**Contentious Politics in the Age of the Botnet**

Xinlan Emily Hu

25 June 2019

Miquela Sousa appears to be your typical Instagram It girl. A Brazilian-Spanish model, she boasts 1.6 million followers on Instagram and produces music in her spare time. More than 80,000 listeners have streamed her songs on Spotify; she has modeled for Prada and flaunted tattoos from Miley Cyrus' tattoo artist[1].

Neil Turner is a vocal right-wing activist. With 29,000 followers on Twitter, Turner has been retweeted by Donald Trump five times[2]. On the eve of the 2016 election, Turner shared a doctored image of a supposedly undocumented immigrant being arrested at the polls. Despite being debunked, his post quickly went viral on Twitter[3].

What do Miquela Sousa and Neil Turner have in common? Both are social media influencers. Both are products of an increasingly technology-mediated world. And both are robots. Neither "Lil Miquela," as she is called, nor Neil Turner exists. Lil Miquela is an artificially-generated digital character[4]. And Neil Turner is a Twitter bot[5].

In this essay, I investigate social media as catalytic force for contentious politics, and the emergent threat that false social media accounts pose to digital activism. Social media has shifted the landscape on citizen advocacy, producing a networked digital space that discusses, advocates, and mobilizes for political issues on an unprecedented scale. This space, however, lends itself to both empowerment and manipulation.

The first part of this essay traces the origins of social media as a political tool. Drawing upon theories such as Helen Margetts's 'chaotic system,' Zizi Papacharissi's 'affective publics,' and Lance Bennett's 'digitally networked action,' this essay offers a three-part analysis of social media as a critical piece in the contentious repertoire. Social media provides accurate information for participants; communicates emotive and personalized messages; and does so at a minimal cost.

However, the rise of false accounts—including 'bots,' state-sponsored propaganda accounts, and other manufactured users—is digital communication's Trojan Horse. Using bots and botnets as a case study, the second part of this essay describes how false accounts appropriate the very features that made social media successful in the first place. If social media is effective because it relays accurate information and enables expressive storytelling, bots instead flood the network with distractions and hijack affective narratives.

---

[1] Hsu, Tiffany. "These Influencers Aren't Flesh and Blood, Yet Millions Follow Them." The New York Times. June 17, 2019. Accessed June 25, 2019. https://www.nytimes.com/2019/06/17/business/media/miquela-virtual-influencer.html.

[2] Jensen, K. Thor. "Inside Donald Trump's Twitter-Bot Fan Club." Intelligencer. June 15, 2016. Accessed June 25, 2019. http://nymag.com/intelligencer/2016/06/inside-donald-trumps-twitter-bot-fan-club.html.

[3] Garcia, Feliks. "Pro-Trump Twitter Bot 'Neil Turner' Posts Fake Image of Immigration Officer Arresting Latino Voter." The Independent. October 29, 2016. Accessed June 25, 2019. https://www.independent.co.uk/news/world/americas/twitter-bot-immigration-arrest-latino-voter-neil-turner-hoax-us-election-2016-a7386061.html.

[4] Hsu, Tiffany. "These Influencers Aren't Flesh and Blood, Yet Millions Follow Them."

[5] Garcia, Feliks. ". "Pro-Trump Twitter Bot 'Neil Turner'."

Finally, this essay grapples with the practical question of bots' undue political influence. The final section examines strategies for identifying, removing, and curbing the activity of online bots. It concludes with new avenues for research as human users learn to coexist with a near-inevitable bot presence.

## *Contentious Politics Through a Digital Network: Three Ingredients*

### *Background*

As I outline a framework for social media and contentious politics, it is useful to first define relevant terms. My paper draws upon Charles Tilly and Sidney Tarrow's vocabulary of *contentious politics*, defined as "interactions in which actors make claims bearing on other actors' interests, leading to coordinated efforts on behalf of shared interests or programs[6]." When these coordinated efforts (*contentious performances*) are sustained with "repeated performances" across "organizations, networks, traditions, and solidarities[7]," they form a *social movement*. Over time, one's repertoire of contentious performances evolves as participants innovate new political strategies.

Digital communication introduces a new element to contentious repertoires. In particular, they involve participants who would otherwise be too risk-averse to join, and they leverage unique avenues of persuasion and influence. Scholars have introduced various terms to describe these strategies of online political organization: Bennett refers to "digitally networked action" and "connective action[8];" Milan refers to "cloud protesting[9];" Margetts uses "online collective action[10]." These movement are often leaderless[11], unaffiliated with traditional institutions[12], and extremely fast-growing—within 24 hours of the first #MeToo tweet in October 2017, it had already been shared in 12 million Facebook posts, by 4.7 million users[13].

Given the diversity of viewpoints on this phenomenon, the first section of this essay will establish a unified framework for understanding digitally-mediated contentious politics. As I incorporate contemporary scholarship, I will sometimes use other scholars' terms interchangeably (e.g., 'online collective action,' 'digitally networked action'). My argument is that the success of digitally-mediated contentious politics can be attributed to three critical ingredients: rectifying information asymmetries, creating emotive and personalized messaging,

---

[6] Tilly, Charles, and Sidney G. Tarrow. *Contentious Politics*. (New York, NY: Oxford University Press, 2015), 7.
[7] Tilly Charles and Sidney Tarrow, *Contentious Politics*, 11.
[8]Bennett, W. Lance, and Alexandra Segerberg. "The Logic Of Connective Action." (*Information, Communication & Society* 15, no. 5 (2012): 739-68. doi:10.1080/1369118x.2012.670661), 743.
[9] Milan, Stefania. "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting." (*Social Media Society* 1, no. 2 (2015): 205630511562248. doi:10.1177/2056305115622481), 4.
[10] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence: How Social Media Shape Collective Action*. (Princeton: Princeton University Press, 2017), 26.
[11] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 13.
[12] Bennett, W. Lance, and Alexandra Segerberg. "The Logic Of Connective Action," 741.
[13] "More than 12M "Me Too" Facebook Posts, Comments, Reactions in 24 Hours." CBS News. October 17, 2017. Accessed June 25, 2019. https://www.cbsnews.com/news/metoo-more-than-12-million-facebook-posts-comments-reactions-24-hours/.

and reducing the cost of mobilization. Crucially, however, each of these ingredients is premised upon the expectation that users share information in good faith. When this premise is broken—as when vast networks of bots generate faux social interest—social media's assets become its liabilities. Thus, this section will also serve as crucial framing for the later discussion of bots and other false accounts.

## *Rectifying Information Asymmetries*

Ultimately, information is the core of digital technology, and the first ingredient to online collective action's success. Helen Margetts argues that, with the advent of social media, an individual's decision calculus becomes more closely-attuned to reality. Due to an informational asymmetry in social movements, it is often difficult for observers to ascertain whether a given political demand has any chance of success. In general, the success of social movements tends to be a right-tailed distribution: that is, a far greater number of social movements fail than succeed. 43% of petitions posted on the UK website die on arrival—garnering fewer than five signatures[14]. These rather poor odds deter risk-averse individuals from participation: if the cause is lost, why expend the effort? If adopted universally, however, the attitude would drain collective action, and nearly all social movements would struggle to leave the ground.

Social media provides a signal that distinguishes between a successful movement and a floundering one. "In an Internet-based environment, people are likely to know how many other people are joining in real time so they can match their joining point to their own threshold[15]." This environment taps into a psychological phenomenon known as *social proof*—when individuals are uncertain about decisions, they conform to others' evaluations[16]. One empirical account of this phenomenon, the BOP (Burden of Social Proof/Balance of Pressures) framework, posits that the social balance of viewpoints for and against one's own view can be mathematically modeled as a conformity function[17]. As the signals of popularity strengthen, an individual becomes more likely to conform with others. Since social media networks heighten the intensity of such social signals (for example, by displaying dozens of friends who had signed a petition), they raise the likelihood that an individual chooses to conform—in this case, participate in the movement.

However, the BOP simulations suggest that social networking tools, in creating "this broadened scope of influence[,] can make societies more volatile[18]." Margetts terms this volatility a *chaotic system*, whereby "a high degree of connectivity…creates positive and negative feedback, causing changes to take place exponentially faster[19]." In other words, where

---

[14] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 26.
[15] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 25.
[16] Sundie, Jill M., Robert B. Cialdini, Vladas Griskevicius, and Douglas T. Kenrick. "The Worlds (truly) Oldest Profession: Social Influence in Evolutionary Perspective." (*Social Influence* 7, no. 3 (2012): 134-53. doi:10.1080/15534510.2011.649890), 144.
[17] Maccoun, Robert J. "The Burden of Social Proof: Shared Thresholds and Social Influence." (*Psychological Review* 119, no. 2 (2012): 345-72. doi:10.1037/a0027121), 346.
[18] MacCoun, Robert J, "The burden of social proof," 366.
[19] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 196.

there previously existed an informational asymmetry—outsiders had far less information about the viability of a social movements than insiders—social media precipitates feedback loops that enable participants to match their actions to real-time updates. Of course, volatility is a double-edged sword. The entire precarious system is premised upon the idea that the information in these feedback loops is accurate. This phenomenon—as I will later show—can be leveraged to generate artificial public support for a nonexistent cause, or enable special interest groups to sway the political electorate. Indeed, the very chaos that gives rise to progress can just as easily collapse society inward.

## *Creating Emotive, Personalized Messaging*

Contentious politics are thus defined by moments of volatility and chaos, during which social movements can spring from obscurity to sudden virality. When mediated by digital networks, these movements also take on a unique pathos. A second ingredient in these movements' success is the creation of emotive and personalized channels. "Through the selection process regarding which social media platforms to use, and how long to spend there, individuals personalize information streams that they receive[20]," which generates custom-tailored political messaging. These messages often travel with first-person narratives, Facebook frames, catchy hashtags and personal images[21]; they circulate among family members and trusted friends.

In the wake of the 2015 *Obergefell* case, which upheld same-sex marriage in the United States, more than 26 million Facebook users added a rainbow overlay to their profile pictures, representing the colors of the LGBT flag. The rainbow filter joined a diverse milieu: "There were green filters for Iranian protesters in 2009, yellow ribbons for Hong Kong in 2014, black dots to oppose sexual violence in India, Arabic "Ns" to support Iraqi Christians[22]." More recently, the #MeToo movement has prominently featured the use of personal narratives in driving an ongoing conversation about sexual assault. With #MeToo in particular, personal narratives have served not only as a rhetorical vehicle for a larger social criticism, but also as a cathartic experience for sexual assault survivors. The ability to disclose information to practical strangers online enables victims who may feel stigmatized over their experience to "disclose personal information without risking embarrassment or adverse reactions from close others[23]."

The networked public produces a rich blend of truth and emotion that Zizi Papacharissi terms the *affective public*—"affective statements…mix fact with opinion, and with emotion, in a

---

[20] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 22.

[21] Bennett, W. Lance, and Alexandra Segerberg. "The Logic Of Connective Action," 742.

[22] Dewey, Caitlin. "More than 26 Million People Have Changed Their Facebook Picture to a Rainbow Flag. Here's Why That Matters." The Washington Post. June 29, 2015. Accessed June 25, 2019. https://www.washingtonpost.com/news/the-intersect/wp/2015/06/29/more-than-26-million-people-have-changed-their-facebook-picture-to-a-rainbow-flag-heres-why-that-matters/?utm_term=.7ab34c1d19b7.

[23] Hosterman, Alec R., Naomi R. Johnson, Ryan Stouffer, and Steven Herring. "Twitter, Social Support Messages, and the #MeToo Movement." (*The Journal of Social Media in Society*7, no. 2 (2018): 69-91. Accessed June 25, 2019. http://thejsms.org/index.php/TSMRI/article/view/475), 73-4.

manner that stimulates the way that we politically react in our everyday lives[24]." This notion of the public stands in stark contrast to the idealist version described by Habermas, in which private individuals voice criticism by "making use of their reason[25]." Through viral personal stories, the first two ingredients of digitally-mediated contentious politics create a virtuous cycle. Emotion contributes to the chaos, fueling moments of activity: "The soft, networked structures of feeling….can drive powerful disruption, help accumulate intensity and tension, or simply sustain infinite loops of activity and inactivity[26]."

### *Reducing Cost*

The final ingredient provides the spark. Online social movements are (almost ludicrously) low-cost. Organizations can cheaply broadcast to millions of followers in seconds; many also share platform infrastructure to drive costs lower[27]. Communication is easily automated—users can set up a simple Twitter bot in five minutes[28].

With society deeply intertwined in digital networks and fueled by a cacophony of affective content, it takes mere minutes for a group of passionate individuals to set the world spiraling in some new direction. "This kind of collective action," Margetts writes, "will continue to act as an important influence on policy change. It will inject turbulence into every area of politics, acting as an unruly, unpredictable influence on political life[29]."

But what if the spark is caused not by some group of passionate activists, but by an army of shadows—bot accounts and computer-generated personas of users who never existed? In the next section, I argue that the rise of the botnet has soured each of the ingredients that made contentious politics online so influential.

## *An Army of Shadows*

A decade ago, bots and other false accounts would hardly have been a substantial concern. Bots were haphazard pieces of code, nameless accounts with no profile information and garbled, at times profane, messages[30]. Since then, however, bots have enjoyed a growing, and increasingly sophisticated, online presence. Bot accounts are now equipped with "carefully staged photos, canned but well-crafted responses to other users," and—perhaps most importantly—political objectives[31].

---

[24] Papacharissi, Zizi. *Affective Publics: Sentiment, Technology and Politics*. (New York: Oxford University Press, 2017.), 27.
[25] Habermas, Jürgen. *The Structural Transformation of the Public Sphere an Inquiry into a Category of Bourgeois Society*. (Cambridge: Polity, 2015), 51.
[26] Papacharissi, Zizi. *Affective Publics*, 29.
[27] Milan, Stefania. "When Algorithms Shape Collective Action," 5.
[28] "Create a Twitter Bot - Getting Started Guide." Digital Inspiration. July 21, 2017. Accessed June 25, 2019. https://digitalinspiration.com/twitter-bots-tutorial-4796.
[29] Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence*, 192.
[30] Howard, Philip N., Samuel Woolley, and Ryan Calo. "Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration." (*Journal of Information Technology & Politics* 15, no. 2 (2018): 81-93. doi:10.1080/19331681.2018.1448735), 84
[31] Ibid., 84

To clarify my use of the term "bot," I draw from Grimme, et. al.'s conception of social bots as any automated or partially automated software agent that produces an "imitation of human communication (behavior)[32]." Bots are a strict subset of the general category of false online accounts, which also include, for example, computer-generated online influencers (such as Lil Miquela). However, the bulk of this essay will focus on social bots. Recent literature indicates that "advanced social bots may no longer aim at mimicking human behavior, but rather at misdirecting attention to irrelevant information[33]." These social bots often work in coordinated patterns known as bot networks, or *botnets*, which can induce a surprising amount of influence in the social landscape.

Bots now generate 55% of all online content[34], and they are often so realistic that humans have trouble differentiating a bot from a real user[35]. By all outward accounts, Neil Turner appears to be a genuine human being who happens to lean far right. But then one notices that his average Twitter response time is three seconds—far shorter than is humanly possible to read and compose a tweet. And rather than reply to the content of the original tweet, Neil's tweets typically consist of a short pro-Trump platitude with an accompanying stock image[36].

However, a reasonable user could easily miss such details, especially when the bot's tweets are featured at the top of the feed. Mobile Twitter feeds are sorted chronologically, which means that, by being the first to respond, bots easily claim the most prominent display locations[37]. Occupying this location enables bots to leverage the connected network; by making a certain political viewpoint appear popular, they hijack social proof. Genuine humans begin to share the bots' social media campaigns, donate to their candidates, and even attend in-person rallies. The crux of bots' influence, therefore, lies not in the ability to shout into the Twitter-void, but to incite action in a human network. Vanessa Kitzie, in examining over 400 social bots, found that, while "bots *amplify* conversation surrounding significant issues (e.g., political elections, crisis situations), people are primarily responsible for the dissemination of bot-generated content[38]."

Social movements thrive in a chaotic system, as small campaigns can experience exponential growth overnight. Disinformation campaigns, however, reap the same benefits. Combining the vast network reach of bots with careful human direction, Russian agents

---

[32]Grimme, Christian, et. al. "Social Bots: Human-Like by Means of Human Control?" p.280.

[33] Ferrara, Emilio, et. al. "Today's social bots are sophisticated and sometimes menacing. Indeed, their presence can endanger online ecosystems as well
as our society." DOI:10.1145/2818717 p, 101.

[34] Woolley, Samuel C. "Automating Power: Social Bot Interference in Global Politics." (*First Monday* 21, no. 4 (2016). doi:10.5210/fm.v21i4.6161), 1.

[35] Gorwa, Robert, and Douglas Guilbeault. "Unpacking the Social Media Bot: A Typology to Guide Research and Policy." (*Policy & Internet*, 2018. doi:10.1002/poi3.184), 17.

[36] Jensen, K. Thor. "Inside Donald Trump's Twitter-Bot Fan Club."

[37] Ibid.

[38] Kitzie, Vanessa L., Ehsan Mohammadi, and Amir Karami. ""Life Never Matters in the DEMOCRATS MIND": Examining Strategies of Retweeted Social Bots during a Mass Shooting Event." (*Proceedings of the Association for Information Science and Technology* 55, no. 1 (2018): 254-63. doi:10.1002/pra2.2018.14505501028), 260.

managed to plan and promote eight political rallies between June and November 2016[39]. In total, Russian pages managed to reach 126 million Americans with "with provocative content on race, guns, immigration and other volatile issues[40]," and Americans were none the wiser. After one event, a few participants simply wondered why "no one from Heart of Texas, which had about 250,000 likes on Facebook, had shown up for the group's own rally[41]." If only they knew that Heart of Texas was based 5,700 miles away from Texas.

Most ironically of all, false online accounts have even helped governments quell protests and confuse activists. "In China, and in the Chinese administrative regions of Tibet and Taiwan, bots have been used to quash sovereignty movements while promoting state ideals." Hashtags such as #tibet and #freetibet—initially used to track protest activity and consolidate messages of solidarity—received so many junk tweets from automated Twitter accounts that they were no longer useful for protestors[42]. Similarly, pro-Trump Twitter bots began "colonizing Clinton hashtags" by tagging negative tweets with #Clinton[43]. Tragically, the features that made online social movements successful in the first place became its Achilles' heel. Impassioned personal narratives are flooded out with canned messages and imitated affect. In the din of it all, one wonders how it is possible to hear the voices that matter.

The next section of the paper revisits the theoretical framework for digitally networked action in light of rampant false accounts and bot networks. The section lays out challenges for detecting and mitigating bot activity, proposing three possible but flawed avenues to curb bots' influence. In the final section, I will conclude that, given challenges to removing bot influence, a promising future direction is to design a more pro-social (rather than pro-bot) social media landscape.

## *Is the Well Poisoned? Challenges to Removing Bot Influence*

With bots generating more than half of online content and posing serious disruptions to social movements, researchers, politicians, and platforms alike have sought solutions. The task at hand is to preserve the values of the digital space—its productive chaos, its emotive and personalized nature, its frictionless exchange of information—while undercutting the ability of bots to take advantage of the same features. This problem is, to some extent, a fool's errand. Often, making social media spaces more difficult for faux accounts will necessarily increase

---

[39] Parlapiano, Alicia, and Jasmine C. Lee. "The Propaganda Tools Used by Russians to Influence the 2016 Election." The New York Times. February 17, 2018. Accessed June 25, 2019. https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html.

[40] Weiss, Brennan. "Russian Trolls Orchestrated Divisive Protests in the US about Trump - Here Are 9 That We Know about." Business Insider. February 21, 2018. Accessed June 25, 2019. https://www.businessinsider.com/russians-organized-pro-anti-trump-rallies-to-sow-discord-2018-2?r=US&IR=T#many-more-possible-russian-rallies-7.

[41] Ibid.

[42] Woolley, Samuel C. "Automating power," 7.

[43] Markoff, John. "Automated Pro-Trump Bots Overwhelmed Pro-Clinton Messages, Researchers Say." The New York Times. November 17, 2016. Accessed June 25, 2019. https://www.nytimes.com/2016/11/18/technology/automated-pro-trump-bots-overwhelmed-pro-clinton-messages-researchers-say.html.

friction for human users; imposing restrictions for bots could poison the well for genuine grassroots activism.

This section offers three possibilities of addressing the prevalence of false accounts in the digital space: deleting false accounts, using CAPTCHA tests, and regulating bot activity. No single solution will entirely capture the problem; however, in conjunction, these solutions promise potential steps in the right direction. Each of these solutions, on their own, could certainly merit its own separate discussion. This section is intended merely to sketch out the directions and challenges for preserving digitally-mediated contentious politics.

## *Deleting False Accounts*

The first possibility is to simply delete suspected bot accounts. In March, Facebook removed nearly 3,000 accounts linked to Russia and Iran for "coordinated inauthentic behavior[44]." Similarly, Twitter made moves to remove as many as 6% of its accounts[45]. These deletions draw on growing academic literature that seeks to computationally detect and remove inauthentic users.

Varol, et. al's seminal 2017 work, trained using Twitter's API, created a bot-detection algorithm that distilled 1,150 features in six different classes[46]. Its accuracy, as measured by Area Under the Curve (AUC), was 0.94[47]. Gilani, et. al. has also focused on specific behavioral differences between humans and bots, concluding that bots retweet more often, are five times more likely to post URL's, and post ten times as much content with their tweets[48]. The empirical research so far indicates that there are sufficient feature differences between humans and bots to make reasonable distinctions between the two.

However, while cutting-edge algorithms are reasonably well-equipped to distinguish between a simple bot and a human user, they do not account for human-bot hybridization. In computational creativity, hybridization is a strategy in which "a human designer and a computer program work together, taking turns, to reach a specific design goal[49]." Thus, hybrid bots are partially automated and partially directed by human users. When used to detect hybrid

---

[44] "Facebook Removes Accounts from Russia, Iran for 'coordinated Inauthentic Behavior'." Reuters. March 26, 2019. Accessed June 25, 2019. https://www.reuters.com/article/us-facebook-fake-accounts/facebook-removes-accounts-from-russia-iran-for-coordinated-inauthentic-behavior-idUSKCN1R70PV.

[45] Confessore, Nicholas, and Gabriel J. X. "Battling Fake Accounts, Twitter to Slash Millions of Followers." The New York Times. July 11, 2018. Accessed June 25, 2019. https://www.nytimes.com/2018/07/11/technology/twitter-fake-followers.html.

[46] Varol, Onur, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." (*International Conference on Web and Social Media (ICWSM)*, 2017. Accessed June 25, 2019. https://arxiv.org/abs/1703.03107), 281.

[47] Suárez-Serrato, Pablo, Margaret E. Roberts, Clayton Davis, and Filippo Menczer. "On the Influence of Social Bots in Online Protests." (*Lecture Notes in Computer Science Social Informatics*, 2016, 269-78. doi:10.1007/978-3-319-47874-6_19), 3.

[48] Gilani, Zafar, Reza Farahbakhsh, Gareth Tyson, Liang Wang, and Jon Crowcroft. "Of Bots and Humans (on Twitter)." (*Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 - ASONAM 17*, 2017. doi:10.1145/3110025.3110090), 6.

[49] Grimme, Christian, Mike Preuss, Lena Adam, and Heike Trautmann. "Social Bots: Human-Like by Means of Human Control?" (*Big Data* 5, no. 4 (2017): 279-93. doi:10.1089/big.2017.0044), 287.

bot accounts, Varol, et. al's classification tool becomes no better than random guessing: AUD dropped to 0.48[50].

The social implications of deletion are also complicated. The attempt to remove bot accounts has led to a number of false positives. Nina Tomasieski, a 70-year-old from Tennessee, had her account repeatedly flagged as "suspicious" when she tweeted genuine support for Donald Trump[51]. The deletion line is an incredibly fine one: on the one hand, platforms hope to remove bots that broadcast the same message in organized unison. On the other, human grassroots activists, while exercising their right to free speech, organize coordinated broadcasts of their own that can be hard to differentiate from bot activity. Tomasieski and other conservatives use Twitter "rooms," which are group messages containing 50 or so members. Each member composes and tweets their own messages, while also re-tweeting the messages of fellow group members. Thus, "a tweet that Tomasieski sends may be seen by her roughly 51,000 followers, but then be retweeted by dozens more people, each of whom may have 50,000 or more followers.[52]" A human botnet, indeed.

False positives in the case of flagging genuine grassroots organizers can be a particularly sensitive issue. Deleting too aggressively erases the very network that enables online contentious politics to thrive in the first place. While platforms have the right to enforce community norms in whatever manner they wish, it is often difficult to avoid accusations of silencing specific political viewpoints or deliberately quelling the most vocal of protestors (after all, those who post the most frequently are also the most likely to be mis-classified as bots). "Even a single false-positive error leading to the suspension of a legitimate account may foster valid concerns about censorship.[53]" Thus, deletion should be employed only in more clear-cut cases, removing egregious cases of bots and trolls, but sparing cases where the risk of misclassification is too great.

### *Requiring CAPTCHAs*

Another strategy is to require CAPTCHAs, or challenge-response tests verifying whether a user is a human, for all social media posts. CAPTCHAs have been successful in combating a variety of online abuse, and using the tests to "limit automatic posting or resharing of news links could stem bot abuse." However, they could also "add undesirable friction to benign applications of automation by legitimate entities, such as news media and emergency response coordinators[54]."

It is also unclear whether CAPTCHA would be effective in stemming the flow of online bots. Truly malicious actors can easily automate CAPTCHA tests; for example, the company Solve reCaptcha offers a web API that automatically solves CAPTCHA tests for a fee of $15 per

---

[50] Ibid.

[51] Burnett, Sara. "Crackdown on 'bots' Sweeps up People Who Tweet Often." AP NEWS. August 04, 2018. Accessed June 25, 2019. https://www.apnews.com/06efed5ede4d461fb2eac5b2c89e3c11.

[52] Ibid.

[53] Shao, Chencheng, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer. "The Spread of Fake News by Social Bots." (July 24, 2017. Accessed June 25, 2019. https://www.a51.nl/sites/default/files/pdf/1707.07592.pdf), 12.

[54] Ibid.

month[55]. As a result, CAPTCHA would likely give legitimate users trouble, while serving as only a weak blockade against the growing botnet.

### *Regulating Bots*

A third option is to pursue regulation. Particularly in the United States, Internet bots are woefully under-regulated. Although the Bipartisan Campaign Reform Act of 2002 (BCRA) places restrictions on public political advertising, online speech is virtually untouched. "[T]he Federal Election Commission (FEC), which is the agency responsible for promulgating regulations pursuant to BCRA, consistently has elected to take a 'light-touch approach' to online activities related to elections"—its definition of "public communications" excludes Internet activity, "except where an advertiser is paying for an advertisement on another person's website (11 C.F.R. 100.26) [56]." As a result, Internet media companies are essentially left to self-regulate—to proactively remove bots from their own platforms. But Internet media companies are loath to reduce their own user base except under extreme cases of backlash: "Twitter naturally wishes to maintain the largest possible user base, and reports 'monthly active users' to its shareholders[57]." While Twitter has recently deleted some of its accounts following intense public pressure, it is clear that Twitter cannot be trusted to be consistently vigilant about deletion. Thus, the prevalence of faux accounts demands intervention from a less invisible hand.

Bot regulation is a clear area where further work will be required. Given the present difficulty of accurately identifying bot accounts, policymakers should treat bots not as a one-off case but rather as a permanent fixture of the communication landscape. Online platforms should be led to re-evaluate ways in which they enable bot manipulation. The final section of this essay concludes by proposing pro-social design as a new area of research that can combat bots' influence.

## *Conclusion: Living With Bots*

On July 13, 2013, Alicia Garza watched the acquittal of George Zimmerman from her living room in Oakland, California. Garza logged onto Facebook and wrote a long post venting her personal anguish, writing that people should come together to ensure "that black lives matter." A friend, Patrisse Cullors, gave the three words a hashtag: #BlackLivesMatter[58]. With that, a movement was born.

Social media is a uniquely dynamic space, where one woman's anguish can become a nationwide movement. #BlackLivesMatter has since been tweeted nearly 30 million times on

---

[55] CodePixar. "Don't Look Anywhere Anymore, We Are the Best ReCAPTCHA Solver Online!" Automatically Solve and Bypass Google's ReCAPTCHA V2 Captcha - Solve ReCAPTCHA. Accessed June 25, 2019. https://www.solverecaptcha.com/.

[56] Howard, Philip N., Samuel Woolley, and Ryan Calo. "Algorithms, Bots, and Political Communication in the US 2016 Election," 89-90.

[57] Gorwa, Robert, and Douglas Guilbeault. "Unpacking the Social Media Bot," 20.

[58] Ibid.

Twitter[59]—in part because Cullors was already building an extensive social media network. "We built platforms on Twitter, on Facebook and on Tumblr. Opal [the third founder of Black Lives Matter] helped really develop the communications plan so that it could go viral[60]."

The networked public, connected through social media platforms, is a powerful tool in the repertoire of contentious politics. Its volatility makes it a springboard for grassroots organizers—and also a cesspool for false accounts, misleading information, and networks of bots. In recent years, bots' political influence has ranged from a mild injection of positive publicity (politicians like Mitt Romney used bots to boost their Twitter following[61]) to diluting and quelling social movements. Bots have thrown the future of social media into question: they drive online networks into a race to the bottom for fast, visceral content, governed by largely unregulated political machinery.

Worse, bots can elude even the most cutting edge technology, making them difficult to delete and control. Perhaps the better question, then, is not whether we can remove bots from the information landscape, but how platforms, users, and other actors may better live with them.

Bots thrive because they take advantage of features in social media that exist by design: frictionless posting. Live news feeds. As the political implications of these designs become clear, online platforms should rethink the presentation of rich information streams. Many reforms can be sparked through regulation: the European Union's General Data Protection Regulation (GDPR) has nudged designers to rethink user interfaces and create new privacy-oriented design frameworks[62].

These reforms can be as simple as re-organizing the news feed. Recall that Twitter organizes its mobile app chronologically—enabling bot accounts to claim the most prominent locations on the interface. Improved variations of the interface could instead prioritize relevant and engaging posts, sorting by metrics that would be more difficult for a bot to "game." Such content display mechanisms have already been implemented. For example, the *New York Times* screens all comments into "reader picks" and "NYT picks[63]."

These design ideas align with the growing area of human-computer interaction known as *pro-social design*: intentionally designing interfaces that facilitate productive discussions and civil engagement. Kim Strandberg's 2017 experiment designed a conversation interface "modelled after offline mini-publics." Standberg's platform required participants to adhere to a list of deliberative guidelines—"being honest and sincere, treating each other with respect,

---

[59] Simon, Caroline. "How Social Media Has Shaped Black Lives Matter, Five Years Later." USA Today. July 15, 2018. Accessed June 25, 2019. https://eu.usatoday.com/story/news/2018/07/12/black-lives-matter-movement-and-social-media-after-five-years/778779002/.

[60] qtd. in Touré. "A Year Inside the Black Lives Matter Movement." Rolling Stone. June 25, 2018. Accessed June 25, 2019. https://www.rollingstone.com/politics/politics-news/a-year-inside-the-black-lives-matter-movement-204982/.

[61] Woolley, Samuel C. "Automating power," 8.

[62] Barrett, Claire, and Claire Barrett. "What Does GDPR Mean for UX?" UX Collective. February 18, 2019. Accessed June 25, 2019. https://uxdesign.cc/what-does-gdpr-mean-for-ux-9b5ecbc51a43.

[63] Reagle, Joseph. "Comments on Articles Are Valuable. So How to Weed out the Trolls?" The Guardian. April 17, 2016. Accessed June 25, 2019.
https://www.theguardian.com/technology/commentisfree/2016/apr/17/comments-articles-trolls-moderate-community.

accepting the presence and merits of differing views, encouragement of arguing ones' viewpoints, etc[64]." Her experiment found that the deliberative design resulted in improved discussion outcomes—higher deliberative efficacy, increased individual trust, and a more coherent discussion[65]. Similarly, Catherine Grevet experimented with a social media platform that simply added a visual highlight to pro-social posts. The highlighting drew attention away from "shock-factor" clickbait content and toward respectful online discussions[66].

Pro-social design could be a means of evading the worst of bot-generated content—the repeated, shock-inducing platitudes accompanied by doctored stock photos. Low-effort accounts, which can be set up in a matter of hours or even minutes, would be stripped of their prominence, whereas genuine grassroots organizers would (theoretically) retain the means to engage their base. At the very least, by prioritizing quality over quantity, the redesigned platforms would increase the cost of creating a bot network.

Bots may be here to stay; however, humans ultimately hold the reins.

---

[64] Strandberg, Kim. "Designing for Democracy?: An Experimental Study Comparing the Outcomes of Citizen Discussions in Online Forums with Those of Online Discussions in a Forum Designed According to Deliberative Principles." (*European Political Science Review* 7, no. 3 (2014): 451-74. doi:10.1017/s1755773914000265), 461.

[65] Ibid., 465.

[66] Grevet, Catherine. "Being Nice on the Internet: Designing for the Coexistence of Diverse Opinions Online." PhD diss., (Georgia Institute of Technology, 2016. Abstract in April 1, 2016. https://smartech.gatech.edu/handle/1853/55002), 76.

**Bibliography**


Barrett, Claire, and Claire Barrett. "What Does GDPR Mean for UX?" UX Collective. February 18, 2019. Accessed June 25, 2019. https://uxdesign.cc/what-does-gdpr-mean-for-ux-9b5ecbc51a43.

Bennett, W. Lance, and Alexandra Segerberg. "The Logic Of Connective Action." *Information, Communication & Society*15, no. 5 (2012): 739-68. doi:10.1080/1369118x.2012.670661.

Burnett, Sara. "Crackdown on 'bots' Sweeps up People Who Tweet Often." AP NEWS. August 04, 2018. Accessed June 25, 2019. https://www.apnews.com/06efed5ede4d461fb2eac5b2c89e3c11.

CodePixar. "Don't Look Anywhere Anymore, We Are the Best ReCAPTCHA Solver Online!" Automatically Solve and Bypass Google's ReCAPTCHA V2 Captcha - Solve ReCAPTCHA. Accessed June 25, 2019. https://www.solverecaptcha.com/.

Confessore, Nicholas, and Gabriel J. X. "Battling Fake Accounts, Twitter to Slash Millions of Followers." The New York Times. July 11, 2018. Accessed June 25, 2019. https://www.nytimes.com/2018/07/11/technology/twitter-fake-followers.html.

"Create a Twitter Bot - Getting Started Guide." Digital Inspiration. July 21, 2017. Accessed June 25, 2019. https://digitalinspiration.com/twitter-bots-tutorial-4796.

Dewey, Caitlin. "More than 26 Million People Have Changed Their Facebook Picture to a Rainbow Flag. Here's Why That Matters." The Washington Post. June 29, 2015. Accessed June 25, 2019. https://www.washingtonpost.com/news/the-intersect/wp/2015/06/29/more-than-26-million-people-have-changed-their-facebook-picture-to-a-rainbow-flag-heres-why-that-matters/?utm_term=.7ab34c1d19b7.

"Facebook Removes Accounts from Russia, Iran for 'coordinated Inauthentic Behavior'." Reuters. March 26, 2019. Accessed June 25, 2019. https://www.reuters.com/article/us-facebook-fake-accounts/facebook-removes-accounts-from-russia-iran-for-coordinated-inauthentic-behavior-idUSKCN1R70PV.

Garcia, Feliks. "Pro-Trump Twitter Bot 'Neil Turner' Posts Fake Image of Immigration Officer Arresting Latino Voter." The Independent. October 29, 2016. Accessed June 25, 2019. https://www.independent.co.uk/news/world/americas/twitter-bot-immigration-arrest-latino-voter-neil-turner-hoax-us-election-2016-a7386061.html.

Gilani, Zafar, Reza Farahbakhsh, Gareth Tyson, Liang Wang, and Jon Crowcroft. "Of Bots and Humans (on Twitter)." *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 - ASONAM 17*, 2017. doi:10.1145/3110025.3110090.

Gorwa, Robert, and Douglas Guilbeault. "Unpacking the Social Media Bot: A Typology to Guide Research and Policy." *Policy & Internet*, 2018. doi:10.1002/poi3.184.

Grevet, Catherine. "Being Nice on the Internet: Designing for the Coexistence of Diverse Opinions Online." PhD diss., Georgia Institute of Technology, 2016. Abstract in April 1, 2016. https://smartech.gatech.edu/handle/1853/55002.

Grimme, Christian, Mike Preuss, Lena Adam, and Heike Trautmann. "Social Bots: Human-Like by Means of Human Control?" *Big Data*5, no. 4 (2017): 279-93. doi:10.1089/big.2017.0044.

Habermas, Jürgen. *The Structural Transformation of the Public Sphere an Inquiry into a Category of Bourgeois Society*. Cambridge: Polity, 2015.

Hosterman, Alec R., Naomi R. Johnson, Ryan Stouffer, and Steven Herring. "Twitter, Social Support Messages, and the #MeToo Movement." *The Journal of Social Media in Society*7, no. 2 (2018): 69-91. Accessed June 25, 2019. http://thejsms.org/index.php/TSMRI/article/view/475.

Howard, Philip N., Samuel Woolley, and Ryan Calo. "Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration." *Journal of Information Technology & Politics*15, no. 2 (2018): 81-93. doi:10.1080/19331681.2018.1448735.

Hsu, Tiffany. "These Influencers Aren't Flesh and Blood, Yet Millions Follow Them." The New York Times. June 17, 2019. Accessed June 25, 2019. https://www.nytimes.com/2019/06/17/business/media/miquela-virtual-influencer.html.

Jensen, K. Thor. "Inside Donald Trump's Twitter-Bot Fan Club." Intelligencer. June 15, 2016. Accessed June 25, 2019. http://nymag.com/intelligencer/2016/06/inside-donald-trumps-twitter-bot-fan-club.html.

Kitzie, Vanessa L., Ehsan Mohammadi, and Amir Karami. "“Life Never Matters in the DEMOCRATS MIND”: Examining Strategies of Retweeted Social Bots during a Mass Shooting Event." *Proceedings of the Association for Information Science and Technology*55, no. 1 (2018): 254-63. doi:10.1002/pra2.2018.14505501028.

Maccoun, Robert J. "The Burden of Social Proof: Shared Thresholds and Social Influence." *Psychological Review*119, no. 2 (2012): 345-72. doi:10.1037/a0027121.

Margetts, Helen, Peter John, Scott A. Hale, and Taha Yasseri. *Political Turbulence: How Social Media Shape Collective Action*. Princeton: Princeton University Press, 2017.

Markoff, John. "Automated Pro-Trump Bots Overwhelmed Pro-Clinton Messages, Researchers Say." The New York Times. November 17, 2016. Accessed June 25, 2019. https://www.nytimes.com/2016/11/18/technology/automated-pro-trump-bots-overwhelmed-pro-clinton-messages-researchers-say.html.

Milan, Stefania. "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting." *Social Media Society*1, no. 2 (2015): 205630511562248. doi:10.1177/2056305115622481.

"More than 12M "Me Too" Facebook Posts, Comments, Reactions in 24 Hours." CBS News. October 17, 2017. Accessed June 25, 2019. https://www.cbsnews.com/news/metoo-more-than-12-million-facebook-posts-comments-reactions-24-hours/.

Papacharissi, Zizi. *Affective Publics: Sentiment, Technology and Politics*. New York: Oxford University Press, 2017.

Parlapiano, Alicia, and Jasmine C. Lee. "The Propaganda Tools Used by Russians to Influence the 2016 Election." The New York Times. February 17, 2018. Accessed June 25, 2019. https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html.

Reagle, Joseph. "Comments on Articles Are Valuable. So How to Weed out the Trolls?" The Guardian. April 17, 2016. Accessed June 25, 2019. https://www.theguardian.com/technology/commentisfree/2016/apr/17/comments-articles-trolls-moderate-community.

Shao, Chencheng, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer. "The Spread of Fake News by Social Bots." July 24, 2017. Accessed June 25, 2019. https://www.a51.nl/sites/default/files/pdf/1707.07592.pdf.

Simon, Caroline. "How Social Media Has Shaped Black Lives Matter, Five Years Later." USA Today. July 15, 2018. Accessed June 25, 2019. https://eu.usatoday.com/story/news/2018/07/12/black-lives-matter-movement-and-social-media-after-five-years/778779002/.

Strandberg, Kim. "Designing for Democracy?: An Experimental Study Comparing the Outcomes of Citizen Discussions in Online Forums with Those of Online Discussions in a Forum Designed According to Deliberative Principles." *European Political Science Review*7, no. 3 (2014): 451-74. doi:10.1017/s1755773914000265.

Sundie, Jill M., Robert B. Cialdini, Vladas Griskevicius, and Douglas T. Kenrick. "The Worlds (truly) Oldest Profession: Social Influence in Evolutionary Perspective." *Social Influence*7, no. 3 (2012): 134-53. doi:10.1080/15534510.2011.649890.

Suárez-Serrato, Pablo, Margaret E. Roberts, Clayton Davis, and Filippo Menczer. "On the Influence of Social Bots in Online Protests." *Lecture Notes in Computer Science Social Informatics*, 2016, 269-78. doi:10.1007/978-3-319-47874-6_19.

Tilly, Charles, and Sidney G. Tarrow. *Contentious Politics*. New York, NY: Oxford University Press, 2015.

Touré. "A Year Inside the Black Lives Matter Movement." Rolling Stone. June 25, 2018. Accessed June 25, 2019. https://www.rollingstone.com/politics/politics-news/a-year-inside-the-black-lives-matter-movement-204982/.

Varol, Onur, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." *International Conference on Web and Social Media (ICWSM)*, 2017. Accessed June 25, 2019. https://arxiv.org/abs/1703.03107.

Weiss, Brennan. "Russian Trolls Orchestrated Divisive Protests in the US about Trump - Here Are 9 That We Know about." Business Insider. February 21, 2018. Accessed June 25, 2019. https://www.businessinsider.com/russians-organized-pro-anti-trump-rallies-to-sow-discord-2018-2?r=US&IR=T#many-more-possible-russian-rallies-7.

Woolley, Samuel C. "Automating Power: Social Bot Interference in Global Politics." *First Monday*21, no. 4 (2016). doi:10.5210/fm.v21i4.6161.