMUg

RT @a7med_afandi Espark-waiting in front of police batons. Be Careful!.#Revolution #direnankara #occupygezi #duranadam #direngeziparki http://t.co/yeFnm48OLA

Although collective action does not require the presence of SMOs (Oliver, 1989), prior work suggests that a significant part of social-change-oriented collective action in Western industrial democracies is "directly fielded by SMOs or proceeds under their auspices" (Edwards & McCarthy, 2004:136). While the importance of SMOs for mobilizing social movement resources is widely acknowledged by the literature, our results show that they are not part of the bot conversation during Brazil's movement. Instead, bots report on the patronage of Anonymous—a hacktivist network known for its distributed denial of service (DDoS) attacks on governments, religion, and corporate websites. Like SMOs, Anons contribute a variety of resources to social movements. For example, the network used social media to organize and mobilize one of the very first calls to Occupy Wall Street (Coleman, 2014). To support Brazil's cause, Anons volunteered their labor and attacked two major targets of the movement—the Brazilian government and FIFA—by hacking many of their websites.

> RT @MaceloLop3s #Anonymous broke into the website of the canopy in Cuiaba, and placed images that the TV did not show. http://t.co/0qdUYclE73 #ChangeBrazil
>
> RT @YourAnonCentral Hackers replace Brazil World Cup 2014 website with protest footage #changeBrazil #Anonymous http://t.co/8JNrQ9r8ly
>
> RT @khairaAlshater #OpBrazil By #Anonymous Gov Sites :- http://t.co/r0AGymNPNu... http://t.co/7eDc8GDwoq

We also find that Anonymous broadcasted a variety of other issues during the week of Brazil's movement, including a rally against fuel hikes in Indonesia, news updates about the Pirate Bay and Snowden investigations, the Taliban's release of a US soldier, and president Obama's lack of actions towards Guantanamo. Finally, the network encouraged others to sign an online petition to free Jeremy Hammond, a fellow Anon who went to prison in 2012 for leaking information from a private intelligence firm.

> RT @rkholil VIDEO violence by riot-police towards students #fuelhike rally in front of #indonesia parliament http://t.co/uBeBnoFg94 #YAL @YourAnonNews
>
> RT @YourAnonLive #BREAKING: Pirate Bay cofounder sentenced to 2 years in prison for hacking http://t.co/vuf7zyTGkY #tpb #YAL
>
> RT @UnToldCarlisle .@ggreenwald about to address charges against #Snowden on CNN. #IStandWithEdwardSnowden #yal
>
> RT @YourAnonLive #BREAKING: Taliban will free US soldier if 5 'operatives' released from Gitmo - @AP http://t.co/Pz6ciso9vq #YAL
>
> RT @07_anonymous it's been a month since Obama said he was going to close Guantanamo, he hasn't done sh*t about it" Debra Sweet #J19 #Snowden #Manning #Yal
>
> RT @OccupyDenver Jeremy has spent 15 months in prison, including weeks in solitary. Please sign this petition to free #Hammond. http://t.co/PoA1znBTcb #yal

Research indicates that celebrity support is germane for acquiring moral resources (King, 2011). Celebrities increase the salience and legitimacy of a cause by giving the movement a much higher profile than it would have otherwise. While we do not find data evidence of celebrities endorsing Brazil's cause, we discover that focal bots of Brazil's movement are not necessarily associated with social activism but instead connected to celebrities. The *Britney Spears* topic is a telling example. Its messages involved, for the most part, the release of "OhLaLa", a single that Spears recorded for the Smurfs 2 movie.

> RT @servullomoreira 'So baby come with me and be my ooh la la' - #BritneySpears acaba de lancar o audio oficial da sua nova musica... http://t.co/SJ2yZxZwWO
>
> RT @domecullen #BritneySpears #OohLaLa on #SPOTIFY , Ooh La La - Britney Spears http://t.co/kewo0VYJVE #NowPlaying
>
> RT @thepopmusiclife #britneyspears #smurfs2 #pop #musicnews Britney Spears releases Ooh La La from Smurfs 2 soundtrack http://t.co/gOJIvUaPpR

According to social movements theory, one of the ways in which activists mobilize social movement resources is by harming the image of target organizations and consequently, damaging their reputation (King, 2008). In line with this argument, we find that bots denigrated the Fédération Internationale de Football Association, also known as FIFA. They accused FIFA of corruption, asked the organization to "go home", and boycotted the 2014 World Cup.

> RT @michaelcdeibert Where does a corrupt operator like #FIFA's Sepp #Blatter get off lecturing #Brasil on how to solve its problems? #changeBrazil #WorldCup
>
> RT @AlexDarlan #ChangeBrazil #FIFAfail #Protest #WorldCup #Soccer #Football #FIFAGoHome #Corruption http://t.co/VpavkVQY2d
>
> RT @kaufpost #Brazil unrest linked to huge #football #WorldCup spending. Lessons from #FIFA mishap in #SouthAfrica cup. http://t.co/t3DcX21u4N
>
> RT @CCKKI Carla #Dauden: No, I'm not going to the #WorldCup #OccupyBrazil #ChangeBrazil http://t.co/UTRQINaFsk via @youtube

Overall, the bots we identified in Brazil's movement shared messages about topics related to political revolutions worldwide alongside topics which are specific to particular social causes, such as the ones in Brazil and Turkey. In doing so, they provided live coverage of street protests, revealed police brutality, demonstrated solidarity with the cause, attempted to mobilize action, expressed movement demands, and justified the reasons of the cause. We also show that bots reported on the activities of Anonymous, and specifically disclosed how the network mobilized resources for Brazil's movement. Finally, we discover that bots promoted celebrities who do not necessarily endorse the movement, harmed FIFA's image and also boycotted the 2014 World Cup. Such findings are consistent with prior literature in dialogue and tactics (e.g., Guo & Saxton, 2014; Theocharis et al., 2015; Penney & Dadas, 2014)—i.e., the content of messages disseminated by bots reflect one-way interactions aimed to broadcast information about movements, mobilize participation in street protests, and demonstrate symbolic support for social causes. Yet unlike existing research, we illustrate that bot generated content is not focused in one specific movement. Rather, they also reach other, more global social causes.

How Do Bot Topics Correlate With One Another to Discuss a Social Cause?

As illustrated in Figure 8, we find a total of four different clusters of topics. Topics in the first cluster are highly correlated with each other, which means that they are more likely to occur together in the same messages. These documents are written, for the most part, in Portuguese and are also specific to Brazil's movement. Altogether, they interact to demonstrate solidarity and sympathetic support for the cause by expressing a desire for a better nation, inviting people to join street protests, and signaling that Brazil is awoke and making history. Below we illustrate examples of correlated messages from three topics in this cluster (*Better Brazil*, *The Giant Has Woken Up*, and *Brazil in the Streets*). Bolded words are the most probable terms in each topic.

RT @cw_victoria #foraDilma **#vemprarua** #gritasemtermedobrasil #verasqueumfilhoteunaofogealuta **#ogiganteacordou** nos top tags, ai sim! Orgulho desse **Brasil**!

#TodosUnidosPorUmBrasilMelhor, #OPovoCansou**Dilma**, #VemPraRua, #changebrazil, **#AcordaBrasil** #oGiganteAcordou http://t.co/N9mDipHVXD

RT @euandrec A tarifa CAIU mas o **povo** nao DESISTIU! e o Brasil rumo a Evolucao! **#verasqueumfilhoteunaofogealuta #AcordaBrasil** #GritaSemTerMedoBrasil

RT @EddieNovaes633 O **mundo** esta com a gente! #AcordaBrasil **#VerasQueUmFilhoTeuNaoFogeALuta #OGiganteAcordou** #ItsNotAbout20Cents http://t.co/KDkzttgMW5

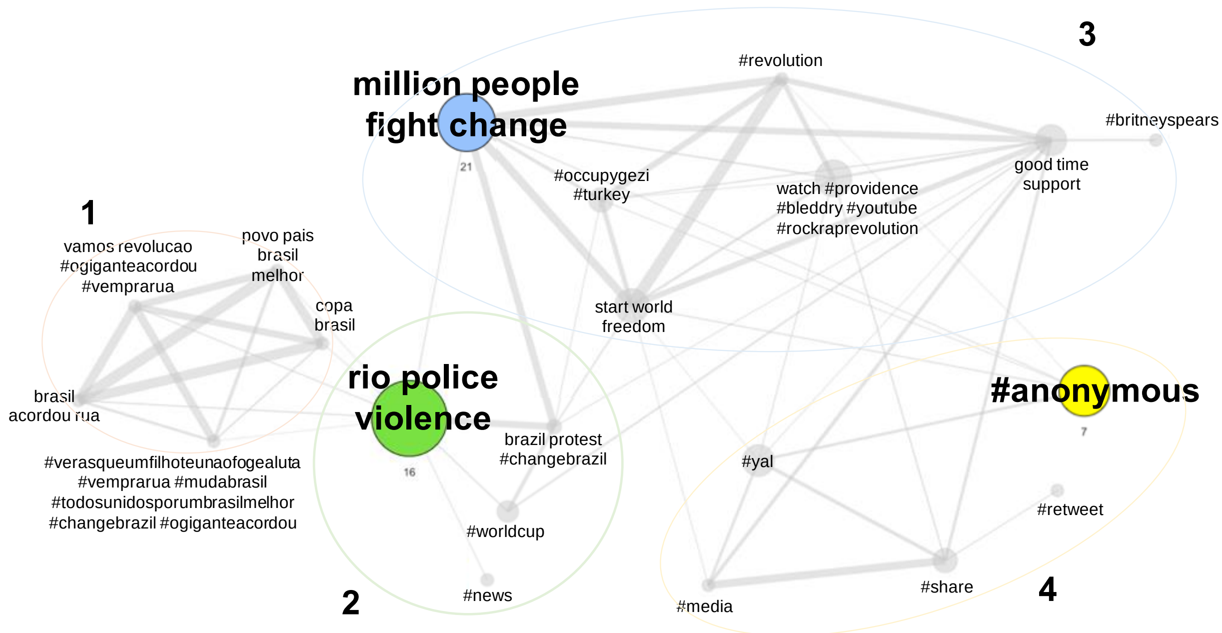

Figure 8. Network of positive correlations between the topics illustrated in Figure 7. Topics near each other, and with a tie, indicate that they are more likely to occur in the same document. Colored ellipses illustrate topic communities detected with the Clauset-Newman-Moore algorithm. Node sizes correspond to betweenness centrality and are graphed using the Fruchterman-Reingold algorithm. Node labels represent the most probable keywords or hashtags in a topic.

While topics in cluster 2 are not as highly correlated as topics in cluster 1, they also interact with each other to support Brazil's cause. For example, the association between *Police Violence in Rio*, *Change Brazil*, and *LiveLeaks* exists because **rio** and **police** (high probability terms in *Police Violence in Rio*) are jointly used in the same messages with **#ChangeBrazil** (a high probability term in *Change Brazil*) and **#news** (a high probability term in *LiveLeaks*) to denounce police violence against activists in Rio de Janeiro. We also find that **#ChangeBrazil** is jointly used in the same messages with **#WorldCup** (a high probability term in *World Cup*) to blame the government and FIFA of corruption and thus legitimize the movement.

> RT @OpManning **Police** shoot at protesters in #**Rio** #Brazil **#ChangeBrazil** #OccupyBrazil #anonymous Video http://t.co/JrQwChDvF5 http://t.co/PnJOQe2L9s
>
> RT @khairaAlshater **Police** fire rubber bullets at a protester in **Rio**. 6/20 #OccupyBrazil **#ChangeBrazil** #Brazil #Revolution... http://t.co/iwytsauu4d
>
> RT @Tkaranasios 1yr before **#WorldCup**, #Brazil erupts in protests. Why? Watch new playlist by @ythumanrights http://t.co/JjYGoSAx8J #ProtestoBR **#changebrazil**
>
> RT @brunoborges This video explains why a million Brazilians protested last night #Brazil http://t.co/U7dcoBjZQm **#ChangeBrazil #WorldCup** #ConfederationsCup
>
> Brazil UPRISING **- Rio Violent** Riot http://t.co/VQ1ekoYXt0 LiveLeak **#News**
>
> Brazil Protests **- Rio** - **Police** brutality http://t.co/cEbKYEHiQC LiveLeak **#News**

We observe a few important differences between the first two clusters. One is related to language. While documents in cluster 1 are mostly in Portuguese the ones in cluster 2 are mostly in English. The second observation relies on the characteristics of topic documents. To support the movement, the first cluster collectively uses terms which are *positive*, based on hope and promise of the future. A few examples are *todosunidosporumbrasilmelhor* (everyone united for a better Brazil), *orgulho* (pride), *vemprarua* (come to the streets), and *ogiganteacordou* (the giant has woken up). Its messages corroborate this as they do altogether express a desire for a *better* nation and signal that Brazil is *awoke, marching*, and *making history*. Yet this is not the dialogue we find in the second cluster, where most topic terms are *negative*, based on anger and shame. Examples include *corrupt*, *violence*, *kill*, and *police*. Our reading of the documents confirms this as topics in this cluster interact with each other to condemn police brutality, accuse FIFA of corruption, and boycott the 2104 World Cup.

Collectively, these two clusters are specific to Brazil's social movement. We term bots' use of social media to support a social movement as *Focal Botivism*, which we define as scripted actions by which a group of bots share positive and negative messages in correlated topics which altogether promote social change in a specific population and for a specific cause. Thus, *Focal Botivism* disentangles activist bots from opponent and bystanders bots. Central to *Focal Botivism* is that topic associations explicitly avoid the promotion of other additional social causes. This construct is also not about topics outside the realm of collective action. Rather, *Focal Botivism's* documents focus on bot topic interactions that altogether provide general support about common social movement issues with an aim of helping to implement the goals of a specific cause.

Correlations between topics in the third cluster are stronger than those in the second but also weaker than the ones in the first. *Revolution* is an especially insightful topic here since it correlates highly with almost every other topic. By reading many different documents, we observe that these associations arise from the co-occurrence of topics most probable terms. **#Revolution** (high probability term in *Revolution*), for example, frequently appears in the same documents with **#freedom** (high probability term in *World Freedom*), **million** and **people** (high probability terms in *Million People Fight for Change*), **#occupygezi** (high probability term in *Occupy Gezi*), and **#providence** (high probability term in *Brazen Providence*).

> RT @Wyte_Bread #OccupyThailand 16.6.2013 **#Revolution** #OccupyAllStreets #OccupyTheWorld **#Freedom** http://t.co/3OntsGnII9

> RT @SaladdinAhme What is the alternative? To struggle for a world in which alternatives are possible. **#Revolution** **#Freedom** #Occupy #Anarchism #Marxism

> RT @strongerhero @ddlovato BRAZIL **#REVOLUTION** one **million people** in the streets fighting for what is promised but not fulfilled! http://t.co/gdinSHp9jW

> Over 1 **million people** on the streets of Brazil last night, in over 100 cities. **#Revolution** #Brazil #ChangeBrazil
>
> RT @70torinoman Turkish police surrounded and closed Taksim Square.#direngeziparki #direnankara #wearegez **#Revolution #occupygezi** http://t.co/YmLdoG3H9o
>
> RT @a7med_afandi Ethem Ankara is calling the resistance.#direngeziparki #direnankara #wearegez **#Revolution #occupygezi** #Taksim #Turkey http://t.co/DzBGESOUD5
>
> RT @BraznProvidence **#PROVIDENCE** is #1 on ReverbNation Alternative charts for The WORLD #CheckItOut http://t.co/RJWgUUAJXZ #RockRap **#Revolution** #Thanks4Support
>
> RT @BraznProvidence 5.4 #MillionViews Join the **#Revolution** Watch **#Providence** - #BledDry on #YouTube #RockRapRevolution, Thanks4Support! http://t.co/ct8xcpygfk

Overall, these topics relate with one another to support political revolutions started by millions of people around the world, including the Gezi Park protests and the Occupy Movement. *Support* and *Britney Spears* are the remaining topics in cluster 3. The former strongly correlates with a variety of other topics in the cluster, such as *Revolution* and *World Freedom*, yet the latter is an isolate—as shown in Figure 8, *Britney Spears* only connects to *Support*.

While topics in cluster 3 also have terms and discourses which are *positive* (freedom and peace) and *negative* (police) they are not restricted to a social movement and so they cannot be part of *Focal Botivism*. Instead, we call this broader space of bot topics as *Global Botivism* and define it as scripted actions by which a group of bots share positive and negative messages in correlated topics which altogether promote social change across the globe. Since many of the topics in *Global Botivism* are general and somewhat abstract (e.g., *Support*) they can easily correlate with topics outside the realm of collective action, such as *Britney Spears*.

> RT @MickeyRoo my video in **support** of #buyoohlalaonitunes #buyoohlala **#britneyspears** @britneyspears @grl @hannahspears http://t.co/MOiQlgS1KF
>
> RT @michaelmor09 You know I'm gonna **support** my girl! **#britneyspears** #oohlala #smurfs #newsingle http://t.co/Gfzc6VjNKE

The fourth cluster also belongs to *Global Botivism*. Here, topics are weakly correlated with each other and revolve around *Anonymous*. We find *Your Anon Live* to be especially insightful. Like *Revolution* in cluster 3, this topic correlates with almost every other topic in the cluster.

> RT @DiffusedDesign To see the value of social **media**, watch what happened in Turkey when the local media failed http://t.co/4Csl2SzsUm **#YAL**
>
> RT @data_overflow @57UN Brazilian police officer threw his weapon and joined the protesters http://t.co/M3eY3LurbQ please **share** it **#YAN** :)
>
> RT @RedAnneBolynn TY! .mt@KitOConnell:#HB60 & related bills up final vote Sun, I call my **#Anonymous** sisters & brothers 4 support. EYES ON #Texas WOMEN, **#YAL**

As the messages above show, topics in the fourth cluster interact to collectively report on the activities of Anonymous or to request resources from the hacktivist network.

In summary, while we find that bots share messages in a variety of topics during a social movement, we also show that these topics correlate with one another to collectively promote social change in a specific population and for a specific cause (*Focal Botivism*) and across the globe (*Global Botivism*). Differences between *Focal Botivism* exist for language. Activist bots which are focal to a movement disseminate messages in multiple idioms. In addition, we find that *Global Botivism* is mostly in English, reaches other political revolutions in the world together with a hacktivist network which is supportive of specific social movements. The abstract nature of *Global Botivism* also enables correlations with topics outside the realm of collective action. Finally, we discover that topics in a cluster relate with one another via term co-occurrences—i.e., by appearing together in the same messages. Table 10 summarizes our findings.

Table 10. Major Findings

Research Question	Analysis	Empirical Findings
What topics do bots discuss in a social	Structural Topic Modeling	See Figure 7

movement on Twitter?		
How do these topics relate to one another?	Network of Topic Correlations & Topic Community Detection	To collectively promote social change in a specific population and for a specific cause (***Focal Botivism***) and across the globe (***Global Botivism***)

We now turn our attention to the synergies between *Botivism* and resource mobilization. Our goal is to theoretically explicate how *Focal Botivism* and *Global Botivism* can be leveraged to mobilize resources for social movements on social media.

How Does Bot Generated Content Mobilize Resources for a Social Cause?

Since the accumulation of a variety of resources is crucial for social movements' success (McCarthy & Zald, 1977), we abstract our empirical findings to a theoretical framework of *Botivism* (see Figure 9). The model offers two paths by which bots help acquire resources for a social movement on social media. The first relies on *Focal Botivism*. Here, bots leverage moral battery to generate indignation and thus attract moral resources toward a specific social cause. They do so by illustrating a social problem, expressing faith that it can be fixed, and eliciting sufficient indignation to mobilize action. Positive messages, on one hand, likely instigate feelings of hope and admiration since they illustrate promise of the future—"The fare WENT DOWN but the people DID NOT GIVE UP! This is Brasil towards evolution! #verasqueumfilhoteunaofogealuta #AcordaBrasil #GritaSemTerMedoBrasil" reads an example. Negative messages, on the other hand, likely get people's attention by provoking an emotional state of shock—"police shoot at protesters in #Rio #Brazil #ChangeBrazil #OccupyBrazil #anonymous Video http://t.co/JrQwChDvF5 http://t.co/PnJOQe2L9s" represents an example. While moral resources are often bestowed by external sources known to possess them (Edwards & McCarthy, 2004), with *Focal Botivism* social movements create moral resources on their own.

The second path relies on *Global Botivism. Botivism* in this form ensures that a group of bots share positive and negative messages in correlated topics which altogether promote social change across the globe. Like *Focal Botivism*, this mechanism increases moral resources via moral battery. The level to which this is accomplished is more abstract though since the stimulus is broader and thus divided across multiple causes worldwide. But because *Global Botivism* foregrounds aggregation (Edwards & McCarthy, 2004) it provides access to many diverse actors and thereby the resources embedded in them can be leveraged. For example, activists of separate social movements, when aggregated together in a social network, may bond over their grievances and therefore exchange solidarity and sympathetic support, increasing moral resources for their respective causes. Example messages include "@DilekKanatli Now we are one... **#ChangeBrazil** #PrayForBrasil #Brazil #Turkey **#OccupyGezi #revolution** #protest http://t.co/6V05X6pee5" and "@Elephantju: We support you #Brazil never give up #occupybrazil **#occupygezi #ChangeBrazil #Revolution** http://t.co/arRilTbVSH." It is important to note that such increase of moral resources is indirect, via social organizational resources, and therefore distinct from the direct effect we propose via moral battery.

Finally, *Global Botivism* raises the amount of human resources provided to a focal movement via patronage. By supporting social causes across the globe, *Global Botivism* reaches networks of patrons which may identify their goals with those of a focal movement and therefore make their labor accessible and usable through participation. An example is Anonymous, which provides expertise and technical assistance by exposing security vulnerabilities of target organizations and institutions of social movements that they sympathize with—"Hackers replace Brazil World Cup 2014 website with protest footage #changeBrazil #Anonymous http://t.co/8JNrQ9r8ly."

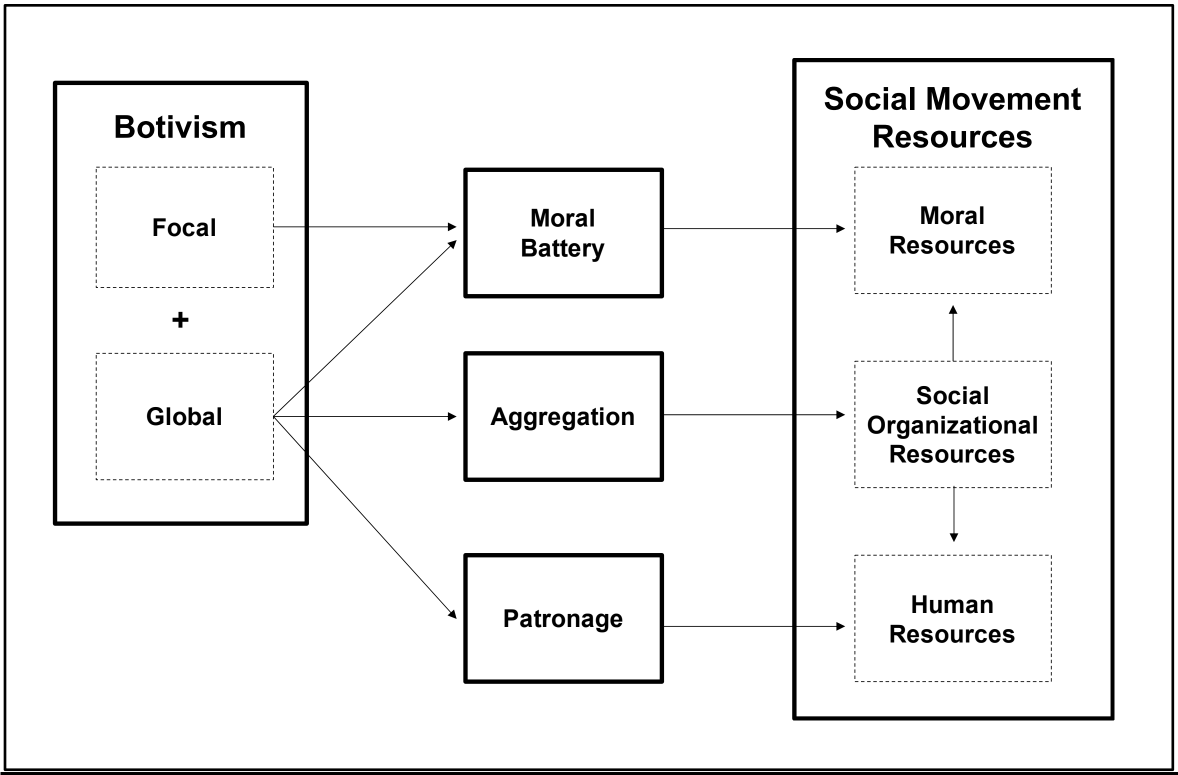

Figure 9. Botivism for Resource Mobilization

Contributions To Theory & Practice

Our study provides three contributions to IS research in social movements. First, we identify two forms of social activism enabled by bots, namely *Focal Botivism* and *Global Botivism*. Prior research has suggested that bots participate and engage in social activism (Salge & Karahanna, 2018; Forelle, Howard, Monroy-Hernández, & Savaga, 2015; Savage, Monroy-Hernández, & Hollerer, 2016). Our central insight is that they do so by promoting political revolutions worldwide alongside with specific social movements.

Second, we offer a theoretical framework which combines the two forms of *Botivism* we identify to explain how activist bots generate resources for social movements on social media. The first mechanism builds on *Focal Botivism* and leverages moral battery to attract resources. Here, bots advantageously share positive and negative messages which collective support social change in a specific population to generate indignation and thus attract moral resources toward a social cause. The second and last mechanism of our model is *Global Botivism*, which builds on aggregation and patronage to mobilize social organization and human resources together with moral resources which are acquired indirectly via aggregation. Like *Focal Botivism*, *Global Botivism* also mobilizes moral resources directly, via moral battery. Overall, our *Botivism* framework indicates that activist bots sharing positive and negative messages in correlated topics which altogether promote social change in a population and across the globe mobilize moral, social organizational, and human resources for social movements via moral battery, social network aggregation, and patronage. It is important to note that *Botivism* is different than the social activism enacted by humans in two ways. First, the marginal cost incurred with *Botivism* is nearly zero since the sharing of content is automated. Second, the scale of *Botivism* is much larger since bots share content at a much higher volume and speed than any human possibly can.

Third, by using moral battery to explicate the acquisition of resources facilitated by digital technologies, we reconcile two streams of research in social movements, namely emotions (Jasper, 2014) and resource mobilization (McCarthy & Zald, 1977). While the first focuses on clarifying the several conceptual confusions associated with emotion and its role in social change (see Jasper, 2011 for a review), the second pays scant attention to emotions, arguing that "social movements may or may not be based upon grievances of the presumed beneficiaries" (McCarthy & Zald, 1977:1216) and instead focuses on how SMOs are structured and organized to acquire resources in order to work toward implementing the goals of a cause.

Fourth, our study contributes to a broader nascent research program examining the synergies between technology and social change (e.g., Vaast et al., 2017; Miranda et al., 2016). While existing research focuses on human activists exchanging content in social media networks, our work focuses on activist bots. Bots provide volume and scalability to actions at a low cost and can thus offer new action repertoires to augment extant research.

Finally, like all research, our work has limitations too. We use a topic modeling approach to examine the thematic content of bot-generated messages during a social movement on Twitter. Our emerging framework is therefore specific to Twitter and may not be applicable to other social media platforms with different social structures, such as Facebook. Nonetheless, our work represents a first step in addressing the content bots discuss through the lens of topics during times of social change. This is important since there is plenty of anecdotal and scientific evidence supporting the idea that these bots are prominent and potentially impactful (see Ferrara et al., 2016 for an example).

Conclusion

Based on the exploration of rich textual social media data, our emergent theoretical framework argues that activist bots mobilize resources for social movements. These resources are of various types and collectively assist in implementing the goals of cause. Our model is especially relevant for grass-root movements facing media censorship and government oppression and lacking external sponsorship, like the cases of Brazil and Turkey. The next step is careful examinations of our framework with larger case samples, in other social media platforms, and across different social movements.

References

Abokhodair, N., Yoo, D., & McDonald, D. W. 2015. Dissecting a social botnet: Growth, content, and influence in Twitter. **In** ***Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing***. ACM.

Agarwal, S. D., Bennett, L. W., Johnson, C. N., & Walker, S. 2014. A Model of Crowd-Enabled Organization: Theory and Methods for Understanding the Role of Twitter in the Occupy Protests. ***Internationa Journal of Communication***, 646–672.

Ahmed, S., Jaidka, K., & Cho, J. 2017. Tweeting India's Nirbhaya protest: A study of emotional dynamics in an online social movement, ***Social Movement Studies***, 16:4, 447-465, DOI: 10.1080/14742837.2016.1192457

Aitchison, J., & Shen, S. M. 1980. Logistic-normal distributions: Some properties and uses, ***Biometrika,*** 67:2, 261–272.

Atkins, D. C., Rubin, T. N., Steyvers, M., Doeden, M. A., Baucom, B. R., & Christensen, A. 2012. Topic models: A novel method for modeling couple and family text data. ***J. Family Psych***. 26(5):816–827.

Auger, G. A. 2013. Fostering democracy through social media: Evaluating diametrically opposed nonprofit advocacy organizations' use of Facebook, Twitter, and YouTube. ***Public Relations Review***, *39*, 369-376. doi:10.1016/j. pubrev.2013.07.013

Bao, Y., Datta, A. 2014. Simultaneously discovering and quantifying risk types from textual risk disclosures. ***Management Sci***. 60(6):1371–1391.

Barker- Plummer, B. 2002. Producing public voice: resource mobilization and media access in the national organization for women. ***Journalism and Mass Communication Quarterly***, **79**(1), 188–205.

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

Becker, G. 1964. ***Human Capital***, New York: National Bureau of Economic Research/ Columbia University Press.

Benford, R. D., and Snow, D. A. 2000. "Framing Processes and Social Movements: An Overview and Assessment," ***Annual Review of Sociology***, (26), pp. 611–39.

Blei, D., Ng, A. Y., & Jordan, M. I. 2003. Latent Dirichlet Allocation. ***Journal of Machine Learning Research***, *3 993–1022*.

Blei, D., & Lafferty, J. D. 2007. A correlated topic model of science. ***The Annals of Applied Statistics***, *1 17–35*.

Blei, D. M. 2012. Probabilistic topic models. ***Communications of the ACM,*** 55: 77–84.

Bortree, D. S., & Seltzer, T. 2009. Dialogic strategies and outcomes: An analysis of environmental advocacy groups' Facebook profiles. ***Public Relations Review***, *35*, 317-319. doi:10.1016/j.pubrev.2009.05.002

Bourdieu, P. 1986. **The forms of capital** in J. Richardson (ed.), *Handbook of Theory and Research for the Sociology of Education*, New York: Greenwood Press, pp. 241–58.

Chavoshi, N., Hamooni, H., & Mueen, A. 2016. DeBot: Twitter bot detection via warped correlation. In ***Proceedings of the IEEE International Conference on Data Mining***.

Coleman, J. S. 1988. Social capital in the creation of human capital. ***American Journal of Sociology*** (Supplement), **94**, S95–S120.

Coleman. G. 2014. ***Hacker, hoaxer, whistleblower, spy: The many faces of Anonymous***. New York, NY: Verso.

Cresci, S., Di Pietro, R., Pretocchi, M., Spognardi, A., & Tesconi, M. 2015. Fame for sale: Efficient detection of fake Twitter followers. ***Decision Support Systems***, *80: 56–71, doi: 10.1016/j.dss.2015.09.003*.

Cress, D.M. & D.A., Snow. 1996. Mobilization at the margins: resources, benefactors,and the viability of homeless social movement organizations. ***American SociologicalReview***, **61**(6), 1089–109.

Cress, Daniel M., & David A. Snow. 2000. The Outcomes of Homeless Mobilization: The Influence of Organization, Disruption, Political Mediation, and Framing. ***American Journal of Sociology*** 105(4): 1063-1104.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. 2016. BotOrNot: A system to evaluate social bots. ***Working Paper.*** 2 pages.

Debortoli, S., Muller, O., Junglas, I., vom Brocke, J. 2016. Text Mining For Information Systems Research- ers: An Annotated Topic Modeling Tutorial. ***Communications of the Association for Information Systems,*** Vol. 39, Article 7.

Earl, J., & Kimport, K. 2011. ***Digitally enabled social change: Activism in the Internet age***. Cambridge: MIT Press.

Edwards, B., & McCarthy, J. D. 2004. Resources and Social Movement Mobilization, **In *The Blackwell Companion to Social Movements*, D. A. Snow, S. A. Soule, and H. Kriesi (eds.)**, Northhampton, MA: EdwardElgar Publishing Inc., pp. 116–152.

Edwards, B., & Kane, M. 2014. Resource mobilization and social and political movements, **In *The Handbook of Political Citizenship and Social Movements*, van der Heijden (eds.)**, Malden, MA: Blackwell Publishing Ltd., pp. 205–232.

Einwohner, R. L. 2003. Opportunity, Honor, and Action in the Warsaw Ghetto Uprising of 1943. ***American Journal of Sociology*** 109(3): 650-75.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. 2016. The rise of social bots. ***Communications of the ACM.*** *59:7 96–104, doi:* http://dx.doi.org/10.1145/2872518.2889302.

Forelle, M., Howard, P., Monroy-Hernandez, A. & Savage, S. 2015. Political bots and the manipulation of public opinion in Venezuela. ***Essay.*** *1–8*.

Gamson, W. A. 1975. **The Strategy of Protest.** Homewood, Ill.: Dorsey.

Halfmann, D. & Michael P. Y. 2010. War Pictures: The Grotesque as a Mobilizing Tactic. ***Mobilization*** 15(1): 1-24.

Hinsley, A. & Lee, H. 2015. #Ferguson strategic messaging: How local journalists and activists used Twitter as a communication tool. ***International Syposium on Online Journalism,*** 6(1) https://isojjournal.wordpress.com/2015/04/15/ferguson-strategic-messaging-how-local-journalists-and-activists-used-twitter-as-a-communication-tool/.

Huang, A., Zang, A., & Zheng, R. 2014. Evidence on the information content of text in analyst reports. ***Accounting Rev.*** 89(6): 2151–2180.

Jackson, P. 1982. ***Tactics of confrontation***. In J. S. Nagelschmidt (Ed.), *The public affairs handbook* (pp. 211–220). New York, NY: American Management Association.

Jasper, J. M. 2011. Emotions and social movements: Twenty years of theory and research. ***Annual Review of Sociology,*** 37:14, 14.1—14.19.

Jasper, J. M. 2014. ***Protest: A cultural introduction to social movements***. New York: Polity.

Jahng, M. R., & Lee, N. 2018. When Scientists Tweet for Social Changes: Dialogic Communication and Collective Mobilization Strategies by Flint Water Study Scientists on Twitter. ***Science Communication***, *40*, 89-108.

Kent, M. L., & Taylor, M. 1998. Building dialogic relationships through the world wide web. ***Public Relations Review***, *24*, 321-334. doi:10.1016/S0363-8111(99)80143-X

King, B. G. 2008. A political mediation model of corporate response to social movement activism. ***Administrative Science Quarterly***, *53:3 395–421, doi: 10.2189/asqu.53.3.395*.

King, B. G. 2011. The tactical disruptiveness of social movements: Sources of market and mediated disruption in corporate boycotts. ***Social Problems***, *58*(4), 491-517.
DOI: 10.1525/sp.2011.58.4.491

Kline, R. B. 2011. ***Principles and practices of structural equation modeling***. New York: The Guilford Press.

Lee, M. & Mimno, D. 2014. Low-dimensional embeddings for interpretable anchor-based topic inference. ***In Proceedings of the Conference on Empirical Methods in Natural Language Processing.*** Doha, Qatar: Association for Computational Linguistics, 1319—1328.

Lee, N. M., VanDyke, M. S., & Cummins, R. G. 2017. A missed opportunity? NOAA's use of social media to communicate climate science. ***Environmental Communication***, 1-10. doi:10.1080/17524032.2016.1269825

LeFebvre, R. K., & Amrstrong, C. 2018. Grievance-based social movement mobilization in the #Ferguson Twitter storm. ***New Media & Society***, *20(1)*, 8-28.

Li, W., Yin, J., & Chen, H. 2017. Supervised topic modeling using hierarchical Dirichlet proceess-based inverge regression: Experiments on e-commerce applications. ***IEEE Transactions on Knowledge and Data Engineering***, *1041–4347*.

Lovejoy, K., & Saxton, G. D. 2012. Information, community, and action: How nonprofit organizations use social media. ***Journal of Computer Mediated Communication***, *17*, 337-353.

Lovejoy, K., Waters, R. D., & Saxton, G. D. 2012. Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. ***Public Relations Review***, *38*, 313-318. doi:10.1016/j.pubrev.2012.01.005

McCammon, H.J., K.E., Campbell, E.M., Granberg & C. Mowery. 2001. How movements win: gendered opportunity structures and U.S. women's suffrage movements, 1866–1919. ***American Sociological Review***, **66**(1), 49–70.

McCammon, H.J. & N. van Dyke (eds). 2010. ***Strategic Alliances: Coalition Building and Social Movements***, Minneapolis, MN: University of Minnesota Press.

McCarthy, J.D. 1996. Mobilizing structures: constraints and opportunities in adopting, adapting and inventing in D. McAdam, J.D. McCarthy and M.N. Zald (eds), ***Comparative Perspectives on Social Movements***, New York: Cambridge University Press, pp. 141–51.

McCarthy, J. D., and Zald, M. N. 1977. Resource mobilization and social movements: A partial theory. ***American Journal of Sociology,*** *82:6, 1212–1241, doi: 10.1086/226464*.

McCright, A. M., & Riley, E. D. 2000. Challenging Global Warming as a Social Problem: An Analysis of the Conservative Movement's Counter-Claims. ***Social Problems*** 47(4): 499-522.

Mimno, D., Wallach, H. M., Talley, E. Leenders, M., & McCallum, A. 2011. Optimizing semantic coherence in topic models. ***In Proceedings of the Conference on Empirical Methods in Natural Language Processing.*** Stroudsburg, PA: Association for Computational Linguistics, 262—272.

Miranda, S. M., Young, A., & Yetgin, E. 2016. Are social media emancipatory or hegemonic? Societal effects of mass media digitization in the case of the SOPA discourse. ***MIS Quarterly***, *40:2 303–329*.

Monroy-Hernández, A., & Spiro, E. 2013. How Brazilian protesters are using Twitter. The full title of the work is part of http://www.theguardian.com/news/datablog/2013/jul/04/brazilian-protesters-twitter-microsoft. Retrieved 7/30/2015.

Morstatter, F., Wu, L., Tahora, N. H., Carley, K., & Liu, H. 2016. A new approach to bot detection: Striking the balance between precision and recall. ***ASONAM.***

Naaman, M., Becker, H., & Gravano, L. 2011. Hip and trendy: Characterizing emerging trends on Twitter. ***Journal of the American Society for Information Science and Technology,*** *62(5), 902–918.*

Newman, M. E. J. 2006. Modularity and community structure in networks. ***PNAS,*** *103 (23), 8577–8582.*

Obar, J. A., Zube, P., & Lampe, C. 2012. Advocacy 2.0: An analysis of how advocacy groups in the United States perceive and use social media as tools for facilitating civic engagement and collective action. ***Journal of Information Policy***, *2*, 1-15.

Oh, O., Eom, C. M., & Rao, H. R. 2015. Research note—Role of social media in social change: An analysis of collective sense making during the 2011 Egypt revolution. ***Information Systems Research,*** *26:1 210–223, doi: http://dx.doi.org/10.1287/isre.2015.0565*.

Oliver, P. E. 1989. Bring the Crowd Back in: The Nonorganizational Elements of Social Movements. ***Research in Social Movements, Conflict and Change***, 5, 133–70.

Oliveira, É. T. C., de França, F. O. D., Goya, D. H., & Penteado, C. L. de Camargo. 2016. The influence of retweeting robots during Brazilian protests. Paper presented at the 2016 49th ***Hawaii International Conference on System Sciences*** (HICSS).

Olzak, S. & E. Ryo. 2007. Organizational diversity, vitality and outcomes in the civil rights movement, ***Social Forces***, **85**(4), 1561–91.

Oselin, S. S., & Corrigall-Brown, C. 2010. A Battle for Authenticity: An Examination of the Constraints on Anti-Iraq War and Pro-Invasion Tactics. ***Mobilization*** 15(4): 511-33.

Papacharissi, Z., & de Fatima Oliveira, M. 2012. Affective news and networked publics: The rhythms of news storytelling on #Egypt. ***Journal of Communication***, 62(2), 266–282. doi:10.1111/j.1460–2466.2012.01630.x

Pedriana, N. 2006. From Protective to Equal Treatment: Legal Framing Processes and Transformation of the Women’s Movement in the 1960s. ***American Journal of Sociology*** 111(6): 1718-61.

Penney, J., & Dadas, C. 2014. (Re)tweeting in the service of protest: Digital composition and circulation in the Occupy Wall Street movement. ***New Media & Society***, 16(1), 74–90. doi:10.1177/1461444813479593.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., Radev, D. R. 2010 How to analyze political attention with minimal assumptions and costs. ***Amer. J. Political Sci.*** 54(1):209–228.

Rybalko, S., & Seltzer, T. 2010. Dialogic communication in 140 characters or less: How Fortune 500 companies engage stakeholders using Twitter. ***Public Relations Review***, *36*, 336-341.

Roberts, M. E., Stewart, B. M., & Airoldi, E. M. 2016. A model of text for experimentation in the social sciences. ***Journal of the American Statistical Association*** (111:515), 988–1003.

Roberts, M. E., Stewart, B. M., & Tingley, D. 2016. Stm: R Package for Structural Topic Models. ***Journal of Statistical Software,*** 1–42, doi: 10.18637/jss.v000.i00.

Salge, C. & Karahanna, E. 2018. Protesting corruption on Twitter: Is is it a bot or is it a person? *Forthcoming at* **Academy of Management Discoveries**.

Savage, S., Monroy-Hernandez, A., & Hollerer, T. 2016. Botivist: Calling volunteers to action using online bots. ***In Proceedings of the 19th ACM Computer-Supported Collaborative Work and Social Computing****, 813–822,* ACM.

Selander, L., & Jarvenpaa, S. L. 2016. "Digital Action Repertoires and Transforming a Social Movement Organization," ***MIS Quarterly*** (40:2), 331–352.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & and Menczer, F. 2017. The spread of misinformation by social bots. ***Technical Report*** 1707.07592, arXiv.

Shi, Z., Lee, G. M., and Whinston, A. B. 2016. "Toward a Better Measure of Business Proximity: Topic Modeling for Industry Intelligence," ***MIS Quarterly*** (40:4), pp. 1035-1056.

Snow, D. A., & Benford, R. D. 1988. Ideology, Frame Resonance, and Participant Mobilization. ***International Social Movement Research*** 1(1): 197-217.

Snow, D., Benford, R. D., McCammon, H. J., Hewitt, L., & Fitzgerald, S. 2014. The emergence, development, and future of the framing perspective: 25+ years since "frame alignment". ***Mobilization*** 19(1): 23-45.

Sommerfeldt, E. J. (2011). Activist online resource mobilization: Relationship building features that fulfill resource dependencies. ***Public Relations Review***, 37, 429-431.

Soule, S. A., McAdam, D., McCarthy, J. D., & Su, Y. 1999. Protest Events: Cause or Consequence of State Action? The U.S. Women's Movement and Federal Congressional Activities, 1956-1979. ***Mobilization*** 4: 239-56.

Suh, D. 2004. Outcome Framing and Movement Dynamics: Korean White-Collar Unions' Political Mobilization and Interunion Solidarity, 1987-1995. ***Mobilization*** 9(1): 17-37.

Suarez-Serrato, P., Roberts, M. E., Davis, C., & Menczer, F. 2016. On the influence of social bots in online protests. ***In Social Informatics, Lecture Notes in Computer Science***, 269–278. Springer.

Taylor, M., Kent, M. L., & White, W. J. 2001. How activist organizations are using the Internet to build relationships. ***Public Relations Review***, *27*, 263-284.

The Economist. 2017. In praise of serendipity. The full title of the work is part of https://www.economist.com/news/books-and-arts/21718464-social-media-should-encourage-chance-encounters-not-customised-experiences-praise. Retrieved 3/14/2018.

Theocharis, Y., Lowe, W., van Deth, J. W., García-Albacete, G. 2015. Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. ***Information, Communication, & Society***. 18(1), 202–220.

Tufekci, Z. 2017. ***Twitter and tear gas: The power and fragility of networked protest***. London, UK: Yale University Press.

Vaast, E., Safadi, H., Lapointe. L, & Negoita, B. 2017. Social media affordances for connective action: An examination of microblogging use during the Gulf of Mexico oil spill. ***MIS Quarterly***, *41:4 1179–1209*.

Vasi, I. B., Walker, E. T., Johnson, J. S., & Tan, H. F. 2015. "No Fracking Way!" Documentary Film, Discursive Opportunity, and Local Opposition against Hydraulic Fracturing in the United States, 2010 to 2013. ***American Sociological Review***, *80(5),* 934-959.

Verkamp, J. P. & Gupta, M. 2013. Five incidents, one theme: Twitter spam as a weapon to drown voices of protest. ***In FOCI***.

Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. Evaluation methods for topic models. ***In Proceedings of the 26th International Conference on Machine Learning***, Montreal, Canada, 2009 *1105–1112*.

Wallsten, K. 2008. Political blogs: transmission belts, soapboxes, mobilizers, or conversation starters? ***Journal of Information Technology & Politics*** 4(3): 19–40.

Wang, G., Mohanlal, M., Wilson, C., Wang, X., Metzger, M., Zheng, H., & Zhao, B. Y. 2012. Social Turing tests: Crowdsourcing Sybil detection. ***Social Information Networks***.

Waters, R. D., & Jamal, J. Y. 2011. Tweet, tweet, tweet: A content analysis of nonprofit organizations' Twitter updates. ***Public Relations Review***, *37*, 321-324.

Watts, J. 2013. Brazil protests: president to hold emergency meeting. The full title of the work is part of https://www.theguardian.com/world/2013/jun/21/brazil-protests-president-emergency-meeting. Retrieved 4/25/2017.

Wetzel, Christopher. 2010. The Dilemma of Differential Mobilization: Framing Strategies and Shaping Engagement in the Occupation of Alcatraz. ***Research in Social Movements, Conflict and Change*** 30: 239-70.

Wu, F., Huberman, B. A., Adamic, L. A., & Tyler, J. R. 2004. Information flow in social groups. ***Physica A,*** *337, 327–335.*

Zwerman, G., Steinhoff, P. G., & della Porta, D. 2000. Disappearing Social Movements: Clandestinity in the Cycle of New Left Protest in the U.S., Japan, Germany, and Italy. ***Mobilization*** 5(1): 85-104.

Appendix A

Bot Detection Approach

We used the funnel process to detect bots (Salge & Karahanna, 2018). Our choice is appropriate for two reasons. First, the method is complementary, combining three different techniques (social networks, machine-learning, and crowdsourcing of humans) in a holistic and precise, but also feasible, approach to detecting bots. Second, the process entails triangulation in that it requires the use of several data sources for detection. This bolsters the confidence of our findings—we are convinced they come from analyzing bots instead of humans. In keeping with our application of the funnel process, we began by identifying potential mechanisms of bot impact. Since existing work has shown bots amplify the magnitude of protests through automated retweets (Oliveira, França, Goya, & Penteado, 2016), we calculated out-degree centrality scores for every node in our sample as to control for that. To quantitatively select broadcasters—i.e., those central in the network with respect to outgoing volume—we applied the median plus two times the standard deviation. This is suitable for many reasons. First, the data is positively skewed. With asymmetrical distributions, the mean is biased while the median is not. Second, broadcasters are outliers and therefore significantly different than common users. A traditional way to identify outliers is to select nodes having scores that are two or three standard deviations above the central tendency of a distribution[67] (Kline, 2011). To have useful variation within a subset of broadcasters, we chose the lowest cutoff value (two instead of three) for the standard deviation. Out of the 142,728 nodes in our sample, 5,314 are broadcasters.

[67] We initially computed interquartile range (IQR) instead of standard deviation but decided to use the latter given the skewness of our data – many IQR scores were zero and therefore not analytically useful for distinguishing broadcasters from other actors.

We then used two computer algorithms to identify which broadcasters were bots. Before selecting BotOrNot (ABTO Software[68]) and DeBot (Chavoshi, Hamooni, & Mueen, 2016) as our algorithms of choice, we explored a variety of options including (but not limited to) the Bot or Not? system API developed by Ferrara et al. (2016)[69] and the decision tree classifier of Cresci, Di Pietro, Petrocchi, Spognardi, and Tesconi (2015). Such exploration is a key strength of our approach since we evaluated many different algorithms. BotOrNot and DeBot were the most appropriate because they have face validity—classifying publicly known bots, such as *anonymousfrai*, *poem_exe* and *thricedotted*, as actual bots—and because they use distinct methodologies[70], enabling a different mathematical procedure to detection. The algorithms flagged 829 out of the 5,314 broadcasters as bots (BotOrNot labeled 722 and DeBot marked 188 for an overlap of 81).

Finally, to avoid Type I error (algorithms labeling an actor as a bot when the node is in fact a person), the first author and a graduate student in computer science manually and independently assessed whether each identified focal bot was in fact a bot (inter-coder reliability = 0.98). Sourcing the work to a graduate student in computer science was appropriate since previous research indicates that they are remarkably accurate in detecting bots (Wang, Mohanlal, Wilson, Wang, Metzger, Zheng, & Zhao, 2012). To rule out Type II error (algorithms labeling an actor as a human when the node is in fact a bot), the coders further rated the validity of 100 randomly selected broadcasters which were identified as humans. None of them were validated as bots (inter-coder reliability = 0.99). This shows that the computer algorithms we chose are fairly accurate in detecting non-bot actors.

[68] http://botornot.co
[69] http://truthy.indiana.edu/botornot/
[70] DeBot uses *between-user* activity correlation (strong association is indicative of bot behavior) while BotOrNot uses *within-user* activity patterns (a set of user behavior criteria (e.g., retweet count) is used to detect bot activity).

Out of the 829 broadcasting bots identified by BotOrNot and DeBot, the coders agreed that only a total of eight were indeed bots (see Table A1). While almost all of them (*Anon_RT*, *blanketRT*, *instant_RT*, *TeamRevoltNow*, *Toni_JoiaRara*, and *WorldCupRetweet*) joined Twitter on 2013—the year in which the movement occurred—we observe significant heterogeneity in their profile description and activity (see Table A1 for descriptive statistics). For example, two bots do not have written biographies (*blanketRT* and BrasilRetwittes) while three (*BrasilRetwittes*, *FreeportIL*, and *TeamRevoltNow*) have a much larger audience (over 2,000 followers) than the others (less than 200). These three bots have also shared more messages (minimum of 32,319 updates) than the rest (maximum of 14,062 updates).

Table A1. Profile Description of Focal Bots

Bot[a]	Join Date	Location	Bio	Photo	Following[b]	Followers[c]	Updates[d]
Anon_RT	2013-06-05	NA	#Anonymous		43	61	1,355
BlanketRT	2013-06-01	NA	NA		5	196	14,062
BrasilRetwittes	2012-09-24	NA	NA		0	2,340	188,672
FreeportIL	2010-02-04	Freeport, USA	Freeport Illinois. Originally called Winnneshiek, affectionately Pretzel City, home to the second debate between Lincoln and Douglas.Followsback #F4F		2,872	2,596	32,319
instant_RT	2013-05-07	NA	Instant #InstantRT to any of our followers who use #InstantRT in their tweets.. Also you can only use 1 other hashtag outside of #InstantRT.. Enjoy..		0	94	12,523
TeamRevoltNow	2013-03-03	NA	@TeamRevoltNow is an open Twitter stream highlighting the ongoing worldwide revolution against Fascism/Oppression #TeamRevolution		234	2,266	97,003
Toni_JoiaRara	2013-06-17	NA	*Character of Thiago Lacerda in #JoiaRara #ToniVidaloka		29	32	2,020

WorldCupReTweet	2013-06-20	All over the world	I will retweet anything related to The World Cup! Keep up to date with the latest News & Gossip with one follow!		7	14	620

[a] All bot names correspond to their public Twitter screen names. They are all *rule-based*.

[b, c, d] Average number of (*following*, *followers*, *updates*) bots accumulated during the movement.

* Translated from Portuguese by the first author.

Appendix B

Topic Modeling

We used topic modeling in our research for two reasons. First, algorithms like LDA and STM are capable of processing a large collection of documents that would otherwise be too costly to code manually. Second, the method is widely-used (Huang, Lehavy, Zang, & Zheng, 2017; Shi, Lee, & Whinston, 2016) and provides a reliable and replicable classification of topics (Huang, Lehavy, Zang, & Zheng, 2017). We use a case example to show its intuition in Figure B1. Because STM builds off LDA, our case example represents the logic of the latter.

The document we illustrate contains a few Twitter messages sampled from Anon_RT—a bot affiliated with Anonymous, a group of hacktivists opposing Internet censorship and control. By hand, we have pointed out different words used in messages shared by Anon_RT. For example, English words about *Anons*[71], such as “anonymous” and “yal” (YourAnonLive) are shown in yellow; words in Portuguese about Brazil’s cause, such as “vemprarua” (come to the street) and “ogiganteacordou” (the giant has woken up) are colored in orange; and English words about social activism, such as “revolution” and “protest” are illustrated in blue. If we were to manually pick out every word in this document, we would likely notice that this bot shared, in different proportions, messages in Portuguese and English that were about Anonymous, the movement in Brazil, and also social activism more broadly. Having this knowledge is useful because we can now situate Anon_RT’s topic content in a collection of other bot documents.

[71] This is how members of the group describe themselves.

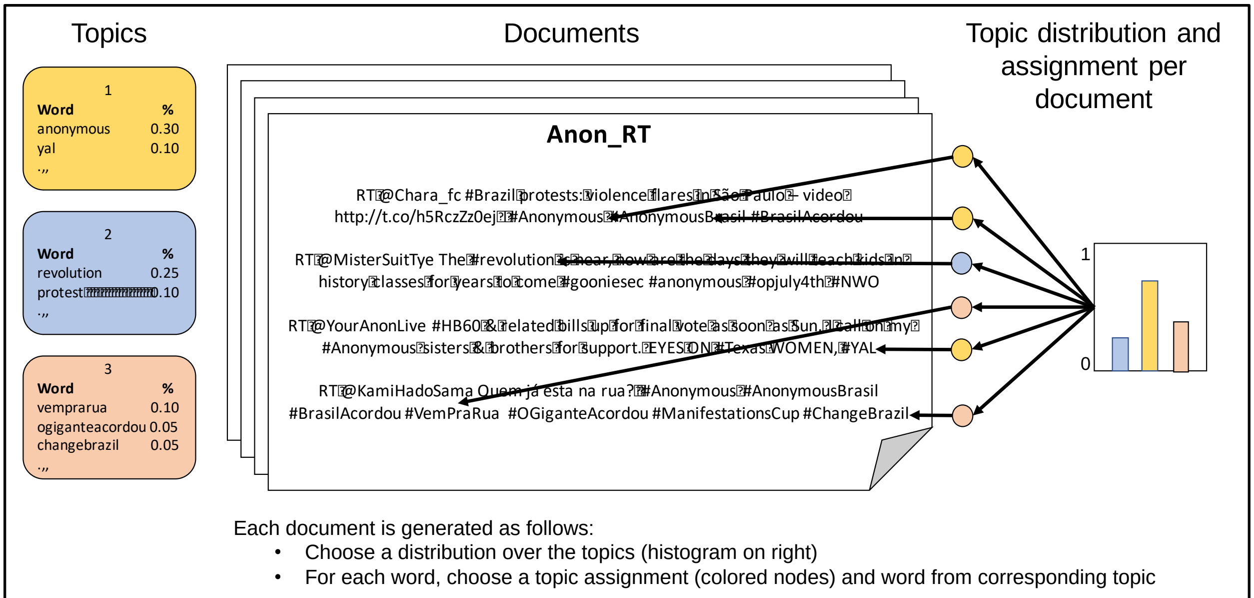

Figure B1. The Intuition Behind latent Dirichlet allocation

Because LDA is grounded in a statistical generative process and observed documents are often used as data inputs to infer hidden topic structures, we can think of topic modeling and LDA specifically as reversing the generative process—i.e., the question that the model is trying to answer is: What is the hidden topic structure that likely generated the observed collection of documents we have? (see Blei, 2012). To do that, the algorithm carries two key assumptions. First, because words carry strong semantic information it presumes that documents with similar topics have similar groups of words. Latent topics are therefore discovered by identifying groups of words that frequently co-occur together in documents. The second assumption is concerned with the structure of documents themselves. The model presumes that documents are probability distributions of relating topics and that topics are probability distributions over words, meaning that every document has a number of topics and each topic has a distribution of words associated with it. Conditional on these assumptions, each document is generated as follows. First, a distribution over topics is drawn from a global prior distribution. Then, for each word in a document, a topic is drawn from a multinomial distribution based on its distribution over topics.

Based on the chosen topic, an observed word is drawn from a distribution over the vocabulary. Finally, the model is completed by assuming a dirichlet prior for the topic proportions such that the local variables (θ and z) are estimated for each document followed by a maximization of the global parameters (β and α in Figure 1 of Blei, Ng, & Jordan, 2003).

LDA has two key limitations. First, it is unable to model topic correlation even though we intuitively know that, for example, a document about Brazil's movement is more likely to also include other social causes rather than something like biology or astronomy. Second, LDA lacks the flexibility to accommodate model covariates. This is problematic because information about documents can be useful for improving the estimation of topics. STM was specifically developed to address these two shortcomings of LDA. Like LDA, each document arises as a mixture of a number of topics. But STM allows for correlations in the topic proportions by using logistic normal distribution (Aitchison & Shen, 1980). In addition, while LDA maximizes global parameters shared by all documents, STM specify these parameters as a function of document-level covariates. In the absence of covariates, the model reduces to the correlated topic model (CTM) introduced by Blei & Lafferty (2007). Two kinds of covariates can be modeled in STM: topic prevalence (X) and topic content (Y). Prevalence controls the allocation of words to topics while content controls the frequency of the terms in each topic. The model therefore allows researchers to capture how document information affects how a topic is discussed (prevalence) and the language used to discuss the topic (content). See Figure B2 for a graphical illustration.

In short, there are three major differences between STM and LDA. First, in STM topics can be correlated while in LDA they cannot. Second, in STM each document has its own prior distribution over topics which are defined by X while in LDA each document shares a global mean. Finally, in STM word use within a topic can vary by covariate Y while this sort of variation is not captured by LDA. STM covariates therefore provide structure to the prior distributions in the model, injecting useful information into the inference procedure.

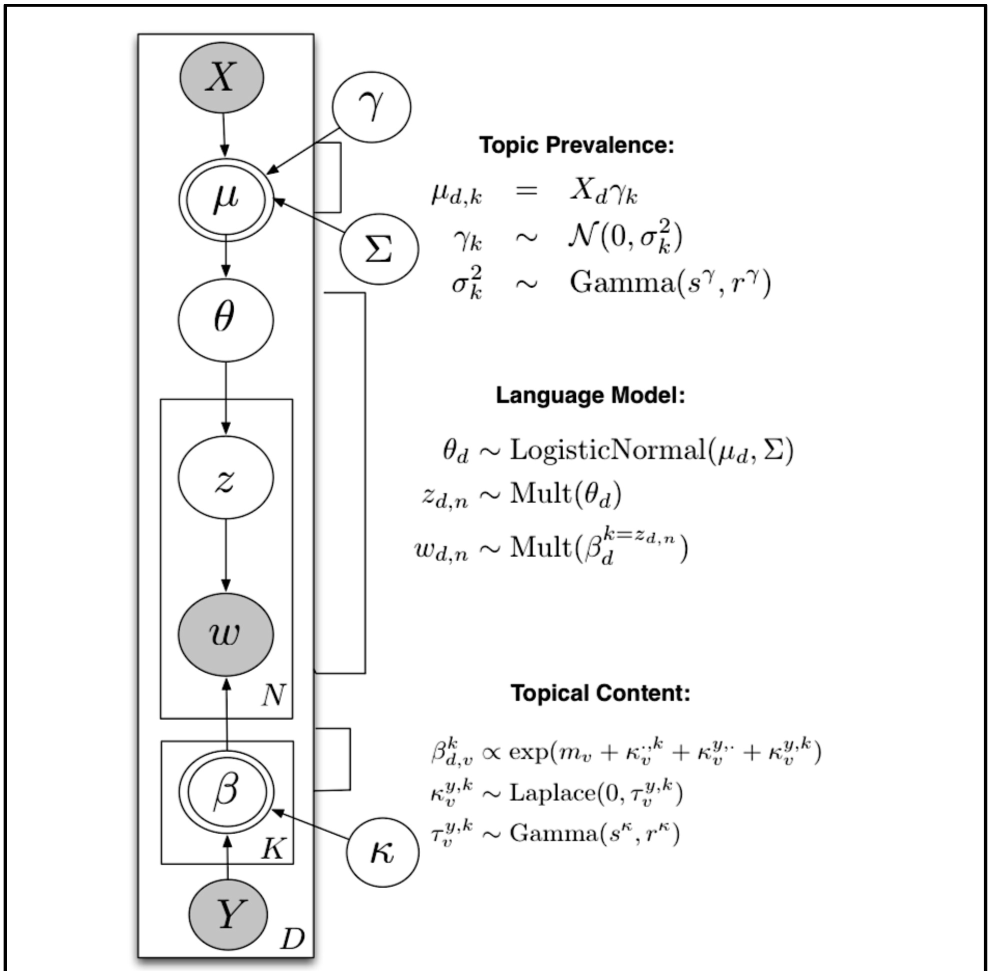

Figure B2. The Structural Topic Model Plate Diagram

Validating STM Model Outputs

We performed a variety of robustness checks for our topic outputs. First, we assessed results for semantic validity. In doing so, we followed Atkins, Rubin, Steyvers, Doeden, Baucom, and Christense (2012), Bao and Datta (2014), and Quinn, Monroe, Colaresi, Crespin, and Radev (2010) in that we manually read the high probability terms in each identified topic and their respective messages. This allowed us to compare and verify every model for its coherence, interpretability, usefulness, and meaningfulness.

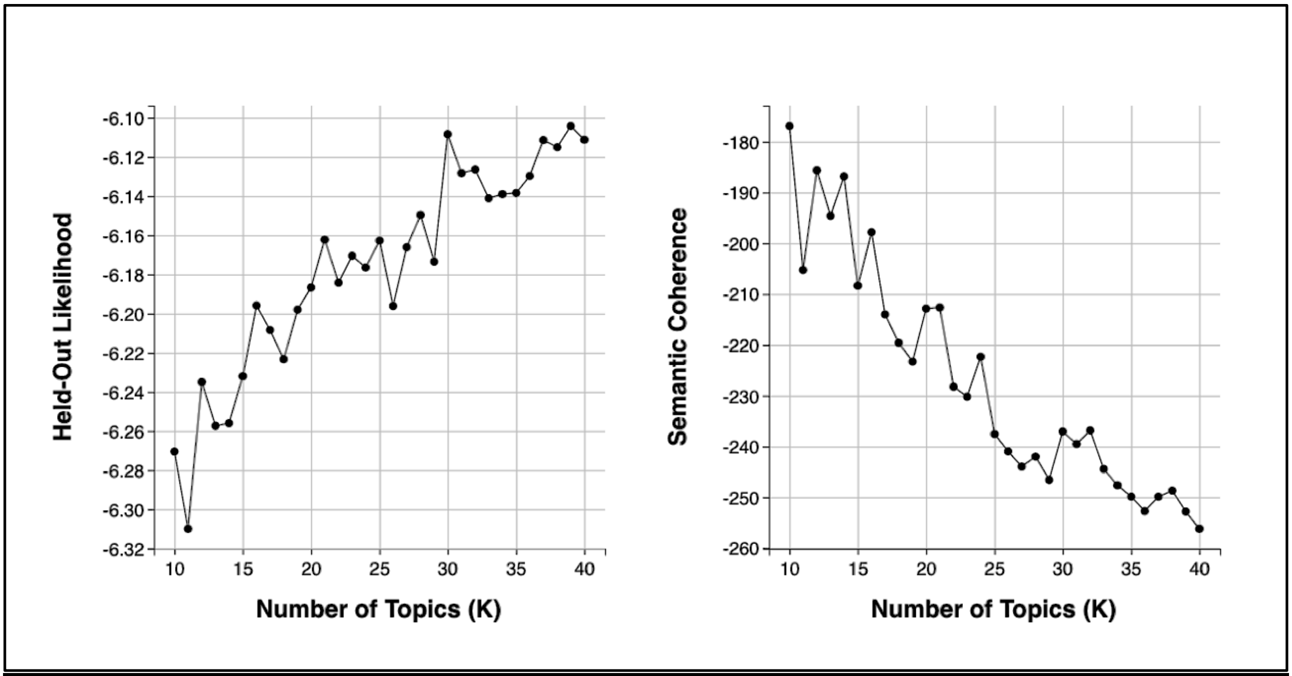

Figure B3. Diagnostic Values by Number of Topics

We also used topic coherence (Mimno et al., 2011) and the held-out likelihood (Wallach et al., 2009) to statistically compare the quality of our models. Topic coherence captures the co-occurrences of the most probable terms within each topic and higher scores are indicative of higher topic quality. The measure has shown high correlation with the judgment of human annotators (Li, Yin, & Chen, 2017). The held-out likelihood test measures how well a model fits a new set of data. Like topic coherence, higher scores are also indicative of superior performance. As shown in Figure B3, we observe that these measures do not agree with each other—increases in topic coherence are associated with decreases in model fit. Since our goal is to develop a model interpretable to humans (instead of a model to serve as a parameter in a follow up regression analysis) we primarily used topic coherence to guide our choice of number of topics. Yet we also consider the held-out likelihood test as to not dismiss significant variance in thematic content across and within topics and thus harm model fit. The structural model we ran with 21 topics represents a compromise between the two measures which errs more favorably toward topic coherence given the focus of our study.

To further check the robustness of our findings, we used the algorithm of Lee and Mimno (2014) which allows us to specify K=0 and thus automatically estimate the number of topics. Because their algorithm is not deterministic like the standard Spectral initialization that we use, running it with a different seed can result in different results and also a different number of topics. To prevent against this issue we ran the K=0 model multiple times in different computers and with distinct seed numbers. Our results remained unchanged in that the suggested number of topics was 39. Contrasting this result with those presented in Figure B3, we notice that the algorithm of Lee and Mimno agree with the results provided by the held-out likelihood test.

Chapter 5. Conclusion

Summary of Results

In the first paper, we discover that what we assumed to be human protestors were in some cases bots—automated accounts in online social networks. To explicate the discovery of bots, we problematize an implicit assumption of online social network research within Information Systems and Management as it pertains to the concept of actors. We also discuss how neglecting bots can threaten research validity and we position bot as a crucial actor with implications for organizational theory and practice. Finally, we provide scholars investigating social phenomena online with a multi-method approach for detecting bots.

In the second paper we inductively study the role of bots in Brazil's 2013 social movement on Twitter. Our results show that they act as *BrokeCasters*—i.e., as brokers and broadcasters of the cause. We also identify two new mechanisms for such *BrokeCasting*. The first mechanism, *Actor Tagging Volume,* relies on effective network volume and is part of broadcast. *Actor Tagging Volume* requires bots to leverage actor tagging features of social media, such as mentions, to share few messages and reach many non-redundant actors. The second mechanism is part of brokerage and relies on *Content Tagging Diversity*, a means by which bots advantageously use content tagging features of social media, such as hashtags, to jointly access information associated and *not* associated with social movements. We conclude the paper with a theoretical framework of *BrokeCasting* which SMOs and activists can use to raise awareness of social movements on social media.

The third paper builds on the findings of the second, especially in regard to content. We examine 12,604 posts generated by bots during Brazil's 2013 political movement on Twitter to understand *what* topics bots share on this platform and *how* topic relations and content characteristics can be leveraged for resource mobilization on social media. We find that that bots engage in two forms of social activism. The first is *Focal Botivism* and requires bots to share positive and negative messages in correlated topics which altogether promote social change in specific a population and for a specific cause. The other is *Global Botivism*, a means by which bots share positive and negative messages in correlated topics which altogether promote social change across the globe. We conclude the paper with a framework of *Botivism* which SMOs and activists can use to mobilize resources for social movements on social media.

Altogether, these papers provide valuable insights into understanding how bots can be leveraged for social change. By discovering that bots are central actors in protest social networks, understanding how they can be designed to raise awareness of and mobilize resources for a social movement on social media, we provide a solid foundation for future research in this area. More specifically, our theoretical models for papers 2 and 3 contribute a theoretical lens to a nascent—mostly empirical—literature on the use of bots in social movements. Given that the majority of the extant social network literature is content-agnostic, our models also contribute to this body of work by highlighting the importance of content and content-tagging social media features in bridging and raising awareness and also amplifying and mobilizing resources for movements—whic can be leveraged not just by bots but by any actor on these social networks. It is important, however, to note that *Botivism* is different than the social activism enacted by humans in two ways. First, the marginal cost incurred with *Botivism* is nearly zero since the sharing of content is automated. Second, the scale of *Botivism* is much larger since bots share content at a much higher volume and speed than any human possibly can. Therefore, bots can augment human action in social movements by providing an unparalleled level of scalability, speed, and volume to content sharing, reach, and resource mobilization.