

Social Media Bots: Implications for Special Operations Forces

Megan K. McBride, Zack Gold, and Kasey Stricklin

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.
Cleared for public release

Abstract

CNA initiated this study of social media bots—automated programs on social media platforms—to explore their implications for US special operations forces (SOF) and the broader national security community. This report explains social media bots and botnets, explores the threat of automation and the role of social media bots as a tool of disinformation, and introduces a taxonomy of six activities that social media bots and botnets can engage in: distributing, amplifying, distorting, hijacking, flooding, and fracturing. It then identifies likely evolutions in the near- to mid-term futures and explores the implications of those futures for SOF. The report examines opportunities and risks for SOF and concludes with examples of potential SOF use in each of the six identified social media bot and botnet activities.

This document contains the best opinion of CNA at the time of issue.

It does not necessarily represent the opinion of the sponsor.

Distribution

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

9/18/2020

This work was performed under Federal Government Contract No. N00014-16-D-5003.

Cover image credit: Cover image compiled by CNA; background image from Vecteezy.com, <https://www.vecteezy.com/vector-art/192151-ai-robot-vector>.

Approved by:

September 2020



Jonathan Schroden, Research Program Director
Center for Stability and Development
Strategy, Policy, Plans, and Programs Division

Request additional copies of this document through inquiries@cna.org.

Executive Summary

Social media bots—automated programs on social media platforms—have been an object of national interest and analysis for nearly four years. Although there has been a flood of reporting on social media disinformation and manipulation, CNA initiated this study to explore the specific implications of social media bots for US special operations forces (SOF) and the broader national security community. In this report, we highlight the tremendous capacity that the successful deployment of social media bots (or networks of bots, known as “botnets”) has to influence public discourse. If deployed effectively, bots could be a powerful asset in the US government’s toolkit, with benefits that outweigh the attendant risks.

The primary object of analysis for this report is the different ways that social media bots can be used to influence conversations between social media users online. In this sense, our report is different from most of the analysis released since the 2016 US presidential election. Many reports about bots focus on the broader issue of disinformation, the disruptive ends of malicious users, or the technological challenge of stopping the activity.

Although we provide examples of disinformation and malicious social media bot activity, the focus of this report is on neither the content of automated social media posts nor the long-term goals of these accounts’ programmers. Social media bots can be programmed for malicious ends, but their activities can also be directed toward neutral purposes or even prosocial goals. For example, a social media bot can be programmed to share the weather just as easily as it can be programmed to share misinformation about a nation’s leader.

It is tempting to think of social media bots exclusively in terms of spreading information (or disinformation). However, our analysis of the objectives of social media bots (i.e., *why* bots are used) identified six distinct activities (see Figure 1).

We populated a taxonomy of these six social media bot activities (taking care to include neutral, prosocial, and malicious examples) to highlight the incredible range of this tool. In some cases, historical examples were difficult to identify. Thus, we have included examples in which analysts believe strongly that social media bots and botnets were responsible for the activity; examples in which the activity was very likely conducted by social media bots or botnets; and examples in which the activity was almost certainly conducted by humans but *could* have been conducted by social media bots or botnets.

Figure 1. Taxonomy of social media bot and botnet activities

ACTIVITY	DESCRIPTION
Distributing	Sharing content
Amplifying	Increasing the reach of a message or user
Distorting	Changing the balance of a conversation
Hijacking	Taking over a conversation
Flooding	Overwhelming a conversation or account
Fracturing	Breaking a large conversation into smaller conversations

Source: CNA.

Implications for SOF

Our study concludes with the implications for SOF of four future trends in social media and automation, which we identified through our literature review and subject matter expert (SME) discussions as likely to evolve in the near- to mid-term:

1. Regulations and authorities
2. Activity in developing countries
3. A technological arms race
4. An increase in active users

The future of social media bot **regulations and authorities** is unclear. In our companion paper, *Social Media Bots: Laws, Regulations, and Platform Policies*, we delve into US and European laws and regulations and platform policies related to social media automation. Even without specific details, the trend line is clearly toward greater regulation, with the following implications for SOF:

- Cyber Command forces and/or SOF have historically had limited ability—i.e., requiring pre-approval by the US secretary of defense (SECDEF) or by certain other persons—to engage in offensive online cyber operations. It is unlikely that SOF will be afforded authorities to deploy social media bots in the immediate future. However, this might change in the years to come.

- Greater government or self-regulation of social media platforms may make it difficult for elements within the government (e.g., SOF) to deploy social media bots.
- Before governments and platforms restrict automation, SOF may seek to gain authorities now for use in prosocial or defensive (e.g., force protection) functions.

It is possible that increased awareness of disinformation tactics and social media bots will leave the West less susceptible to malign activities. It is less clear, though, what will happen as bot **activity in developing countries** increases along with internet usage. Populations in developing countries are vulnerable targets for bot influence operations, with the following implications for SOF:

- Violent extremist organizations (VEOs), such as al-Qaeda and the Islamic State, could use social media bots to exploit vulnerabilities in developing, failing, and failed countries to further weaken governments, sow discord, and stimulate civil wars and internecine conflicts. As a result of increased instability in these areas, SOF could see increased requirements for counterterrorism (CT) forces in these countries, whether for direct action or for training partner nation security forces.
- In the era of great power competition (GPC), the US, China, and Russia are likely to seek increased influence worldwide. Given the relatively weak media environments in many developing countries, the use of social media—and social media bots and botnets—could become a primary avenue for competitive activities.
- The Joint MISO WebOps Center (JMWC)—a new organization within US Special Operations Command (SOCOM)—could be a frontline entity for detecting and combatting these types of adversary operations, against both VEOs and state actors. The JMWC and other SOF entities—such as the US Army Special Operations Command’s two psychological operations groups (POGs)—could partner with State Department entities, such as the Global Engagement Center, to provide training for partner nation forces on establishing their own social media botnets for defensive purposes. It may even be possible to leverage existing authorities to provide this type of training.
- SOF could also work in partnership with US embassy public diplomacy and media development efforts to employ bots for prosocial activities in developing countries, such as promoting public health and education initiatives or providing information during crisis events, such as natural disasters.

The cat-and-mouse between bot programmers and bot hunters will continue to be a **technological arms race**, with the following implications for SOF:

- Even if governments try to regulate the use of botnets, the enforcement of such regulations would require successful and unambiguous detection. This means that the space may yet exist for actors such as SOF to employ botnets even if a trend toward

government attempts to restrict that space move ahead. In addition, SOF may consider investing in artificial intelligence (AI)/machine learning (ML) technologies to improve the quality of social media bots that they may be allowed to employ eventually.

- The increased sophistication of bots could also pose challenges to the ability to detect their use. SOF may need also to invest in improving technologies to detect the use of social media bots and botnets to play defense.
- The ability of US adversaries to employ social media bots to distract US forces (including SOF) in a military deception (MILDEC) campaign will be enhanced by further improvements in the quality of bots themselves, unless the US can keep up in terms of technologies to detect their use. The potential force protection impacts of this increased MILDEC capability are likely to be felt most acutely by SOF, who typically operate as small teams in contested environments.
- US military forces (including SOF) could also deploy botnets for MILDEC against US adversaries, given the requisite resources and authorities to do so.

As better technology becomes more accessible, there is likely to be **an increase in active users** of social media bots, with the following implications for SOF:

- As access to high-quality bots spreads, it may become increasingly difficult to detect bots and their activities. The democratization of bot use is another reason for SOF to consider investments in bot detection technologies.
- A global increase in social media bot and botnet use will likely mean a corresponding increase in the difficulty of countering bot campaigns. This challenge could also be exploited by US government actors (including SOF), who also might be able to hide their activities in the noise.
- An increase in the ease of use and proliferation of social media botnets could eventually result in the creation of a global social media indications and warning (I&W) network for SOF activities (e.g., botnets that could look for indicators of SOF activity and immediately amplify them for the sake of exposure). To combat such a possibility, SOF may consider investing in technology to create force protection bots—botnets that might employ flooding or fracturing techniques to counter the amplification of SOF activity indicators. The US could also seek to create its own amplification botnets for adversary activities of various kinds.

Opportunities and risks for SOF

Social media bots can be incredibly useful for SOF. To begin, this tool has three key features that make it an especially powerful means of influencing social media conversations and public opinion:

- **Reduced need for cultural expertise:** SOF could flood individuals or hashtags and fracture online conversations without deep cultural knowledge.
- **Rapidly deployable capability:** Social media bots could offer SOF an exceptionally quick response to a detected threat.
- **Wide range of applications:** Using the taxonomy we developed, it is not difficult to identify instances in which SOF might use social media bots and botnets, as seen in the table below.

Social media bots and opportunities for SOF

Bot Activities	Examples of SOF use
Distributing	SOF might partner with a US embassy to share information about a natural disaster overtly
Amplifying	SOF might partner with a US embassy to amplify information about a public health campaign overtly
Distorting	SOF might covertly attempt to increase discord by posting culturally relevant incendiary content in a foreign-language social media campaign with US national security implications
Hijacking	SOF might covertly or overtly respond to an identified adversary disinformation campaign by hijacking the relevant hashtag and turning the campaign into something innocuous or prosocial
Flooding	SOF might covertly overwhelm a social media account that threatened an ongoing direct action by live-tweeting details of the event
Fracturing	SOF might covertly break a social media campaign into multiple parts to diffuse the impact of messaging that threatens US forces by revealing their locations

Source: CNA.

There is, of course, a reputational risk inherent in deploying social media bots. In some cases, SOF might decide that potential backlash against their use of bots may not be worth the benefit. Just as malicious “Russian bots” are the focus of social media disinformation operations in the US, “US bots” (if their origins were identified) could be framed as malicious activity against the backdrop of perceived US hegemony and cultural aggression. Yet there is no question that there are other instances in which the risks attendant to the use of social media bots or botnets

might be outweighed by the benefits. As one example, flooding a new hashtag with irrelevant information to bury information about the location of an ongoing SOF operation might be worth the potential backlash.

As this report makes clear, although the future of social media bots is largely unknowable, the paths forward are ripe with both challenges and opportunities. Social media bots are relatively simple mechanisms that have a tremendous capacity to influence discourse; if deployed correctly (e.g., in accordance with US values and under appropriate authorities and oversight), they could be a powerful tool for SOF actors and the US government more broadly.

Contents

Introduction.....	1
Key questions.....	3
Approach.....	3
Organization.....	4
Caveats.....	5
What Is a Bot? What Is a Botnet?	7
Social media bots.....	10
Social media botnets	15
The Threat of Automation	18
Why write about social media bots?.....	18
Identifying Bots and Botnets	22
Bot-like activity	24
Trolls.....	27
Taxonomy of Bot and Botnet Activity	28
Scope of the taxonomy	28
Neutral, prosocial, and malicious activity	30
Social Media Bot and Botnet Taxonomy	32
Distributing: sharing content	32
Amplifying: increasing the reach of a message or user	38
Distorting: changing the balance of a conversation	41
Hijacking: taking over a conversation.....	43
Flooding: overwhelming a conversation or account	45
Fracturing: breaking a large conversation into smaller conversations.....	51
Implications of Future Trends	55
Regulations and authorities.....	55
Activity in developing countries	57
A technological arms race.....	59
An increase in active users	60
Opportunities and Risks for SOF	63
Reduced need for cultural expertise	63
Rapidly deployable capability.....	63
Wide range of applications.....	65
Reputational risk.....	65

Conclusion	67
Appendix A: The Legal Landscape and Platform Policies	68
The legal landscape	68
Platform policies.....	69
Figures	73
Abbreviations	75
References	76

Introduction

Social media bots—simply put, automated programs on social media platforms—are, from a national security perspective, an interesting and underexplored internet phenomenon. CNA initiated this study of social media bots as part of a series of studies on technology and the internet and the implications for SOF.¹ This report focuses on the challenges and opportunities that social media bots offer to US national security professionals, specifically SOF.

To understand the implications of social media bots for SOF, this report addresses a handful of specific elements. We define social media bots, explain how they work, and explore why people use them. Social media bots are tools that can be harnessed to spread disinformation or programmed to share accurate information. Rather than focusing on who deployed the bots (i.e., motivation) or what information the bots are sharing (i.e., content), this report focuses on the tasks that these bots can be programmed to perform in the context of social media discourses (i.e., objective).

In focusing on objectives, this report makes space to explore instances in which social media bots do precisely the opposite of what we have come to expect them to do. That is, we emphasize the fact that few of these activities are inherently problematic. A social media bot programmed to distribute information can do so maliciously (e.g., distributing disinformation on COVID-19 vaccines) or prosocially (e.g., sharing accurate information about the spread of COVID-19). Our goal in emphasizing the proximal objectives of social media bots (versus the distal objectives of, for example, sowing social discord) is to highlight the influence that social media bots can have on social media discourse and the impressive utility of social media bots.

This report differs from existing literature in five key ways. First, as mentioned above, the defining feature of this report is that *it is about the objective that the social media bot pursues and not about the motivation of the programmer or the content that it posts*. This is not a report about disinformation or misinformation; these topics have been covered widely in the wake of the 2016 US election. Instead, we have generated a list of potentially disruptive tactics that we

¹ For other research in this series, see: Vera Zakem, Megan McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaigns*, CNA, 2018, IV. https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf; Megan McBride and Zack Gold with contributions by Jonathan Schroden and Lauren Frey, *Cryptocurrency: Implications for Special Operations Forces*, CNA, 2018. https://www.cna.org/CNA_files/PDF/CRM-2019-U-020186-Final.pdf.

can expect to see deployed in US and allied social media ecosystems. By focusing on social media bots as technological tools, we hope to derive a broader potential applicability for national security professionals.

Second, *this report is an analysis of social media bots that interact with individual users on social media platforms*. Internet bots make up a significant amount of web traffic, but they are also “considerably different than automated social media accounts.”² Automated activity, as one report noted, is “an infrastructural element of search engines and other features of the modern World Wide Web.”³ Of course, there are neutral, prosocial, and malicious internet bots, but we have circumscribed our analysis to social media bots (i.e., those operating on social media platforms).

Third, *this report does not approach social media bots from a technical perspective*. We are not concerned primarily with the identification or disruption of social media bots. Our goal is to inform SOF and other national security professionals of the types of activities in which social media bots might engage. We chose this focus to improve the capacity of US actors to identify such efforts; to alert US actors to the types of online conversations vulnerable to interference, so that they can be more closely monitored; and to contribute to the toolkit of potential approaches that US actors might deploy (openly or covertly) in online spaces.

Fourth, *this report emphasizes that social media bots are unbiased tools*. When possible, we have included examples of social media bots engaged in prosocial or positive activities. We did so in part to highlight the potential utility of bots for overt actors: for example, a US embassy might deploy a bot that tweets information about earthquake activity in a country where access to reliable information is inconsistent. We did this also to emphasize the unbiased definition of automated social media activity. In public discourse, social media bots are recognized mostly because of their deployment by nefarious actors. However, it is important to avoid contributing to a rhetorical shift that labels all social media bots as nefarious.

Fifth, *this report is intended for national security readers, with a particular focus on SOF*. As with our 2018 report on memetic warfare, we hope this work on social media bots will inform US government (USG) elements engaged in information activities—from the Department of State’s Global Engagement Center to JMWC. With that mandate, we examine a toolkit that SOF can both protect against and deploy for their own purposes. In our exploration of potential futures, we outline implications, opportunities, and risks of these futures for SOF (although many of these have broader implications for the USG as well).

² Robert Gorwa and Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, arXiv, 5–6.

³ *Ibid.*

Key questions

We aim our discussion at national security audiences generally (e.g., policy-makers and staffers in the executive and legislative branches) and specifically at Department of Defense (DOD) entities most likely to be involved potentially with social media bots. Based on their information capabilities and authorities to operate in the information space, SOF are that entity. This report addresses the following questions:

- What are social media bots? What are social media botnets? How can we identify social media bots and botnets?
- What is the range of activities for which social media bots/botnets can be deployed? Where, how, and why are they being used—both legitimately and illegitimately?
- What are the legal restrictions on—and authorities relevant to—the deployment of social media bots or botnets?
- What are the likely next evolutions that can be anticipated in this space? What are the implications of these changes for SOF?
- What are the opportunities and risks that social media bots/botnets represent in the national security space? How might SOF deploy social media bots productively? How might SOF anticipate, identify, and respond to the use of social media bots by others?

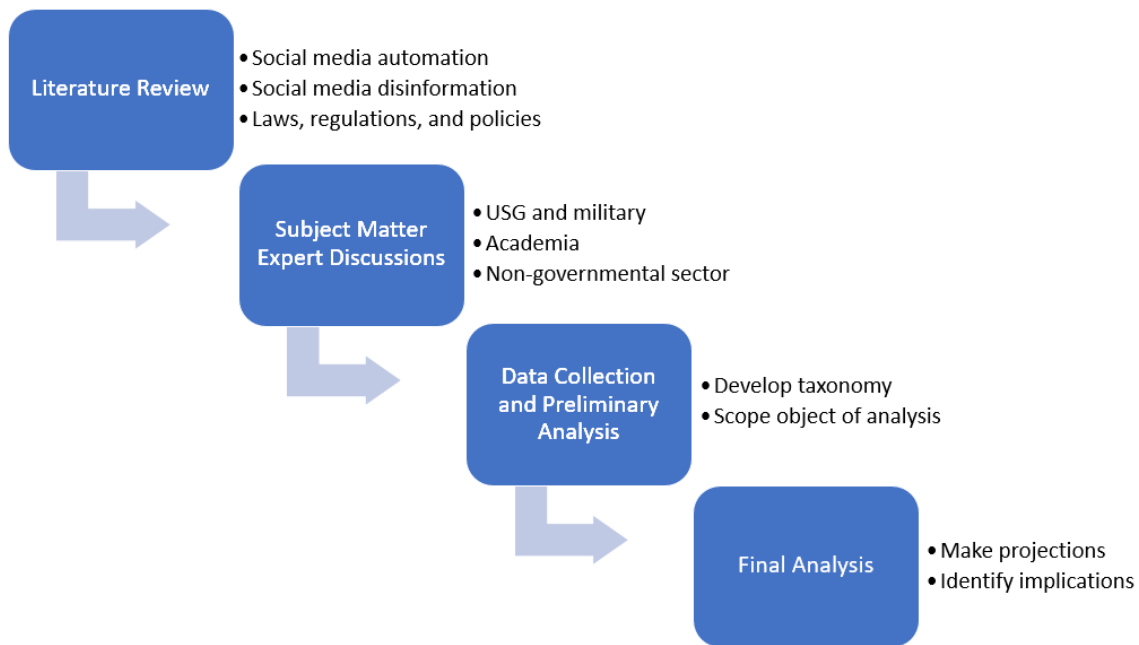
Approach

To address these questions, we began with a literature review of academic works on social media automation and detection; policy papers on “computational propaganda,” or social media disinformation operations; and the legal, regulatory, and private-sector approaches to social media automation. Concurrent to this research, we collected and tracked examples of social media bot and botnet activity to understand the range of activities in which these tools have engaged.

After we completed our initial literature review, we had discussions with SMEs in the USG and military, private-sector, academia, and non-governmental sectors. Following our SME interviews, our team worked simultaneously to (1) develop a taxonomy capturing the full range of social media bot and botnet activity that we had collected over the preceding months and (2) clearly articulate the types of activity that would be our object of analysis. This was an iterative process, as the discovery of new examples prompted a reconsideration of scoping, while issues raised in the scoping prompted the search for new types of examples.

Finally, with a fuller understanding of social media bots gained from the previous sections and our SME discussions, we identified likely future evolutions and explored the implications, opportunities, and risks that social media bots and botnets represent for SOF and the USG.

Figure 2. Research approach



Source: CNA.

Organization

We organized this paper around the research questions outlined above. The first section provides an understanding of bots, generally, and social media bots and botnets more specifically. The second section highlights why national security professionals, and especially SOF, should care about bots by unwrapping the threat of automation and the role of social media bots as tools of disinformation. The third section acknowledges the difficulty in identifying bots and botnets.

Readers well versed in these issues can proceed directly to the fourth section: a taxonomy of bot and botnet activity. In this section, we introduce a taxonomy that defines six activities of social media bots and botnets, which includes numerous malicious, neutral, and prosocial examples of each activity.

In the fifth section, we identify likely evolutions in the near- to mid-term futures (which our SME discussions were particularly helpful in exploring) and the implications of those futures for SOF. The sixth and final section outlines the opportunities and risks for SOF, and includes a version of our taxonomy that includes examples of potential SOF use in each of the six social media bot and botnet activities that we identified. The paper then ends with a short conclusion.

We also completed a full review of the legal and regulatory landscape. A brief summary of this work can be found in Appendix A: The Legal Landscape and Platform Policies a more detailed assessment can be found in our companion paper, *Social Media Bots: Laws, Regulations, and Platform Policies*.

Caveats

This is the third in a series of reports that CNA has completed on technology, the internet, and implications for SOF. It was by far the most conceptually difficult, for the following three reasons.⁴

- First, there is no question that in each of these reports (on memes, cryptocurrency, and social media bots) we struggled with issues of definition and scope. These terms, which refer to evolving technologies and the boundaries establishing what should and should not be included, are still fluid. In this case, these issues were compounded by the fact that it is nearly impossible to state with certainty that a social media account is a bot. This fact left us in the undesirable position of wanting to write a report about the activities of social media bots, activities that might be performed by social media bots, activities that we thought were performed by social media bots, and activities that could have been performed by social media bots. Although we have tried to be clear in what we have included, we apologize for what might seem like an unnecessarily complicated effort to scope the taxonomy.
- Second, the websites of the social media platforms themselves were far less helpful than we anticipated in their articulation of relevant policies. In most cases, we confronted the dual challenges of platform websites that (a) posted their policies and policy changes ad hoc in a variety of places including blogs, community standards, developer policies, and newsrooms, and (b) had no clear policies on social media bots themselves, but instead a patchwork set of relevant policies on a variety of related issues (e.g., automation, fake

⁴ For other research in this series see: Vera Zakem, Megan McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaigns*, CNA, 2018, IV, https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf; Megan McBride and Zack Gold with contributions by Jonathan Schroden and Lauren Frey, *Cryptocurrency: Implications for Special Operations Forces*, CNA, 2018, https://www.cna.org/CNA_files/PDF/CRM-2019-U-020186-Final.pdf.

accounts). Although much of our research on regulations and policies can be found in our companion paper, *Social Media Bots: Laws, Regulations, and Platform Policies*, even the brief summary in this report required far more research than we anticipated.

- Third, it was very difficult to separate the concept of social media bots—which we firmly believed could be explored in a neutral capacity as a potential tool for SOF—from (a) the mountain of literature on disinformation and (b) the imprecise literature on social media trolls. Although the threat posed by disinformation campaigns is obviously of critical importance, the potential utility of social media bots has been lost in the wake of this conversation. Maintaining focus on social media bots as a tool was thus difficult as we were consistently extracting data from articles that began with the threat of disinformation. Similarly, there is a problematic imprecision in some of the writing on social media bots that blurs the line between bots and trolls. Articles with promising headlines suggesting relevant anecdotes about the behavior of social media bots turned out, on more than one occasion, to be articles about the activities of online trolls.

What Is a Bot? What Is a Botnet?

As mentioned in the introduction, bots operate in nearly all sectors of the internet and in a wide variety of roles. Many of these bots, though, are completing tasks that could best be categorized as routine internet maintenance. To be clear about the object of analysis in this report, it is important both to define “bots” and to scope the types of bots relevant to this study.

The terminology associated with bots can be overwhelming and confusing given the frequency with which conversations about bots, social media bots, and internet trolls overlap. Compounding this disadvantage is that many of the terms used in analysis of these topics are effectively internet slang (e.g., astroturfing, sockpuppets, crowd turfers, sybils). We have defined these terms in a short glossary on bot slang (see Figure 3). However, the core terminology for discussing bots is far less complicated: “bot” is short for “robot”; a “botnet” is a network of bots; and a “botmaster” is an individual who controls a botnet.

Bots are most essentially defined as “automated programs designed to perform a specific task.”⁵ Like the robots of science fiction, bots assist human efforts by performing menial or routine work.⁶ Once programmed, they are able to carry out activities that would be time-consuming for humans.

The function that a bot serves, or the task that it performs, is dictated by conditions set by the bot’s programmers. Most bots are “rule-based,” which means they are programmed to follow instructions that define and limit their behavior. Generally speaking, these programs follow an “if ... then” pattern. For example, a rule-based bot on a shopping website could be programmed so that *if* a consumer adds laundry detergent to his basket, *then* the bot would recommend he also buy fabric softener.

⁵ Paris Martineau, “What Is a Bot?” *Wired*, Nov. 16, 2018. <https://www.wired.com/story/the-know-it-alls-what-is-a-bot/>.

⁶ “What Are Software Bots?” *ThinkAutomation*. Accessed Sept. 8, 2020. <https://www.thinkautomation.com/bots-and-ai/what-are-software-bots/>.

Figure 3. Bot slang glossary



ASTROTURF

A technique used to manufacture an artificial online presence for an interest group. The name is derived as a contrast to an authentic “grassroots” campaign.

BOT

Automated programs that perform tasks online.

BOTMASTER

An individual or group that controls a botnet; this botnet might be coordinated to promote content, to alter the site’s trends, etc.

BOTNETS

A network of bots that perform a botmaster’s commands. Individual accounts in the botnet often appear to be bona fide.

BOT-LIKE ACTIVITY

Activity that displays key characteristics of social media bots, particularly on the Twitter platform.

CLICK FARM

A service that allows an account to boost its profile by purchasing clicks (e.g., likes and followers). The fake followers are typically humans using multiple accounts and IP addresses.

CLICKWORKERS

Human agents who manually complete many of the same functions typically associated with bots. Unlike bots, clickworkers’ activities are usually undetected by anti-spam and bot identification programs..

CROWDTURFERS

Human agents who take part in astroturfing campaigns, typically related to promoting a product with positive reviews and comments.

CYBORGS

Hybrid bot-human actors used to defeat bot identification systems. This term may refer to bots that require human inputs to function, or bots that are designed to supplement human activity. Feature-based detection methods struggle to detect these hybrid accounts.

FALSE AMPLIFIERS

Networks of fabricated accounts deployed to affect public opinion by driving political discussions.

HEADLESS BROWSING

An automated browsing tool with no user interface, legitimately used to test server environments. Illicit “headless browser farms” mimic internet users and artificially inflate ad traffic, which increases the ad’s revenue.

POLITICAL BOTS

Social media bots used specifically to supplement political campaigns and pursue political objectives.

SMOKESCREEN

Deliberate misdirection to a topic that is unrelated to a target issue. Effective smokescreens require significant coordination among bots.

SOCKPUPPET

A false persona developed to engage with bona fide users on social media.

SYBILS

A term that can reference an entity controlling a series of false nodes on a platform or to the nodes themselves.

TROLL

Actors who sabotage online chats with inflammatory remarks or images. Often specifically used to denote political sockpuppets deployed by government affiliates.

TWITTER BOMB

Coordinated activity among multiple Twitter accounts to use hashtags and keywords to achieve more views of an idea or a product.

Source: CNA, based on Robert Gorwa and Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, *arXiv*; Emilio Ferrara et al., “The Rise of Social Bots,” *Communications of the ACM* 59 (7), 96-104, 2016; Jen Weedon et al., *Information Operations and Facebook*, Facebook, Apr. 27, 2017; and John P. Mello Jr., “Headless Web Traffic Threatens Internet Economy,” *E-Commerce Times*, Mar. 25, 2014.

Bots are increasingly programmed ML—a subset of the field called AI.⁷ ML bots start their tasks following programmed rules, but over time their activities will adapt based on the data they collect. For example, a shopping website’s ML bot will analyze all purchases and recognize that a significant percentage of shoppers who buy detergent also buy cat food; the bot will then adjust to recommend cat food as an additional purchase when a shopper puts detergent in his basket.

Many of these bots function behind the scenes. For example, web crawlers are a type of bot that trawls the internet for data; these bots are critical to the functionality of the internet and they perform tasks that would take humans so much time as to be impossible. Similar bots are essential for services such as search engines and social media platforms.

As mentioned above, bots are unbiased actors. They can be programmed to perform routine maintenance, but they can also be programmed to perform less constructive or innocuous tasks. A bot may be programmed to scrape proprietary data from a competitor’s website.⁸ Or an ML bot may be programmed to scrape personal data to compose an effective spear phishing email.⁹

Moreover, the impact of a bot can be amplified when combined with others in a botnet (i.e., a network of bots working in coordination).¹⁰ Much like a single bot, a botnet can perform neutral work to keep the internet running, prosocial work that performs a service, or malicious work that furthers a nefarious objective. A botmaster,¹¹ the person or entity in control of the botnet, might use a botnet to attack a website to overload its server (known as a distributed

⁷ Andrew Ilachinski, *AI, Robots, and Swarms: Issues, Questions, and Recommended Studies*, CNA, Jan. 2017, DRM-2017-U-014796-Final. While machine learning is sometimes referred to as artificial intelligence (AI), it is important to note that the two are not equivalent. Machine learning is a subset of AI; thus, all machine learning is AI, but not all AI is machine learning.

⁸ Distil Networks, “Bad Bot Report 2019: The Bot Arms Race Continues,” Blue Cube Scurity, PDF file. Accessed Sept. 8, 2020, <https://www.bluecubesecurity.com/wp-content/uploads/bad-bot-report-2019LR.pdf>.

⁹ Danny Palmer, “Hackers Don’t Just Want Your Credit Cards, Now They Want the Pattern of Your Life,” *ZDNet*, Apr. 16, 2016, <https://www.zdnet.com/article/hackers-dont-just-want-your-credit-cards-now-they-want-the-pattern-of-your-life/>.

¹⁰ Johnathan Azaria, “The Challenges of DIY Botnet Detection – and How to Overcome Them,” *Imperva* (blog), Feb. 4, 2019, <https://www.imperva.com/blog/the-challenges-of-diy-botnet-detection-and-how-to-overcome-them/>.

¹¹ Eiman Alothali, N. Zaki, E.A. Mohamed, and H. Alashwal, “Detecting Social Bots on Twitter: A Literature Review,” *2018 International Conference on Innovations in Information Technology (IIT)*, Al Ain, UAE, 2018, pp. 175–180, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8605995&tag=1>.

denial-of-service, or DDoS, attack).¹² On a social media platform, a botmaster may coordinate bot accounts to post content repeatedly to spread disinformation or to trick the platform—and its users—into assuming a topic is popular or trending.¹³

Social media bots

The examples above illustrate that bots can perform a wide variety of (neutral, prosocial, and malicious) tasks. This paper, however, is not concerned with the maintenance of the internet's infrastructure, automated data scraping, or DDoS attacks (although there are national security implications to such activity).¹⁴ Our focus is on bots that operate on social media platforms.

Broadly speaking, two types of bots operate on social media platforms. The first type is bots that operate *behind the scenes*. In some cases, social media companies themselves deploy these bots to support and monitor operations of their social media networks. This type of bot is effectively a maintenance bot and works for the social media network. In other cases, people other than social media companies (i.e., third parties) deploy these bots to monitor data and collected information. Some of these bots might have neutral agendas (e.g., collecting information for a university research project) and some might have a more nefarious agenda (e.g., collecting data to support a disinformation campaign).

This report focuses not on the bots that operate *behind the scenes* but those that *interact with humans*.¹⁵ Some of these accounts are self-identified as bots. These are easy to find, and they perform all sorts of interesting and amusing functions. On Twitter, @censusAmericans uses 2009–2013 US Census data to “[tweet] the census one real american [sic] at a time.”

¹² Steve Weisman, “What Is a Distributed Denial of Service Attack (DDoS) and What Can You Do About Them?” NortonLifeLock (blog), Norton, modified Jul. 6, 2020. Accessed Sept. 8, 2020, <https://us.norton.com/internetsecurity-emerging-threats-what-is-a-ddos-attack-30sectech-by-norton.html>.

¹³ Kai-Cheng Yang, et al., “Arming the Public with Artificial Intelligence to Counter Social Bots,” *Human Behavior and Emerging Technologies 1.1*, Feb. 6, 2019, arXiv:1901.00912v2 [cs.CY].

¹⁴ Joseph Cox, “This Bot Tweets Photos and Names of People Who Bought ‘Drugs’ on Venmo,” Motherboard Tech by Vice (blog), Jul. 19, 2018. https://www.vice.com/en_us/article/qvmkvx/twitter-bot-venmo-buying-drugs-photo-names.

¹⁵ Some literature (e.g., Gorwa and Guilbeault, 2018) differentiates between “social bots” (two words), which are social media bots as we have defined them, and “socialbots” (one word), which are inherently nefarious social bots. We use the term “social media bots” throughout this paper.

Figure 4. @censusAmericans Twitter bot



Source: <http://www.twitter.com/censusAmericans>.

Figure 5. Weather Bot Twitter feed

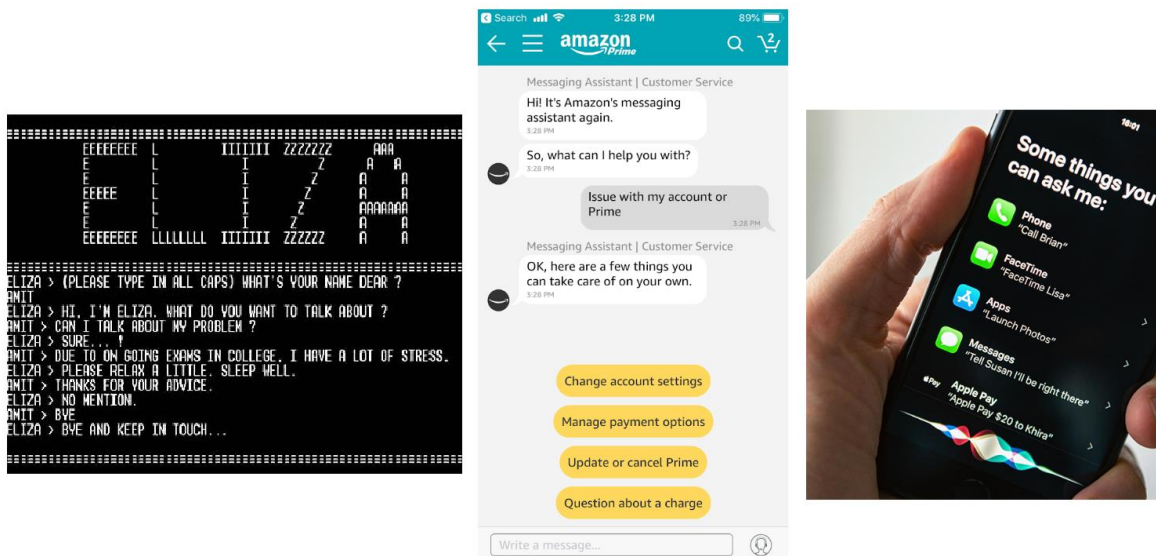


Source: <http://www.twitter.com/tempextremes>.

A similar, though perhaps more useful, self-identified bot that generates and posts content is Weather Bot (@tempextremes). This account posts the highest and lowest temperatures in the US each day (see Figure 5).¹⁶

Another type of bot that is often self-identified is a chatbot. Chatbots are “programs that approximate human speech and interact with humans directly through some sort of interface”; these bots are the foundation for social media bots that “converse.”¹⁷ One of the first chatbots, developed in the 1960s, was the rule-based ELIZA. Today we are more familiar with far more sophisticated programs, such as customer service chatbots that attempt to solve our problems and the ML chatbots like Apple’s Siri.

Figure 6. Chatbots



Sources: Siraj Abbas, “History and Future of Chatbots,” Medium (blog), Apr. 25, 2017; Ajit Ghuman, “Chatbots for Customer Service: Why You Need to Rethink Your Strategy,” *helpshift* (blog), May 8, 2019; Marc Saltzman, “Simple Tips for Mastering Apple’s Siri and Other Digital Voice-Enabled Assistants,” AARP, Oct. 21, 2019.

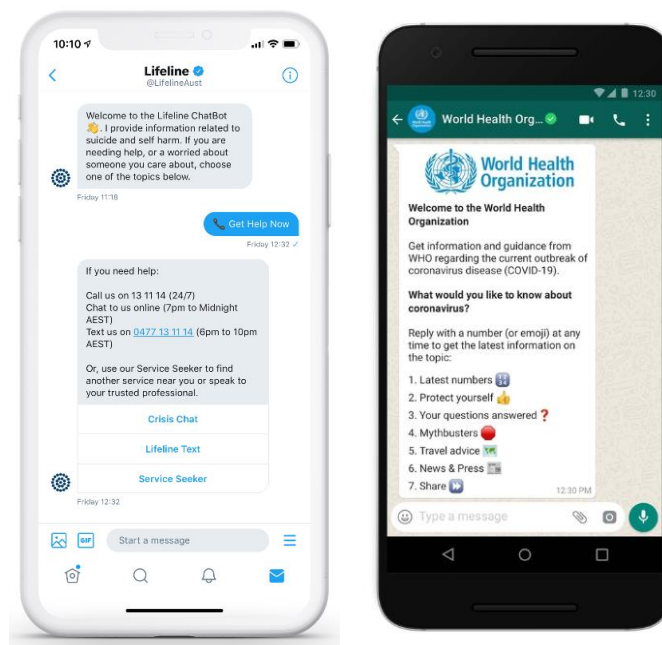
A social media chatbot is different from these examples because it exists on a social media platform. One example is @LifelineAust, an approved Twitter bot designed for individuals,

¹⁶ Weather Bot (@tempextremes). Accessed Sept. 8, 2020. <https://twitter.com/tempextremes>.

¹⁷ Robert Gorwa and Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, arXiv.

family members, and friends of those struggling with depression, self-harm, and suicide. After sending a Direct Message to the account, the user is prompted to select from three options, including “Get Help Now.”¹⁸ More recently, the World Health Organization released a chatbot on WhatsApp to “to answer questions from the public about coronavirus, providing reliable information worldwide to as many as 2 billion people.”¹⁹

Figure 7. Social media chatbots



Sources: Kara Hinesley, “Lifeline’s New Twitter DM Chatbot Helps Friends and Family #BeALifeline,” Twitter Blog, Oct. 17, 2018 and Kelsey Warner, “WHO Rolls Out WhatsApp Chatbot To Answer Questions and Debunk Myths about Coronavirus,” *The National*, Mar. 25, 2020.

All of these bots function in the open and none are attempting to masquerade as human beings. As we discuss later in Appendix A: The Legal Landscape and Platform Policies some social

¹⁸ Kara Hinesley, “Lifeline’s New Twitter DM Chatbot Helps Friends and Family #BeALifeline,” Twitter Blog, Oct. 17, 2018, https://blog.twitter.com/en_au/topics/company/2018/Lifeline-launches-Twitter-DM-chatbot-to-help-BeALifeline.html.

¹⁹ Kelsey Warner, “WHO Rolls Out WhatsApp Chatbot To Answer Questions and Debunk Myths about Coronavirus,” *The National*, Mar. 25, 2020, <https://www.thenational.ae/uae/health/who-rolls-out-whatsapp-chatbot-to-answer-questions-and-debunk-myths-about-coronavirus-1.997415>.

media platforms, specifically Twitter, actively support some degree of user automation and thus these social media bots are able to function in the open.

However, this is not the case for all social media bots. In some instances, social media bots masquerade as human users. This type of bot is an instance of:

Automation software that controls an account on a particular [social media network], and has the ability to perform basic activities such as posting a message and sending a connection request...[and] that it is designed to be stealthy, that is, it is able to pass itself off as a human being.²⁰

In other words, a bot in this category is a program “that automatically produces content and *interacts with humans on social media, trying to emulate and possibly alter their behavior*” (emphasis added).²¹

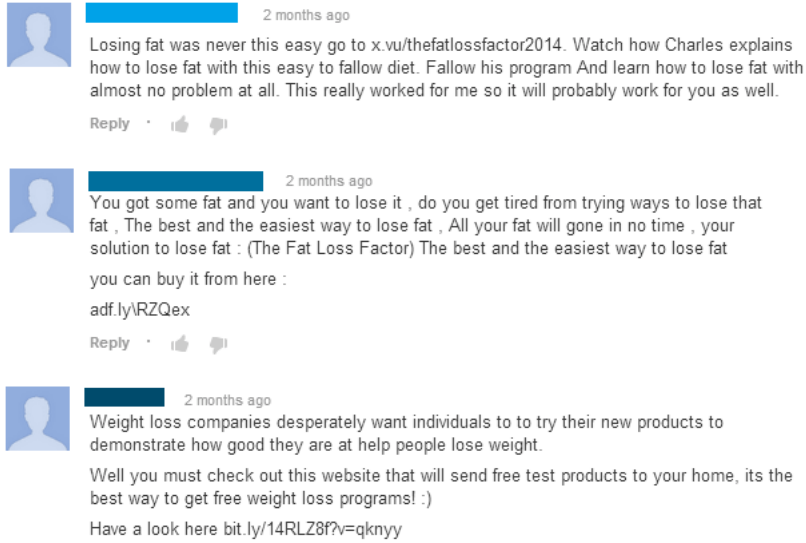
Sometimes this activity is simply annoying. For example, we are all familiar with spambots that inundate our inboxes with advertisements and malware.²² A social media spambot engages in this type of activity on social media networks. On YouTube, for example, a social media spambot might be used to advertise a product or service (typically by prompting users to click on a link).

²⁰ Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov and Matei Ripeanu, “The Socialbot Network: When Bots Socialize for Fame and Money,” in *Proceedings of the 27th Annual Computer Security Applications Conference (ACSAC’11)*, Dec. 2011, <http://lersse-dl.ece.ubc.ca/record/272>: - LERSSE-RefConfPaper-2011-008.

²¹ Emilio Ferrara, Onor Varol, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini, “The Rise of Social Bots,” *Communications of the ACM*, 59 (7), 96–104, Jun. 2016. Available at SSRN. <https://ssrn.com/abstract=2982515>.

²² Robert Gorwa and Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, arXiv.

Figure 8. Social media spambot activity



Source: Martin Brinkmann, "YouTube's New Commenting System Aims to Push Google+, Nothing More," ghacks.net, Sept. 25, 2013.

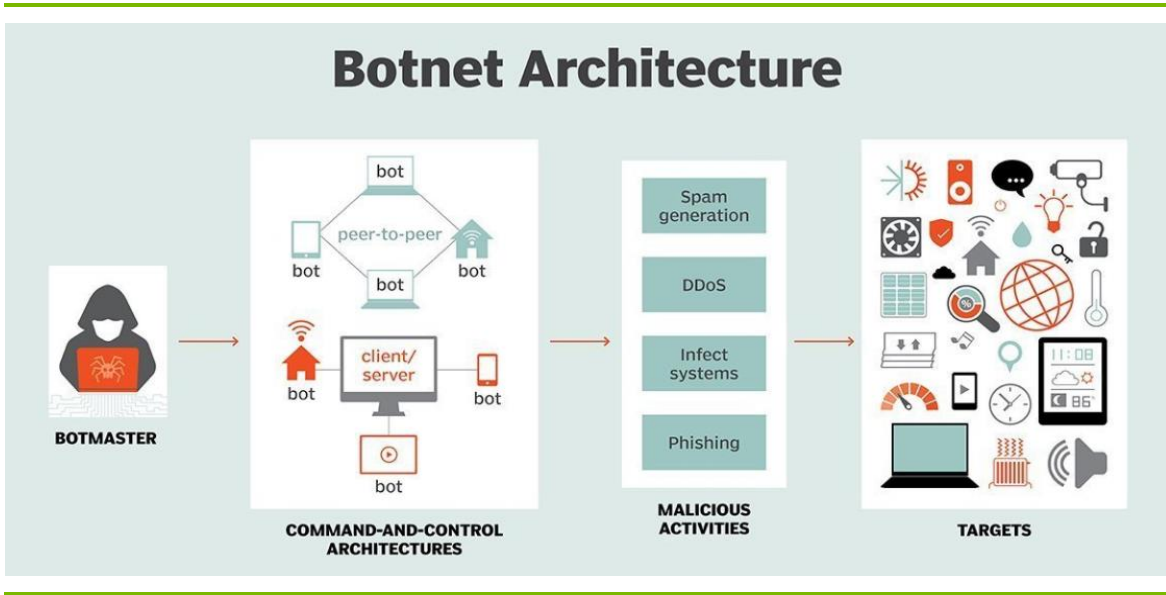
In other instances, a bot will engage in more subtle behaviors, such as those observed during the 2016 US presidential election or the UK Brexit vote.

Social media botnets

In all the cases above, a single bot was acting in isolation. As on the wider web, though, botmasters can deploy social media botnets for "coordinated activities" on social media networks.²³ The range of such activities is considerable, and will be explored in greater detail later in the section titled Taxonomy of Bot and Botnet Activity, but a sense of the botnet's architecture can be seen in the figure below (though the examples of malicious activity are not specific to social media).

²³ Onur Varol, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini, "Online Human-Bot Interactions: Detection, Estimation, and Characterization," in *Proceedings of the 11th International Conference on Web and Social Media (ICWSM '17)*, last revised Mar. 27, 2017, arXiv:1703.03107v2 [cs.SI], <https://arxiv.org/pdf/1703.03107.pdf>.

Figure 9. Botnet architecture



Source: Krishna Gupta, "Define Botnets and Their Types?" The Tech Win (blog), Dec. 13, 2018.

The nodes of a social media botnet are sometimes called "sybils," following the computer security terminology of a botmaster controlling multiple "personalities."²⁴ Although a single social media bot has the potential to influence the public discourse, particular concern should be paid to coordinated social media bots that operate at scale.

For example, multiple accounts controlled by a single entity can quickly generate posts and make specific content trend or amplify misinformation. They can trick humans and engagement-based ranking algorithms alike, creating the appearance that some person or opinion is popular.²⁵

²⁴ Onur Varol, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini, "Online Human-Bot Interactions: Detection, Estimation, and Characterization," in *Proceedings of the 11th International Conference on Web and Social Media (ICWSM '17)*, last revised Mar. 27, 2017, arXiv:1703.03107v2 [cs.SI], <https://arxiv.org/pdf/1703.03107.pdf>. The term derives from the 1973 book—later a television movie—about the treatment of "Sybil Dorsett" for what at the time was called "multiple personality disorder." See: <https://www.geeksforgeeks.org/sybil-attack>.

²⁵ Kai-Cheng Yang et al., "Arming the Public with Artificial Intelligence to Counter Social Bots," *Human Behavior and Emerging Technologies* 1.1, Feb. 6, 2019, arXiv:1901.00912v2 [cs.CY].

By feigning a popular trend through coordinated efforts, botnets can not only drive human public opinion on a social media platform but also set the agenda among mainstream media and society.²⁶

²⁶ Alice Marwick and Rebecca Lewis, “Media Manipulation and Disinformation Online,” Data & Society Research Institute, *datasociety.net*, May 15, 2017, https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf; and Clint Watts, “Testimony on Extremist Content and Russian Disinformation Online: Working with Tech to Find Solutions before the U.S. Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism,” Oct. 31, 2017, <https://www.judiciary.senate.gov/imo/media/doc/10-31-17%20Watts%20Testimony.pdf>.

The Threat of Automation

Why write about social media bots?

In 2019, a *Washington Post* article on influence and disinformation in social media spaces included the disturbing, but perhaps unsurprising, quote that there is “never something that’s trending that’s not in some way promoted by bots.”²⁷ That might seem like a melodramatic or exaggerated claim, but the ubiquity of bots—in conversations of grave importance and staggering triviality—cannot be responsibly ignored by those concerned with national security.

Social media bots have been observed participating in—and sometimes interfering with—social discourse on a wide range of subjects: the 2016 US presidential election, the UK’s decision to leave the European Union (EU), the COVID-19 crisis, discussions about vaccine safety, debates about climate change science,²⁸ and the 2019 season of *The Voice Kids* (a Russian talent show).²⁹ Although some of these topics elicit more concern than others, even cases that appear trivial merit serious examination when they are successful. For example, a technique that works in the context of a televised talent contest can be easily adapted to a televised presidential debate.

Bots have been active in political discussions for nearly a decade. As one report noted, “one of the first political uses of social bots was during the 2010 Massachusetts special Senate election in the United States, where a small network of automated accounts was used to launch a Twitter smear campaign against one of the candidates.”³⁰ In the wake of these events—but particularly following the 2016 US presidential election—hundreds of articles and reports have been written about the influence of bots online. Bots have also been long active in online

²⁷ Elyse Samuels and Monica Akhtar, “Are ‘Bots’ Manipulating Conversation? Here’s What’s Changed Since 2016,” *Washington Post*, Nov. 20, 2019. <https://www.washingtonpost.com/politics/2019/11/20/are-bots-manipulating-conversation-heres-whats-changed-since/>.

²⁸ Thomas Marlow, Sean Miller, and J. Timmons Roberts, “Twitter Discourses on Climate Change: Exploring Topics and the Presence of Bots,” *SocArXiv*, Feb. 26, 2020, doi:10.31235/osf.io/h6ktm.

²⁹ “Russian Bots Rigged Voice Kids TV Talent Show Result,” *BBC News Europe*, May 16, 2019. <https://www.bbc.com/news/world-europe-48293196>.

³⁰ Eni Mustafaraj and P. Takis Metaxas, “From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search,” in *Proceedings of The World Congress on Engineering and Computer Science (WCECS 2010)*, 2010. <https://repository.wellesley.edu/object/ir122>.

discussions about medicine and science. For example, in 2015, the US Defense Advanced Research Projects Agency's (DARPA's) Bot Challenge asked participants to identify "influence bots" in an online discourse on vaccines.³¹ Thus, the current crisis of bots as they pertain to politics and COVID-19 is simply a continuation of a long-standing pattern.

As mentioned above, social media bots are not inherently nefarious. They are also not inherently effective. In isolation, it is unlikely the hundreds of millions of real social media users will be influenced by a single bot account's activity—and it is even less likely that a single bot will make a meaningful impression. That is not to suggest that a single bot cannot be influential. However, social media bots are worrying primarily because they are tools of a nefarious activity known as "computational propaganda."³²

According to researchers at the University of Oxford, "Computational propaganda is the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks."³³ In 2014, the World Economic Forum identified "the rapid spread of misinformation online" as a key global trend.³⁴ This trend includes disinformation and distorted information, and often social media bots are among the means by which nefarious online actors spread such disinformation.

In May 2019, when Senator Richard Burr filed a bill to authorize US intelligence activities for fiscal years 2018–2020, he included a section (Section 404) that presented findings related to Russian efforts to "deploy information warfare operations against the United States, its allies and partners, with the goal of advancing the strategic interests of the Russian Federation":

(2) One line of effort deployed as part of these information warfare operations is the weaponization of social media platforms with the goals of intensifying societal tensions, undermining trust in governmental institutions within the United States, its allies and partners in the West, and generally sowing division, fear, and confusion.

³¹ David A. Broniatowski et al., "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate," *American Journal of Public Health* 108, no. 10 (Oct. 1, 2018): 1378–1384. <https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2018.304567>.

³² For more on the broader subject, see the Computational Propaganda Project at the Oxford Internet Institute. <https://navigator.oii.ox.ac.uk/>.

³³ Samuel C. Woolley and Philip N. Howard, "Computational Propaganda Worldwide: Executive Summary," Working Paper No. 2017.11, Oxford, UK: Project on Computational Propaganda. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf>.

³⁴ Farida Vis, "The Rapid Spread of Misinformation Online," *Outlook on the Global Agenda 2014*: 28a–29b, The Network of Global Agenda Councils. <http://reports.weforum.org/outlook-14/top-ten-trends-category-page/10-the-rapid-spread-of-misinformation-online/>.

(3) These information warfare operations are a threat to the national security of the United States and that of the allies and partners of the United States. As Director of National Intelligence Dan Coats stated, “These actions are persistent, they are pervasive and they are meant to undermine America’s democracy.”

(4) These information warfare operations continue to evolve and increase in sophistication.

(5) Other foreign adversaries and hostile non-state actors will increasingly adopt similar tactics of deploying information warfare operations against the West.³⁵

Compounding the concern of bots and other computational propaganda in the social media space is the fact that conversations over social media can spread into print and television media. Reporters and columnists for US news outlets are active on social media, and they often turn to social media posts to present the “person on the street” perspective for stories. One study of the online reporting of 33 US news outlets found that 32 of them cited Twitter accounts controlled by the Kremlin-linked Internet Research Agency.³⁶

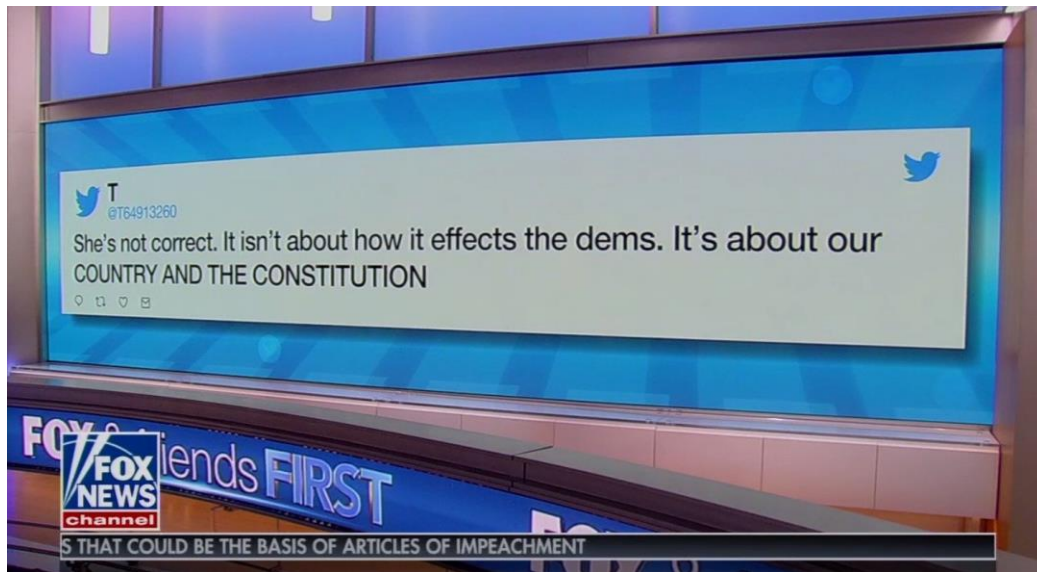
Social media bots can also reach broadcast audiences, where—as in print and online media—social media messages may be cited as evidence of support for particular opinions. In Figure 10, we display a tweet, broadcast during a cable news show, that appears to have characteristics of a bot account.³⁷

³⁵ Damon Paul Nelson and Matthew Young Pollard Intelligence Authorization Act for Fiscal Years 2018, 2019, and 2020, S. 1589, 116th Cong., 1st sess., May 22, 2019.

³⁶ Lukito, Josephone, and Chris Wells, “Most Major Outlets Have Used Russian Tweets as Sources for Partisan Opinion: Study,” *Columbia Journalism Review*, Mar. 8, 2018. <https://www.cjr.org/analysis/tweets-russia-news.php>.

³⁷ The assessment that the account is a bot is based on a few characteristics noted by an NPR journalist at the time it appeared on TV (retrospective analysis is not possible as the account has been suspended). The journalist did not characterize the account as a bot, but pointed out that the account had a suspicious name consisting of a single letter and a string of eight numbers; that there was no name or biographical information; that it had been created approximately a month before its appearance on TV and yet had tweeted hundreds of times; and that it had just three followers.

Figure 10. Tweet displayed during cable news broadcast



Source: Meridith McGraw, Twitter post, 6:59AM, Nov. 22, 2019,
<https://twitter.com/meridithmcgraw/status/1197847059539382272>.

In other words, we care about social media bots because of their use as tools of computational propaganda and the affect they can have on public discourse—on social media platforms and on society more broadly.

Identifying Bots and Botnets

Establishing which social media accounts are bots is effectively an impossible task. Experts have identified a number of key characteristics to aid in the identification of bots, but none is conclusive.³⁸

The Atlantic Council's Digital Forensic Research Lab (DFRL), for example, has suggested that one "clue" that might aid in identifying bots is the number of posts that the account produces.³⁹ Some individuals, however, post hundreds of times a day. In 2018, for example, a "70-year-old grandmother [who spent] up to 14 hours a day tweeting the praises of President Trump and his political allies" had her account temporarily frozen as part of Twitter's effort to identify and stop bots.⁴⁰

Another organization, First Draft, compiled a list of more than two dozen bot indicators (see Figure 11). As First Draft emphasized, any attempt to identify a bot should consider these indicators in tandem—the group says 10 or more of these characteristics make it likely the account is a bot.⁴¹

³⁸ See, for example, Roberts, Siobhan, "Who's a Bot? Who's Not?" *New York Times*, Jun. 16, 2020. <https://www.nytimes.com/2020/06/16/science/social-media-bots-kazemi.html>; and "#BotSpot: Twelve Ways to Spot a Bot," Medium (blog), August 28, 2017. <https://medium.com/dfrlab/botspot-twelve-ways-to-spot-a-bot-aedc7d9c110c>.

³⁹ Digital Forensic Research Lab, "#BotSpot: Twelve Ways to Spot a Bot," Medium (blog), Aug. 28, 2017. <https://medium.com/dfrlab/botspot-twelve-ways-to-spot-a-bot-aedc7d9c110c>.

⁴⁰ Burnett, Sara, "Crackdown on 'Bots' Sweeps Up People Who Tweet Often," *AP News*, Aug. 4, 2018. <https://apnews.com/06efed5ede4d461fb2eac5b2c89e3c11>.

⁴¹ Carlotta Dotto and Seb Cubbon. "How to Spot A Bot (or Not): The Main Indicators of Online Automation, Co-Ordination and Inauthentic Activity," *First Draft News*, Nov. 28, 2019. <https://firstdraftnews.org/latest/how-to-spot-a-bot-or-not-the-main-indicators-of-online-automation-co-ordination-and-inauthentic-activity/>.

Figure 11. Indicators to identify a bot account on social media

HOW TO SPOT A BOT (OR NOT)

Just because it acts like a bot doesn't mean it is a bot. These indicators of automated or co-ordinated online activity can help, but look for a combination of signs, not just one.

Account

- Recent creation date
- Lack of personal information
- Profile photo is ambiguous, stolen or nonexistent
- Divisive words, hashtags, URLs or emojis in bio
- Suspicious handle e.g. lots of numbers

Activity

- High volume of tweets (more than 100/day)
- High percentage of retweets (more than 80%)
- Posting persistently day and night
- Posting only at specific times of day
- Sudden spike in activity or change in interests

Content

- Tweeting in more than one language
- Engaging in multiple international narratives
- Signs of automation or account management software like buff.ly
- Posting inflammatory memes and GIFs
- Hashtag spamming
- Occasional off-brand retweets
- Very few reliable news sources
- Awkward turns of phrase

Network

- Followers and following is high and almost identical
- High number of following and no followers
- Following a suspicious mix of sources
- Connected to other suspicious accounts
- Duplicated account
- Previously circulating suspicious content
- Previously identified by other organisations as suspicious

FIRSTDRAFT

Source: Carlotta Dotto and Seb Cubbon. “How to Spot A Bot (or Not): The Main Indicators of Online Automation, Co-Ordination and Inauthentic Activity,” *First Draft News*, Nov. 28, 2019.

There are also automated online resources that can run accounts through an algorithm to determine the likelihood that the account is a bot (i.e., bots hunting bots). The most well known is Botometer, which tests Twitter accounts.⁴² Again, these techniques are imperfect. Twitter itself criticized these automated bot identifiers for using a methodology that is “an extremely limited approach.”⁴³ Further, as with human deduction, automated bot-hunters can be wrong. One study pointed out the potential for false positives (i.e., humans tagged as bots) and false negatives (i.e., bots that fool the algorithm) of such efforts.⁴⁴

⁴² Indiana University Observatory on Social Media, “Botometer.” <https://botometer.iuni.iu.edu/#/>.

⁴³ Yoel Roth and Nick Pickles, “Bot or Not? The Facts About Platform Manipulation on Twitter,” Twitter Blog, May 18, 2020. https://blog.twitter.com/en_us/topics/company/2020/bot-or-not.html.

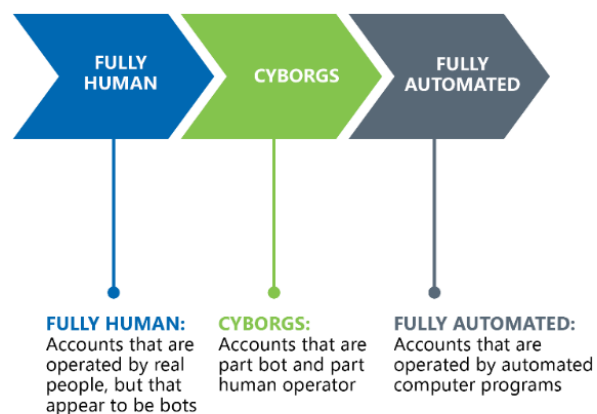
⁴⁴ Adrian Rauchfleisch and Jonas Kaiser, “The False Positive Problem of Automatic Bot Detection in Social Science Research,” *Berkman Klein Center Research Publication No. 2020-3*, Mar. 2020, Available at SSRN: <https://ssrn.com/abstract=3565233>.

The difficulty in identifying bots has drawn the attention of the US national security community, including DARPA’s “bot competition” to identify automated accounts tweeting about vaccinations in 2014.⁴⁵ It is appropriate that the US military took an interest in detecting bots, as one group of academics described as “a never-ending arms race” the parallel loop of the design of sophisticated automated accounts and the detection of those accounts.⁴⁶

Bot-like activity

As highlighted above, our study takes into account “bot-like” activity (i.e., activity that displays key characteristics of social media bots regardless of whether this has been proven definitively) precisely because of the challenges in identifying bots.

Figure 12. Spectrum of social media accounts



Source: CNA.

⁴⁵ Emerging Technology from the arXiv, “How DARPA Took On the Twitter Bot Menace with One Hand Behind Its Back,” *MIT Technology Review*, Jan. 28, 2016. <https://www.technologyreview.com/s/546256/how-darpa-took-on-the-twitter-bot-menace-with-one-hand-behind-its-back/>.

⁴⁶ Onur Varol, Emilio Ferrara, Clayton A. Davis, Filippo Menczer and Alessandro Flammini, “Online Human-Bot Interactions: Detection, Estimation, and Characterization,” in *Proceedings of the 11th International Conference on Web and Social Media (ICWSM '17)*, last revised Mar. 27, 2017, arXiv:1703.03107v2 [cs.SI]. <https://arxiv.org/pdf/1703.03107.pdf>; and Emilio Ferrara, Onur Varol, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini, “The Rise of Social Bots,” *Communications of the ACM*, 59 (7), 96-104, Jun. 2016. Available at SSRN: <https://ssrn.com/abstract=2982515>.

By including bot-like activity, we have captured the full spectrum of social media activity that we are interested in: activity ranging from the fully human to the completely automated (see Figure 12). Bot-like activity falls into two sub-categories: fully human activity and cyborg activity.

Bot-like activity that is fully human includes accounts run by human beings that simply *appear* to be social media bots. In some cases, as with the woman mentioned above, this appearance is unintentional. In other instances, the human user is purposefully masquerading as a bot. In one example, human-run accounts embedded bits of code in their tweets to create the impression that they were bots to undermine British Prime Minister Boris Johnson's credibility.⁴⁷

Bot-like activity that is part bot and part human is sometimes described as "cyborg" activity.⁴⁸ By including cyborg activity, we avoid the academic debate over how much human intervention is permissible before something stops being a bot (e.g., is it a bot if humans intervene 10 percent of the time? 20 percent? 80 percent?). Currently, neither human nor automated detection methods are able to successfully identify a bot account in which there is human intervention (i.e., a cyborg account).⁴⁹ The ability to fool bot-detection systems with a little human assistance has created an industry of "cyborg farms," where humans sit around waiting for bots to be blocked with the question "Are you a human?"—at which point the humans confirm their humanness in order for the bot to continue operating (see Figure 13).⁵⁰

⁴⁷ Joey D'Urso. "The Real People Pretending to Be 'Boris Bots' on Facebook," *BBC News*, Nov. 1, 2019, <https://www.bbc.com/news/blogs-trending-50218615>.

⁴⁸ Robert Gorwa and Douglas Guilbeault. *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, arXiv.

⁴⁹ Kai-Cheng Yang et al. "Arming the Public with Artificial Intelligence to Counter Social Bots," *Human Behavior and Emerging Technologies* 1.1, Feb. 6, 2019, arXiv:1901.00912v2 [cs.CY].

⁵⁰ Steven Puddephatt. "Bots x Humans: a Solution Is Needed," *Infosecurity Magazine* (web), Oct. 11, 2019, <https://www.infosecurity-magazine.com/opinions/bots-humans-solution/>.

Figure 13. Example of a cyborg click farm



Source: Ben Makuch, "We Talked to Phone Farmers Who Use Ad Fraud to Earn Beer Money," *Motherboard Tech by Vice* (blog), Aug. 5, 2019.

Industry actors have largely denied that they are unable to detect automated account activity. In 2017, the acting general counsel of Twitter testified to the Senate Judiciary Committee:

Our systems are built to detect automated and spam accounts across their lifecycles, including detection at the account creation and login phase and detection based on unusual activity (e.g., patterns of Tweets, likes, and follows). Our ability to detect such activity on our platform is bolstered by internal, manual reviews conducted by Twitter employees.⁵¹

However, the evidence belies these claims. As one 2018 literature review noted, bots "have become increasingly sophisticated in their designs and capabilities to avoid social bot detection techniques."⁵² This is true for techniques designed both inside and outside social media companies.

⁵¹ Sean J. Edgett, "Testimony before the U.S. Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism," Oct. 31, 2017, <https://www.judiciary.senate.gov/imo/media/doc/10-31-17%20Edgett%20Testimony.pdf>.

⁵² Eiman Alothali, N. Zaki, E. A. Mohamed and H. Alashwal, "Detecting Social Bots on Twitter: A Literature Review," *2018 International Conference on Innovations in Information Technology (IIT)*, Al Ain, UAE, 2018, pp. 175-180, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8605995&tag=1>.

Trolls

Although we have tried to avoid most of the slang that characterizes writing on social media bots, one term—troll—requires some attention. The line between bots and trolls is blurry because the two are frequently linked. It is likely that language on this issue is imprecise partly because trolls are antisocial actors, and most people assume that social media bots are also antisocial actors. It is also possible that this language is blurry because the imprecision is politically useful. As a 2018 article on social media bots noted, “regimes around the world have already begun to label dissidents as ‘bots’ or ‘trolls’” thus further compromising the integrity of online discourse and creating political cover under which to demand that social media networks remove these “false accounts” (i.e., individuals “that have espoused anti-government views”).⁵³ As the authors of the report noted, though, in most cases these words refer to very different things.

Social media bots are automated social media accounts that perform tasks online; trolls are social media users that sabotage online chats with inflammatory remarks or images. This means that a bot (and a bot-like account) can be a troll. In other words, an account that is automated (and an account that appears to be automated) can engage in inflammatory behavior that attempts to sabotage online discussion. It does not mean, though, that all bots (or all bot-like accounts) are trolls.

Although it is increasingly difficult to identify social media bots confidently, the activities of these accounts offer challenges and opportunities. To discuss the implications of social media bots and botnets for US SOF and national security professionals, in the next section we describe the activities that they can conduct.

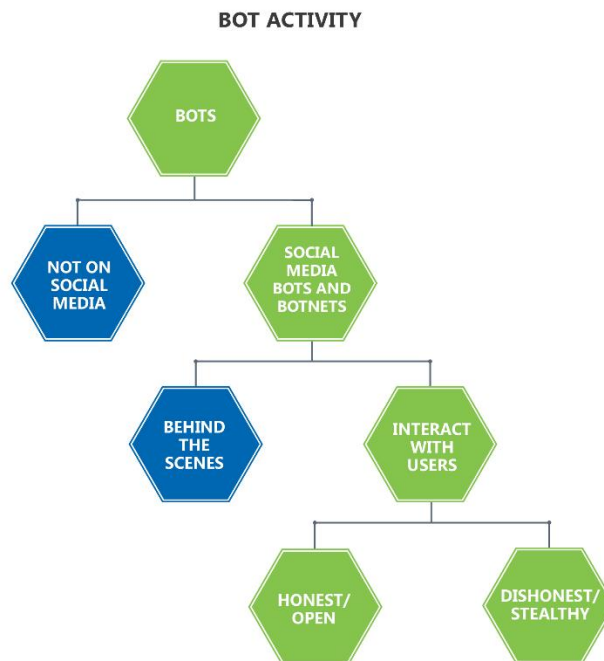
⁵³ Robert Gorwa and Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, Jul. 28, 2018, arXiv, 15.

Taxonomy of Bot and Botnet Activity

Scope of the taxonomy

This report takes as its central concern the activity of (a) social media bots and botnets that (b) interact with humans in ways that are (c) both overt and covert/ clandestine.

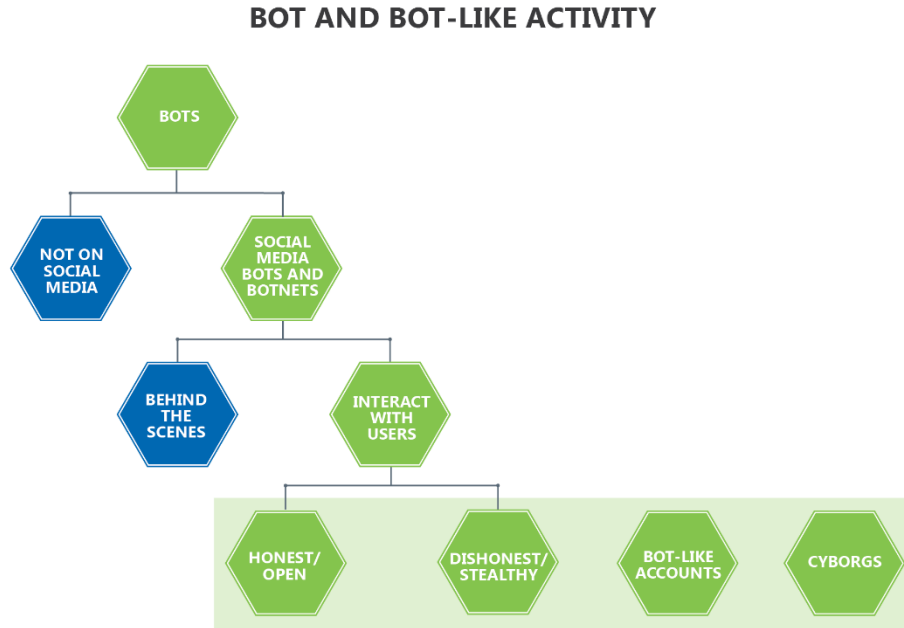
Figure 14. Types of bot activity



Source: CNA.

The report also includes “bot-like activity” in its analysis. As we discussed in the previous section (Identifying Bots and Botnets), it is not always possible to determine whether an account belongs to an authentic user or a social media bot. As a result, we are including not only fully automated social media bot and botnet activity but also a range of bot-like activities that are likely only partly automated (and, in the case of actors intentionally pretending to be bots or inadvertently identified as bots, are fully human but appear to be bots).

Figure 15. Activity under analysis in this report



Source: CNA.

In a few instances, we have chosen to include examples in which the activity was undertaken by humans but *could have been* executed by social media bots. We did so because our aim in creating this taxonomy was to be not only descriptive (i.e., outlining the scope of the threat) but also prescriptive (i.e., identifying potential uses that may not yet have been fully exploited). The central concern of this report is the potential influence that social media bots and botnets wield—regardless of whether, in the particular examples we provide, it was actually bots, cyborgs, or humans taking action.

Finally, while many of the examples in this report come from Twitter, bots are not a phenomenon unique to Twitter. The activities highlighted in this report are relevant for a large number of social media networks including Twitter, Facebook, Reddit, WhatsApp, and YouTube. Where possible, we have included examples from other platforms. The emphasis on Twitter in the following pages is attributable to the fact that the activities of bots on Twitter are to some degree easier to identify, resulting in more mainstream media coverage and more information for analysts.

Neutral, prosocial, and malicious activity

As mentioned in the introduction, social media bots are tools that can be harnessed to spread disinformation or programmed to share accurate information. These tools offer both challenges and opportunities to those working on national security issues. Rather than focusing on the motivation behind a bot or the content it is sharing, in this section we emphasize the objectives that social media bots can be programmed to perform.

During our analysis, we identified six primary actions that a social media bot or botnet could be programmed to do: distributing, amplifying, distorting, hijacking, flooding, and fracturing. In executing these tasks, though, a bot or botnet might be engaged in activity that could be characterized as neutral, prosocial, or malicious. In some cases, making the assessment that a bot is neutral, prosocial, or malicious is simple. A bot that shares the time can be easily described as neutral; a bot offering help to someone posting about suicide can be easily described as prosocial; and a bot spreading disinformation about COVID-19 can be easily described as malicious. The reality, though, is that a judgment call is being made even in these easy cases.

The Twitter bot @the_nationalbot self-describes as a bot that “posts lyrics from [the band] the national every hour” (see Figure 16).⁵⁴ It is tempting to describe this activity as neutral; on the other hand, fans of the band might characterize this bot as prosocial because it offers a pleasant break from the monotony of the day. Conversely, those who do not like the band might describe the bot as an annoying daily nuisance. Similarly, the Twitter bot @OEFTracker, otherwise known as “Is the US still at war in Afghanistan?” tweets a single word once a day: “Yes.” This bot could be viewed as neutral in that it simply provides a piece of factual information; it could be seen as prosocial in that it reminds users that the US is still keeping the homeland safe by fighting terrorists overseas; or it could be seen as malicious in that it undermines the US government by consistently reminding people that the US is still embroiled in a costly and seemingly endless war in Afghanistan.⁵⁵

⁵⁴ the national bot (@the_nationalbot). Accessed Sept. 8, 2020. https://twitter.com/the_nationalbot.

⁵⁵ Is the US still at war in Afghanistan? (@OEF tracker). Accessed Sept. 14, 2020. <https://twitter.com/OEFTracker>.

Figure 16. Images of the national bot and OEFTracker



Source: https://twitter.com/the_nationalbot and <https://twitter.com/OEFTracker>.

In the taxonomy below, we have attempted to include neutral, prosocial, and malicious examples. To be clear about the examples we chose, we have included a graphic highlighting what we have included and we have labeled each individual example. It is important to acknowledge, though, that we have made judgment calls in assigning these labels. Although some bots are very obviously one of these—i.e., neutral, prosocial, or malicious—there is a wide universe of bot and botnet activity that is difficult (if not impossible) to classify.

Finally, some bots are capable of performing more than one function in the taxonomy. For example, bots that generate and post spam are distributing messages, but they can also hijack trending topics or flood a hashtag. We classified bots in the taxonomy based on the primary purpose for which they were programmed.

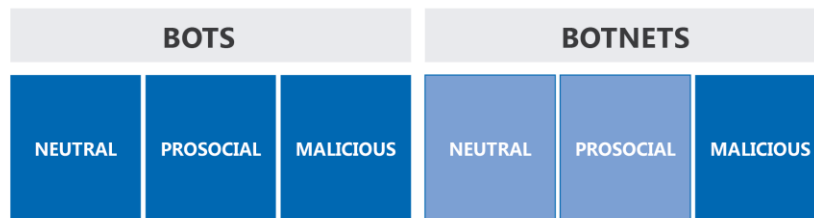
Social Media Bot and Botnet Taxonomy

Distributing: sharing content

A number of bots and botnets simply distribute content. In some cases, this might be described as content creation as many of these bots “create” their own content and exist to disseminate it. A bot in this category might, for example, retweet another account (creating content via the retweet) or post information from a public database (creating content by sharing the data). Although the use of a bot or botnet to distribute content is often an effort to amplify a specific message, what makes this effort a distributive one in our taxonomy is that the primary purpose of the bots/botnet is to distribute its own unique content. By contrast, in the amplification category, the primary purpose of the bot or botnet is to increase the reach of content that comes from somewhere else.

Social media bots and botnets are useful for distributing messages because they can create and push content much more quickly and consistently than humans can, which allows the bot or botnet to reach a much larger audience. In the case of prosocial bots, such as @ParityBOT featured below, the ability to rapidly identify the messages it is targeting, and generate response tweets, is almost instantaneous. Humans would need to search for these messages and manually tweet responses, which would likely reduce the impact significantly. Similarly, for malicious bots, the ability to produce content at extremely high speeds increases the number of individuals they can target and, in the case of spambots, the chances someone will click on their links.

Figure 17. Types of examples included in this section

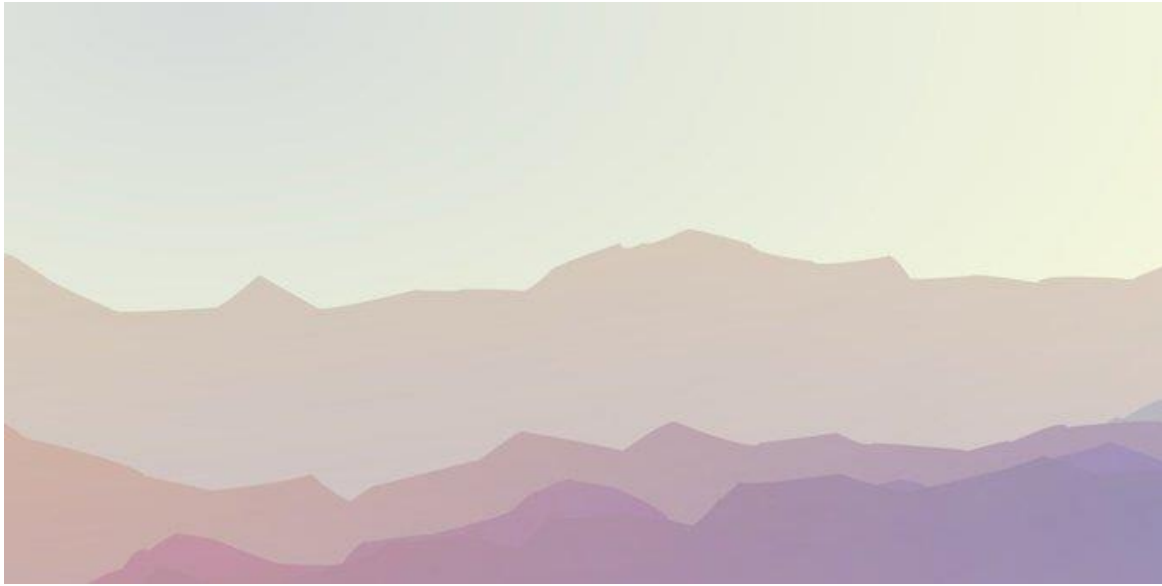


Source: CNA.

Neutral examples—bots

The Twitter bot @softlandscapes generates an original landscape in muted colors every six hours without accompanying text.⁵⁶

Figure 18. Example of a landscape generated by the bot @softlandscapes, a neutral example of a distribution bot



Source: <https://twitter.com/softlandscapes/status/1293259977612812288>.

Another example that fits into the category of neutral distribution is that of Twitter account @mothgenerator.⁵⁷ This account creates and tweets images of fake species of moths, while assigning each an official-sounding name. Interestingly, this account might be described as one creating and distributing benign disinformation.

⁵⁶ soft landscapes (@softlandscapes). Accessed Sept. 8, 2020. <https://twitter.com/softlandscapes>.

⁵⁷ moth generator (@mothgenerator). Accessed Sept. 14, 2020. <https://twitter.com/mothgenerator>.

Figure 19. Example of tweet generated by the bot @mothgenerator



Source: <https://twitter.com/mothgenerator>.

Prosocial example—bots

The Twitter account @ParityBOT posts a positive tweet every time its algorithm “detects an abusive tweet directed at a woman in politics.” Examples of these tweets include motivational messages, such as “Women in politics, YOU ARE KILLING IT,” and facts about women in politics, such as “1948: The right to vote is extended to Asian women.”⁵⁸

⁵⁸ ParityBOT (@ParityBOT). Accessed Sept. 8, 2020. <https://twitter.com/ParityBOT>.

Figure 20. The Twitter biography of @ParityBOT, a prosocial example of a distribution bot



Source: <https://twitter.com/ParityBOT>.

Malicious example—bots

Spambots often generate and post malicious links, typically at high intervals and possibly concealing damaging malware. In Figure 21, the spambot @asians_cute posts the same link multiple times in a row, promising to take the user to pornography websites, while potentially concealing malicious code in the links.⁵⁹

⁵⁹ Richard J. Oentaryo et al., "On Profiling Bots in Social Media," in *Proceedings of Social Informatics: 8th International Conference, SocInfo 2016*, Bellevue, WA, 2016, Research Collection School Of Information Systems. Available at: https://ink.library.smu.edu.sg/sis_research/3648.

Figure 21. Example of a spam bot spreading potentially malicious links



Source: Richard J. Oentaryo et al., "On Profiling Bots in Social Media," in *Proceedings of Social Informatics: 8th International Conference, SocInfo 2016*, Bellevue, WA, 2016, Research Collection School Of Information Systems.

Malicious example—botnets

On September 11, 2014, a large number of Twitter accounts whirled into motion, sending out photos, videos, and text about an explosion at the Columbian Chemicals plant in St. Mary Parish, Louisiana. The tweets used a variety of hashtags (such as #LouisianaExplosion and #ColumbianChemicalsInNewOrleans) before eventually coming together to use the single hashtag #ColumbianChemicals. The accounts put out photos depicting flames and plumes of black smoke rising over the site, photoshopped screenshots of the CNN homepage showing the story, and videos of ISIS allegedly taking blame for the attack. They also tagged high-profile journalists and individuals from around the country in their posts to try to gain attention for the story. As it turned out, the entire attack (and all the supporting evidence) was a fabrication. Researchers subsequently discovered that many of those posting were part of a botnet, which

distributed relevant content at about one tweet per second. This is an example of a distributive botnet because the bots posted unique content and media on the topic.⁶⁰

Figure 22. Twitter account posting a fake CNN homepage showing the story of the Columbian Chemicals plant explosion



Source: John Borthwick, "Media Hacking," Medium (blog), March 7, 2015.

⁶⁰ Adrian Chen, "The Agency," *New York Times*, Jun. 2, 2015, <https://www.nytimes.com/2015/06/07/magazine/the-agency.html>; and Borthwick, John, "Media Hacking," Medium (blog), March 7, 2015, <https://web.archive.org/web/20150309221009/https://medium.com/in-beta/media-hacking-3b1e350d619c>.

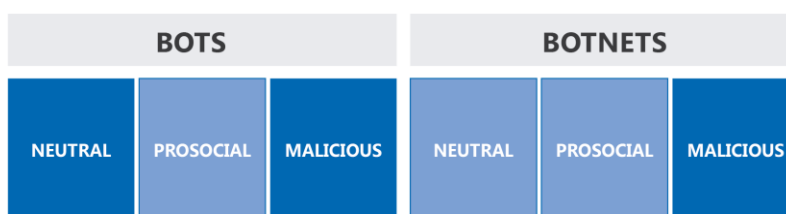
Amplifying: increasing the reach of a message or user

Bots and botnets are also useful for amplifying a particular message, individual, or group. As noted above, this activity is different from that of distribution because the primary goal is not to create and post new content but to increase the reach of content that someone else has created and posted (effectively making posts appear more popular than they would be without this artificial support). Although amplification bots also sometimes generate new content, the distribution of their own message is not their primary purpose; they are generating content to amplify a pre-existing message.

Botnets are particularly useful for amplification, because they can retweet relevant messages at much higher rates than humans can. This speed vastly increases both the number of people who might see the post and the likelihood that it will trend.

Social media users who want to appear more famous or popular sometimes hire botnets, which often use identities stolen from real people, to amplify their profiles and boost their views. The bots can like, retweet, and repost the users' content depending on the platform. As a result of this increased engagement, real humans are more likely to view the user as important, which can lead to further amplification of their content. Social media algorithms may, as one example, become more likely to push that user's content to the forefront, making it show up more often to more people.⁶¹

Figure 23. Types of examples included in this section



Source: CNA.

⁶¹ Nicholas Confessore et al., "The Follower Factory," *New York Times*, Jan. 27, 2018, <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>.

Neutral example—bots

According to a 2018 *The New York Times* story, a company called Devumi has made millions of dollars by selling fake bot followers and engagements to celebrities, companies, and individuals, pulling from its base of more than 3.5 million bot accounts, which are each sold many times. Overall, Devumi has provided its clients with over 200 million fake Twitter followers, as well as plays on YouTube, clicks on the music site SoundCloud, and more.⁶²

Malicious example—bots

A study from February 2020 showed that about a quarter of tweets skeptical about climate change on any given day are produced by bots, which amplify the messages of climate change deniers and make up a high percentage of tweets related to “fake science” and Exxon. By contrast, only five percent of tweets promoting actions to reverse climate change were bots.⁶³

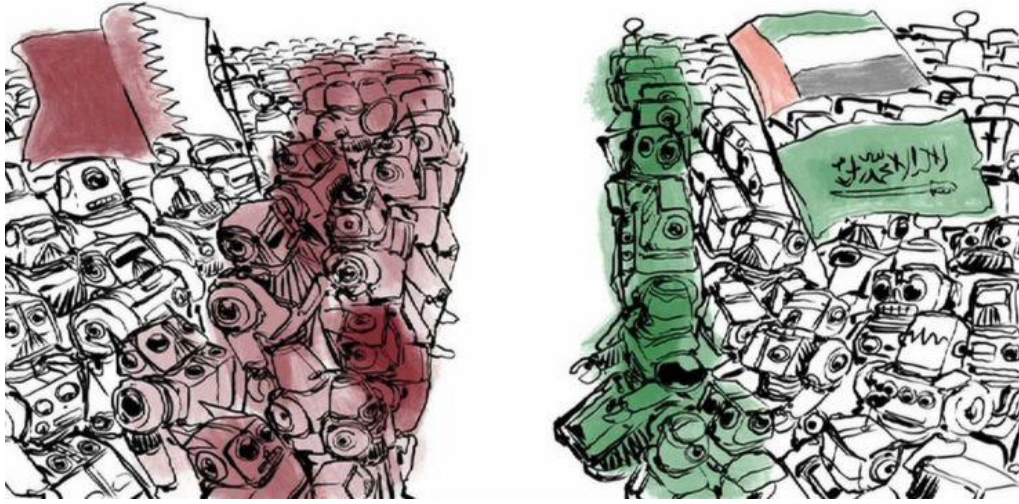
Malicious example—botnets

On June 6, 2017, four Gulf countries (Saudi Arabia, the UAE, Bahrain, and Egypt) severed diplomatic ties with the neighboring state of Qatar over allegations Qatar was sponsoring terrorism and attempting to destabilize the region. Two months before the crisis began, a network of social media bots began sending out a high volume of tweets containing disinformation and promoting specific hashtags. These bots, with biographies sporting anti-Qatar messages, laid the groundwork for the diplomatic dispute and occupied a central place in the online discourse on the blockade that eventually occurred. Years after the crisis began, the bot accounts had amplified the reach of certain political messages, including some anti-Qatar accounts and some accounts associated with the Saudi royal family. Pro-Qatar bots also began to appear, albeit at a lesser quantity, and amplified the accounts of some Qatari royal family members and prominent businessmen. The use of bots on both sides was so pronounced that a 2018 BBC article published a cartoon, seen in Figure 24, depicting bots lined up for battle, with those of Qatar on one side and those of Saudi Arabia and the UAE on the other.

⁶² Ibid.

⁶³ Oliver Milman, “Revealed: Quarter of All Tweets About Climate Crisis Produced by Bots,” *The Guardian*, February 21, 2020, <https://www.theguardian.com/technology/2020/feb/21/climate-tweets-twitter-bots-analysis>; Yarno Ritzen, “The Fake Twitter Accounts Influencing the Gulf Crisis,” *Al-Jazeera*, Jul. 21, 2019, <https://www.aljazeera.com/news/2019/07/fake-twitter-accounts-influencing-gulf-crisis-190717052607770.html>; and Owen Pinnell, “The Online War Between Qatar and Saudi Arabia,” *BBC*, Jun. 3, 2018, <https://www.bbc.com/news/blogs-trending-44294826>.

Figure 24. BBC cartoon of bot use by the Gulf countries in their diplomatic dispute



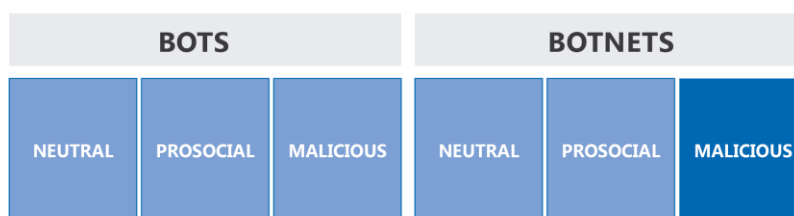
Source: Owen Pinnell, "The Online War Between Qatar and Saudi Arabia," *BBC*, Jun. 3.

Distorting: changing the balance of a conversation

Distortion is known as one of the four “Ds” of Russian disinformation tactics (i.e., distort the situation, dismiss critics, distract from the larger issue, and dismay listeners).⁶⁴ The use of distortion can feed into an adversary’s objectives by allowing them to cast doubt on facts and introduce other (likely false) sub-narratives that cause chaos and make the true story seem unknowable.⁶⁵

Distortion is functionally similar to distribution, because the bots and botnets distorting a conversation also tend to generate their own content. However, distortion is distinct from distribution in that its primary purpose is not simply to disseminate a message for the sake of the message itself but to disseminate a message to achieve increased dissonance. Botnets are particularly useful for this purpose, as the use of a single bot is unlikely to achieve the desired degree of chaos.

Figure 25. Types of examples included in this section



Source: CNA.

Malicious examples—botnet

Following the 2018 disappearance (and, ultimately, the murder) of journalist Jamal Khashoggi, a botnet began pushing coordinated pro-Saudi tweets “imploing users to express doubt about news stories that Khashoggi was killed at the Saudi consulate in Turkey on October 2 at the

⁶⁴ Ben Nimmo, “Anatomy of an Info-War: How Russia’s Propaganda Machine Works, and How to Counter It,” *StopFake.org*, May 19, 2015, <https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/>.

⁶⁵ *Ibid.*

order of the Saudi government.”⁶⁶ One such tweet stated: “From the very beginning, false statements have tried to link the disappearance or killing of #Jamal_Khashoggi to the kingdom. This is a campaign they are waging against the kingdom.”⁶⁷

Another, less drastic example, occurred just after voting ending in the 2019 Kentucky gubernatorial election, when a network of accounts displaying bot-like behavior came alive and posted messages that the election was rigged. One tweet by user @Overlordkraken1, seen in Figure 26, claimed that the poster had just “shredded a box of Republican mail-in ballots.” Although there was no evidence of voter fraud, the incumbent requested a vote recount, alleging widespread irregularities.⁶⁸

Figure 26. A tweet sent after the Kentucky gubernatorial elections alleging voter fraud



Source: Joe Sonka, “Thousands of Twitter ‘Bots’ Targeted Kentucky with Fake News on Election Night,” *USA Today*, Nov. 11, 2019.

⁶⁶ Ben Collins and Shoshana Wodinsky, “Twitter Pulls Down Bot Network That Pushed Pro-Saudi Talking Points About Disappeared Journalist,” *NBC News*, Oct. 18, 2018, <https://www.nbcnews.com/tech/tech-news/exclusive-twitter-pulls-down-bot-network-pushing-pro-saudi-talking-n921871>.

⁶⁷ *Ibid.*

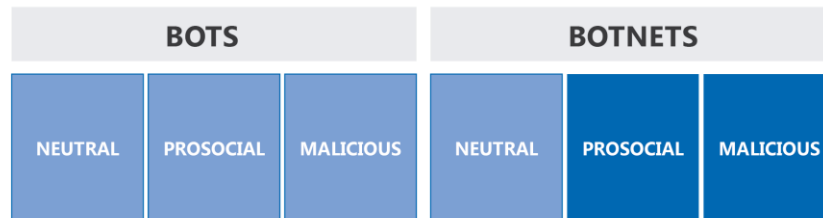
⁶⁸ Joe Sonka, “Thousands of Twitter ‘Bots’ Targeted Kentucky with Fake News on Election Night,” *USA Today*, Nov. 11, 2019, <https://www.usatoday.com/story/news/politics/2019/11/11/kentucky-elections-2019-thousands-twitter-bots-spread-fake-facts/2564439001/>.

Hijacking: taking over a conversation

Instances in which social media bots and botnets hijack a conversation occur largely around hashtag campaigns. In these cases, a hashtag campaign is being used to facilitate a conversation and/or cause a topic to trend. In response, a cohort of users (who may, or may not, be deploying social media bots or botnets) begins to post using the same hashtag. If this intervening cohort generates enough content, then they will successfully hijack the hashtag. In many instances, users have been seen hijacking hashtag campaigns on Twitter by appending the hashtag to “contrary or irrelevant messages.”⁶⁹

The role of social media bots and botnets in this activity is to increase the likelihood that the intervening cohort is able to hijack the hashtag successfully by increasing the rate of irrelevant or contradictory posts.

Figure 27. Types of examples included in this section



Source: CNA.

Prosocial example—botnet:

This example is one in which a botnet was not observed, but in which it *could have been* programmed to do what a collection of individual users did. In this case, Twitter account @LGBTfacts began tweeting messages that contained hate speech and/or were prejudicial toward LGBTQ+ individuals, using the hashtag #LGBTfacts. Another group of Twitter users discovered the hashtag and began using it to tweet their own messages and poke fun at the original @LGBTfacts account as seen in Figure 28. The responding users effectively hijacked the hashtag to spread the opposite message, turning #LGBTfacts into a hashtag that appeared

⁶⁹ Nathalie Marechal, “When Bots Tweet: Toward a Normative Framework for Bots on Social Networking Sites,” *International Journal of Communication 10* (2016), Feature 5022-5031, 2016, <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2016/10/marechal.pdf>.

to support the LGBTQ+ community through the use of humor and satire.⁷⁰ Again, although there is no indication that automated accounts were behind these tweets, this is certainly the kind of operation that bots could carry out were they programmed to do so.

Figure 28. A tweet hijacking the #LGBTfacts hashtag from its original anti-homosexual purpose



Source: Chris Matyszczyk, "Anti-Gay Twitter Hashtag Hijacked by Wit," *Cnet*, Jan. 24, 2012.

Malicious examples—botnets

Amid Egyptian antigovernment protests in September 2019, pro-Islamic State Twitter networks and accounts used the protestors' hashtags to try to persuade people to abandon the protests and join the Islamic State instead. Arabic hashtags such as "The people demand the fall of the regime," "Leave," and "Friday of Rage," and the English hashtag #sisi_out, all saw the posting of Islamic State propaganda videos as the group attempted to hijack the hashtags for its own purposes.⁷¹

In another incident, during the fourth democratic debate on October 15, 2019, coordinated botnets on social media used #DemDebates to promote unrelated issues and advance disinformation. One botnet capitalized on the viral hashtag to push anti-vaccine messaging and to retweet specific anti-vaccine accounts to promote the user's merchandise shop.⁷²

⁷⁰ Chris Matyszczyk, "Anti-Gay Twitter Hashtag Hijacked by Wit," *Cnet*, Jan. 24, 2012, <https://www.cnet.com/news/anti-gay-twitter-hashtag-hijacked-by-wit/>.

⁷¹ Alistair Coleman, "Analysis: Spammers and Terrorist Groups Exploit Egyptian Protest Hashtags," *BBC Monitoring*, Sept. 23, 2019, <https://monitoring.bbc.co.uk/product/c2013ul2>.

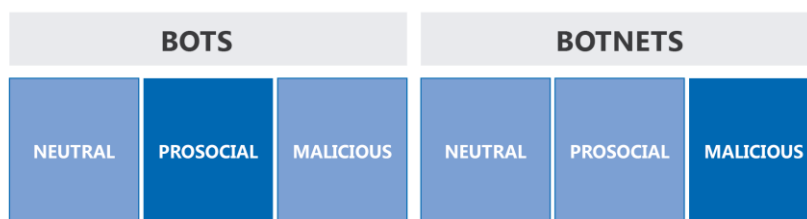
⁷² Elyse Samuels and Monica Akhtar, "Are 'Bots' Manipulating the 2020 Conversation? Here's What's Changed Since 2016," *Washington Post*, Nov. 20, 2019, <https://www.washingtonpost.com/politics/2019/11/20/are-bots-manipulating-conversation-heres-whats-changed-since/>.

Flooding: overwhelming a conversation or account

Social media bots and botnets can be programmed to flood either a message or an individual account. Flooding a message typically occurs by overwhelming a hashtag campaign with erroneous or irrelevant information. As a result of this action, those who click on the hashtag to follow the discussion are unable to find its core message. Although this might sound like hijacking, there is a clear difference. In hijacking, the primary purpose is to take over a conversation and change its message (e.g., changing a pro-ISIS message to an anti-ISIS message). In flooding, the primary purpose is to end a conversation by burying its message (e.g., making a pro-ISIS message hard to find by flooding the conversation with pictures of koala bears). In some cases, this activity might be referred to as hashtag spamming or hashtag poisoning: “the practice of affixing a specific hashtag to irrelevant content renders the hashtag unusable.”⁷³

Flooding an account is a bit different, and is reminiscent of traditional DDoS attacks in which a system (previously a website, and in this case a social media account) is overwhelmed with activity to shut it down and/or intimidate the individual. The objective in this case is to “[flood one’s] enemies with followers” in hopes that this will intimidate the user, overwhelm the account, or attract the attention of the platform and result in an account suspension.⁷⁴

Figure 29. Types of examples included in this section



Source: CNA.

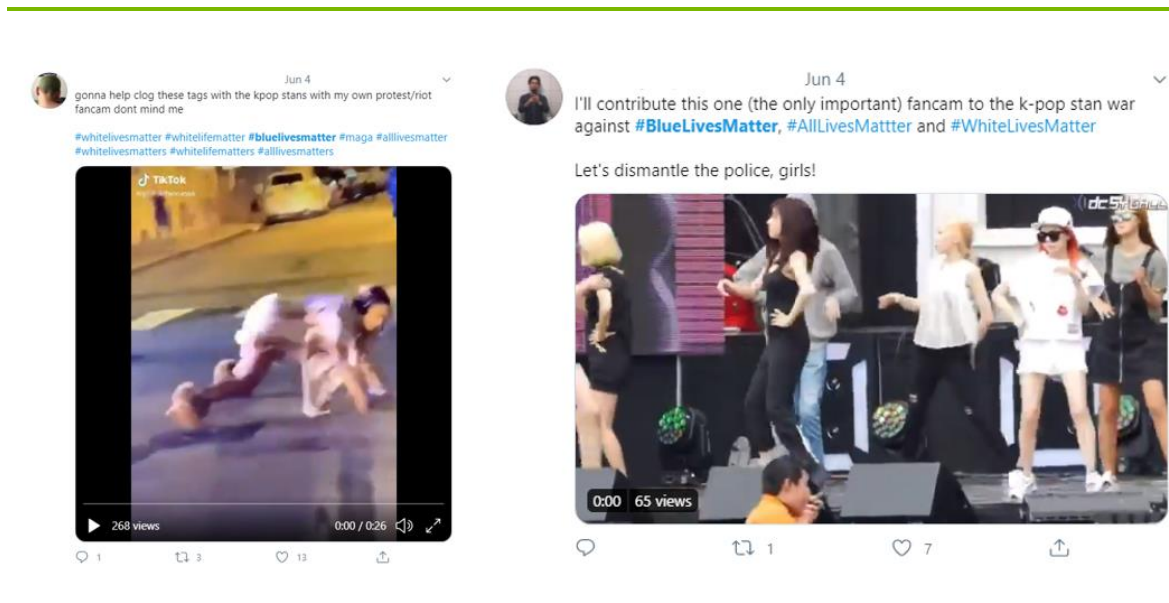
⁷³ Nathalie Marechal, “When Bots Tweet,” see footnote 75; and Erin Gallagher, *Mexican Botnet Dirty Wars*, presented at the Chaos Communication Camp 2015, Zehdenick, Germany, 2015, retrieved from https://media.ccc.de/v/camp2015-6795-mexican_botnet_dirty_wars#video.

⁷⁴ “The Surprising News Strategy of Pro-Russia Bots,” *BBC News*, Sept. 12, 2017, <https://www.bbc.com/news/blogs-trending-41203789>; and Brian Krebs, “Twitter Bots Use Likes, RTs for Intimidation,” *KrebsonSecurity* (blog), Aug. 30, 2017, <https://krebsonsecurity.com/2017/08/twitter-bots-use-likes-rt-for-intimidation/>.

Prosocial example (message)—bots

One of the more surprising developments of 2020 has been the political activism of K-Pop fans (i.e., fans of Korean pop music). In June 2020, it was reported that this group was flooding hashtags such as #whitelivesmatter with memes, photos, and videos of famous K-Pop stars.⁷⁵ Although in this example there is no evidence that social media bots or botnets were involved, this is a type of activity that social media bots or botnets *could have been* programmed to execute.

Figure 30. K-Pop fan videos posted in an effort to flood the hashtag #whitelivesmatter



Source: [https://twitter.com/search?q=\(%23whitelivesmatter\)%20until%3A2020-06-05%20since%3A2020-06-01&src=typed_query](https://twitter.com/search?q=(%23whitelivesmatter)%20until%3A2020-06-05%20since%3A2020-06-01&src=typed_query).

Malicious example (individual)—botnets

A clear example of botnets being used to flood an account is the case of Ben Nimmo, a nonresident senior fellow at the Atlantic Council's DFRLab. Nimmo has had a number of skirmishes with botnets over the years, and he has tweeted multiple times about Russian bots meddling in the US. In one instance, someone responded to Nimmo by making a copy of a

⁷⁵ James Vincent, "K-Pop Stans Are Flooding Right-Wing Hashtags Like #Bluelivesmatter And #MAGA," *The Verge*, Vox Media, Jun. 3, 2020, <https://www.theverge.com/2020/6/3/21278950/k-pop-stans-social-media-flooding-hashtags-bluelivesmatter-maga>.

colleague's profile page, creating an inaccurate tweet claiming that Nimmo had died, and using a botnet to amplify that tweet by retweeting it 21,000 times.⁷⁶ This experience angered Nimmo—particularly because friends and family had seen the news and reached out to confirm that he was okay—and he worked with the Atlantic Council to report on the Russian botnet. In response, the account posting the story was flooded as the botnet retweeted the story 106,000 times by the end of the day. It is possible that this type of attention will sound appealing because it ensures that the original tweet (with the story about the Russian botnet) will get more attention, but the accounts doing the retweeting did not have any followers so the story was not spreading, and the notifications flooding the target account quickly became unmanageable.⁷⁷ Because DFRLab's analysis led them to believe that the botnet was relatively unsophisticated they set a trap. Nimmo posted a tweet including the terms they believed would trigger the botnet, which was retweeted 500 times in the first nine minutes.

Figure 31. Ben Nimmo tweet flooded with responses



Source: Digital Forensic Research Lab, "#BotSpot: The Intimidators." Medium (blog), Aug. 30, 2017.

Nimmo and his colleagues alerted Twitter by using the same keywords, and including the handle @TwitterSupport, to trigger a botnet attack that would be very hard for Twitter to ignore. The botnet cooperated, and by the next morning, over 50,000 bots had tweeted @TwitterSupport.⁷⁸

⁷⁶ Digital Forensic Research Lab, "#BotSpot: The Intimidators," Medium (blog), Aug. 30, 2017, <https://medium.com/dfrlab/botspot-the-intimidators-135244bfe46b>.

⁷⁷ Digital Forensic Research Lab, "#BotSpot: Bots Boost NFL Divide," Medium (blog), Sept. 30, 2017, <https://medium.com/dfrlab/botspot-bots-boost-nfl-divides-abec2e025ddb>.

⁷⁸ Digital Forensic Research Lab, "#BotSpot: The Intimidators."

Figure 32. Ben Nimmo tweet alerting Twitter to the botnet



Source: Digital Forensic Research Lab, “#BotSpot: The Intimidators.” Medium (blog), Aug. 30, 2017.

Malicious examples (message)—botnets

Flooding hashtag campaigns has been observed in Mexico repeatedly. As one blogger noted, social media bots have “followed protesters from hashtag to hashtag...drowning out real conversations with noise.”⁷⁹

For example, in 2014 activists began to tweet using the hashtag #YaMeCanse (i.e., IAmTired) after 43 students went missing and were assumed dead. The hashtag became “a central hub for organizing protests and disseminating information,” but was soon “flooded with tweets that included the hashtag but no other content, except for a few random characters such as commas, [semicolons], and angle brackets.”⁸⁰ Although the article on this effort described the accounts as spambots, the ultimate objective was to flood the discussion and drown out real content. As the article noted, “It became difficult, if not impossible, for activists to actually share information with each other through the #YaMeCanse hashtag, and as a result it quickly

⁷⁹ Klint Finley, “Pro-Government Twitter Bots Try to Hush Mexican Activists,” *Wired*, Aug. 23, 2015, <https://www.wired.com/2015/08/pro-government-twitter-bots-try-hush-mexican-activists/>.

⁸⁰ *Ibid.*

dropped out of Twitter's trending topics. Bots, it seemed, had effectively jammed the protesters' communications channel.”⁸¹

In another example, journalist Erin Gallagher noted similar activity following protests in Mexico City. In this instance, the hashtag #RompeElMiedo (i.e., BreakTheFear) was being used by “a network of journalists, activists and human rights defenders to document human rights abuses in Mexico during protests.”⁸² As part of this effort, protestors were using the hashtag to share information about police activity so that protestors and journalists could evade arrest. One protestor, @anonopshispano, used the hashtag to share a map with a red outline of the area in which police were currently active.

Figure 33. Activist tweet using #RompeElMiedo to spread information about police activity



Source: Klint Finley, “Pro-Government Twitter Bots Try to Hush Mexican Activists,” *Wired*, Aug. 23, 2015.

⁸¹ Ibid.

⁸² J.M. Porup, “How Mexican Twitter Bots Shut Down Dissent,” *Motherboard Tech by Vice* (blog), Aug. 24, 2015, https://www.vice.com/en_us/article/z4maww/how-mexican-twitter-bots-shut-down-dissent.

According to Gallagher, the same bot-facilitated flooding occurred, with the result that the hashtag stopped being an effective way to share information.⁸³ As Gallagher noted in a talk:

"Arbitrary arrests began shortly thereafter...Activists were beaten. Protesters were beaten... In the end, the [protests] ended in brutal police repression...It's possible that these folks would still have been beaten regardless of whether they received the notifications or not, but clearly this was putting them in real danger by not being able to access this hashtag."⁸⁴

⁸³ This is obviously a complicated example as it is not clear whether the protestors were violent, or whether the police were corrupt. We have identified it as a case of malicious activity by taking the article at face value. The article said that the goal of the hashtag was to "share information about police locations so that protesters, journalists and bystanders could exit protest without being arrested or beaten." Because there is no indication that the protestors were being violent, we assessed the activist's effort to be prosocial, and thus assessed efforts to undermine it as antisocial.

⁸⁴ J.M.Porup, "How Mexican Twitter Bots Shut Down Dissent," *Motherboard Tech by Vice* (blog), Aug. 24, 2015, https://www.vice.com/en_us/article/z4maww/how-mexican-twitter-bots-shut-down-dissent.

Fracturing: breaking a large conversation into smaller conversations

Fracturing occurs exclusively around social media hashtag campaigns. In these cases, the interceding entity deploys a purposefully misspelled hashtag to break a larger conversation into a series of smaller conversations. These “misspelled hashtags are decoys, aimed at diffusing the reach of the original by breaking the conversation into smaller groups. These decoys can dilute certain voices and distort public perception of beliefs and values.”⁸⁵

Fracturing sometimes happens accidentally as a result of a simple spelling mistake. In 2015, following a mass shooting in California, the hashtags #SanBernardino and #SanBernadino both gained traction; in fact, the incorrectly spelled #SanBernadino was tweeted over 300,000 times and actually trended to the number 1 spot on Twitter.⁸⁶ In 2019, the hashtag #HurricaneDorian was fractured accidentally when some users posted using the misspelled hashtag #HurricaneDorain.⁸⁷ And on July 4, 2019, the US Air Force, the city of Boston, the US first lady of the, and the Canadian Space Agency used the incorrectly spelled #IndependenceDay to celebrate.⁸⁸

⁸⁵ Eoin O’Carroll, “From Russia, ‘With Hastags? How Social Bots Dilute Online Speech,” *The Christian Science Monitor*, Jul. 18, 2018, <https://www.csmonitor.com/Technology/2018/0718/From-Russia-with-hashtags-How-social-bots-dilute-online-speech>.

⁸⁶ Taylor Goldenstein, “#Sanbernadino Has Been Shared Over 333,000 Times Even Though It’s Misspelled,” *Los Angeles Times*, Dec. 2, 2015, <https://www.latimes.com/local/la-me-san-bernadino-misspelled-hashtag-20151202-htmlstory.html>.

⁸⁷ Johnny Diaz, “Hurricane Dorain? On Social Media, Dorian Is Getting Misspelled All Over The Place,” *SunSentinel*, Sept. 2, 2019, <https://www.sun-sentinel.com/news/weather/hurricane/fl-ne-hurricane-dorian-dorain-trending-social-media-20190902-3zdgrlaktjbanbmslca6nfuh5q-story.html>.

⁸⁸ Marcus Gilmer, “Twitter Can’t Even Celebrate Independence Day Without Misspelling the Hashtag,” *Mashable.com*, Jul. 4, 2018, <https://mashable.com/article/independence-day-twitter-hashtag-misspelled/>.

Figure 34. Incorrectly spelled Independence Day tweets



Source: Marcus Gilmer, "Twitter Can't Even Celebrate Independence Day Without Misspelling the Hashtag," Mashable.com, Jul. 4, 2018.

Successful fracturing works best when a botnet is able to make the misspelled hashtag trend. When this happens, Twitter begins to suggest the misspelled hashtag to people, which creates a cascading effect as people accidentally select the misspelled version and continue to fracture the conversation. The fracturing, at this point, is no longer contingent upon individual people accidentally misspelling the hashtag as Twitter is now an (inadvertent) contributor to the effort.

Importantly, although fracturing can function to dilute the power of political protests on social media by preventing hashtags from trending (by siphoning users onto the misspelled hashtag), it can also make it harder for people to find information (an issue raised following the San Bernardino shooting) and can undermine a user's ability to assess the degree to which others agree.⁸⁹

Although these examples appear to be the result of simple spelling errors, in the cases below we have focused on intentional instances propagated by bots.

⁸⁹ Taylor Goldenstein, "#Sanbernadino Has Been Shared Over 333,000 Times Even Though It's Misspelled," *Los Angeles Times*, Dec. 2, 2015, <https://www.latimes.com/local/la-me-san-bernadino-misspelled-hashtag-20151202-htmlstory.html>.

Figure 35. Types of examples included in this section



Source: CNA.

Malicious examples—botnets

In 2018, as the US found itself embroiled in a debate about how to deal with immigration at the southern border, activists used the hashtag #FamiliesBelongTogether to oppose policies that supported the separation of families. However, analysts found that Russian bots were active in supporting a number of related hashtags, including #FamiliesBelongTogther and #FamiliesBelongTogether. According to one article, a dashboard that “monitors the top 600 pro-Kremlin Twitter accounts, found that the decoy hashtag #FamiliesBelongTogether was the third most-tweeted hashtag on June 30 and July 1, the weekend that thousands took to the streets to march against the president’s immigration policies.”⁹⁰

In supporting these misspelled hashtags, the Russian botnet aspired to “train Twitter’s search algorithms to see the misspelled versions as trending topics.”⁹¹ And then, because they were trending, Twitter began to suggest them to users who were searching for the correctly spelled hashtag.

⁹⁰ Eoin O’Carroll, “From Russia, ‘With Hastags? How Social Bots Dilute Online Speech,” *Christian Science Monitor*, Jul. 18, 2018, <https://www.csmonitor.com/Technology/2018/0718/From-Russia-with-hashtags-How-social-bots-dilute-online-speech>.

⁹¹ *Ibid.*

Figure 36. Twitter recommends incorrectly spelled hashtag



Auto Fill Search Results in Twitter Last Week Showing the Decoy Hashtag as second most recommended

Source: Tim Chambers, "#FamiliesBelongTogether Robotic Attack This Week," Medium (blog), July 1, 2018.

As a result, the incorrect hashtags gained even more attention. Ultimately, the list of people who used an incorrect hashtag included two senators, one congresswoman, and the American Civil Liberties Union.⁹² The goal of such an effort is not to sow discord by pitting the different hashtags against one another, but to fragment the original conversation into a series of small conversations. This action both decreases the likelihood that any single hashtag will trend, and makes it difficult for a single community to share information.

⁹² Ibid.

Implications of Future Trends

The landscape in which social media bots and botnets operate is so complex and dynamic that it is nearly impossible to predict the most likely near- to mid-term evolutions, short- or medium-term threats, or the most concerning trends. To mitigate against this challenge, we have identified four areas that our literature review and/or SME interviews identified as likely to evolve in the near- to mid-term—regulations and authorities, activity in developing countries, a technological arms race, and an increase in active users. In the following section, we provide some background on each of these issues and identify the challenges and opportunities that we believe it represents for SOF—taking into consideration SOF’s role as both an offensive and defensive actor.

Laws, regulations and authorities

Since the 2016 US presidential election, awareness of the largely unregulated nature of social media platforms has been on the rise and calls have grown for greater regulation of the platforms. In 2019, Facebook founder and CEO Mark Zuckerberg wrote, in a *Washington Post* editorial, that the responsibility for moderating decisions was too great for just the platforms to bear and more regulation was needed in certain areas.⁹³

US legislators and politicians have recently attempted to force social media companies into greater regulation through a number of proposals. These efforts have taken aim at the platforms’ protection from liability for content posted on their sites, as enshrined in the Communications Decency Act Section 230, which is discussed in more detail in our companion report, *Social Media Bots: Laws, Regulations, and Platform Policies*. Some lawmakers have proposed changes to the law, including Senator Josh Hawley, who would require the companies to prove they use politically neutral moderation policies to retain their Section 230 immunity. Others, such as former Vice President Joe Biden, have called for the complete repeal of the provision.⁹⁴ President Trump also issued an executive order in May 2020 that requires the

⁹³ Mark Zuckerberg, “Mark Zuckerberg: The Internet Needs New Rules. Let’s Start in These Four Areas,” *Washington Post*, Mar. 30, 2019, https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html.

⁹⁴ “Senator Hawley Introduces Legislation To Amend Section 230 Immunity For Big Tech Companies,” Press Release On Senator Josh Hawley’s Website, Jun. 19, 2019, <https://www.hawley.senate.gov/senator-hawley-introduces-legislation-amend-section-230-immunity-big-tech-companies>; Bambauer, Derek, “How Section 230

Federal Communications Commission to present recommended regulations revising Section 230.⁹⁵ It is unclear exactly what form platform regulation will take in the future. By extension, it is also unclear whether the use of bots for both offensive and defensive purposes will become more difficult, although the trend line is clearly toward greater regulation.

DOD authorities related to information operations are also evolving, albeit in the opposite direction, as the government moves toward an increased ability to compete in the information space. Cyber Command and/or SOF forces have historically had limited ability—i.e., requiring pre-approval by SECDEF or by certain other persons—to engage in offensive online cyber operations. However, there are indications that those restrictions may be loosening. For example, in 2018 the executive and legislative branches authorized greater permissions for engaging in offensive cyber capabilities, resulting in enhanced cyber operations for certain DOD entities.⁹⁶ In May 2019, officials from Cyber Command stated the new authorities have led them to “conduct more cyberspace operations in the last few months than in the previous 10 years.”⁹⁷ In addition, US SOCOM is in the process of fully standing up the JMWC. The JMWC, expected to be fully operational by 2025, will “[support] the combatant commands with improved messaging and assessment capabilities, shared situational awareness of adversary influence activities, and coordinated internet-based MISO globally,” which presumably will involve interaction with, if not the deployment of, social media bots.⁹⁸

Although certain concerns persist, and it is unclear how these new capabilities might relate to social media bots, a greater ability to employ offensive capabilities could also lead to enhanced USG activity in this space.

The trend lines of government regulations and DOD authorities may ultimately come into conflict, as the government seeks enhanced regulation of the social media companies just as

Reform Endangers Internet Free Speech,” *The Brookings Institution*, Jul. 1, 2020, <https://www.brookings.edu/techstream/how-section-230-reform-endangers-internet-free-speech/>.

⁹⁵ “Executive Order on Preventing Online Censorship,” White House, May 28, 2020, <https://www.whitehouse.gov/presidential-actions/executive-order-preventing-online-censorship/>.

⁹⁶ Ellen Nakashima, “U.S. Cyber Command Operation Disrupted Internet Access of Russian Troll Factory on Day of 2018 Midterms,” *Washington Post*, Feb. 27, 2019, https://www.washingtonpost.com/world/national-security/us-cyber-command-operation-disrupted-internet-access-of-russian-troll-factory-on-day-of-2018-midterms/2019/02/26/1827fc9e-36d6-11e9-af5b-b51b7ff322e9_story.html.

⁹⁷ Mark Pomerleau, “New Authorities Mean Lots of New Missions at Cyber Command,” *Fifth Domain*, May 8, 2019, <https://www.fifthdomain.com/dod/cybercom/2019/05/08/new-authorities-mean-lots-of-new-missions-at-cyber-command/>.

⁹⁸ Richard D. Clarke, “Statement of General Richard D. Clarke, U.S. Army Commander United States Special Operations Command,” House Armed Services Committee Intelligence, Emerging Threats and Capabilities Subcommittee, Apr. 9, 2019, https://armedservices.house.gov/_cache/files/7/9/7970f176-0def-4a2d-beb3-a7d5d69e513b/9C80F888EEE40D8E82ABFF5336C012C3.hhrg-116-as26-wstate-clarker-20190409.pdf.

DOD grants greater room for operating in the information space. It is possible the latter trend could restrict SOF's ability to use or defend against botnets, even if it gains greater DOD authority to carry out relevant operations.

Implications

- Cyber Command and/or SOF forces have historically had limited ability—i.e., requiring pre-approval by SECDEF or certain other persons—to engage in offensive online cyber operations. SMEs we spoke with for this study were not sanguine that DOD entities such as SOF would be afforded authorities to deploy social media bots or botnets (either offensively or defensively) in the immediate future. However, they held out hope that this might change in the years to come.
- As the trend toward greater government and/or self-regulation of social media platforms persists, the hopes of those elements within the government (e.g., SOF) for possible use of social media bots/botnets may prove even more difficult to realize.
- SOF may therefore want to try and get ahead of the possible trend toward increased government regulation—and possible increased platform restrictions on the use of bots and botnets—by advocating strongly to gain authorities now for the use of social media bots and botnets in prosocial or defensive (e.g., force protection) functions.

Activity in developing countries

Across the West, there is a growing awareness of the tactics of disinformation and the use of social media bots by malign actors, prompting increasing attention and regulation focused on this area. However, developing countries, particularly those with native languages outside of those spoken regularly in the West, often receive less focus from the mostly US-based social media companies, which also may not understand the cultural and political context in these countries. Although malicious use of bots may pose less of a threat to Western democracies in the future because of possibly greater USG regulation, it is possible that their use could accelerate in developing nations.

Although internet usage is steadily increasing in many emerging and developing economies, these countries often have weak institutions, lower media and digital literacy, and prohibitive costs for access to unlimited data (which may limit a user's ability to fact check), making their

populations more susceptible to the influence of social media bots.⁹⁹ Many of these countries also have simmering ethnic tensions ripe for exploitation by malign actors.¹⁰⁰ At the same time, developing countries are increasingly acting as a home base for actors wishing to sow discord abroad. For example, in March 2020, Facebook discovered a network of Russian-based professional trolls outsourcing their disinformation work to Ghanaian and Nigerian individuals and paying them to target the US.¹⁰¹ These African citizens used a number of deceitful tactics, including fake accounts where they posed as non-governmental organizations or bloggers, for example, toward that aim.¹⁰² Although there is no evidence that they used social media bots in these efforts, it is easy to see how adversaries could outsource bot work to individuals in such countries in the future.

Implications

- The US fight against VEOs is likely to persist for many years, and the bulk of the affiliates of groups like al-Qaeda and the Islamic State reside in developing, failing, or failed states. These groups have a long history of using social media to further their agendas; the Islamic State of Syria and Iraq even had an application, Dawn of the Glad Tidings, that used automation to amplify its Twitter posts.¹⁰³ It is reasonable to assume, then, that these groups will attempt to use social media bots to exploit existing tensions or vulnerabilities in vulnerable countries to further weaken governments, sow discord, and stimulate civil wars and internecine conflicts that these groups exploit to survive and expand.¹⁰⁴ The implications of this for SOF could ultimately be increased requirements

⁹⁹ Conor Sanchez, "Misinformation is a Threat to Democracy in the Developing World," CFR Net Politics (blog), Council on Foreign Relations (CFR), Jan. 29, 2019, <https://www.cfr.org/blog/misinformation-threat-democracy-developing-world>.

¹⁰⁰ Ibid.

¹⁰¹ Alex Hern and Luke Harding, "Russian-Led Troll Network Based in West Africa Uncovered," *The Guardian*, Mar. 13, 2020, <https://www.theguardian.com/technology/2020/mar/13/facebook-uncovers-russian-led-troll-network-based-in-west-africa>.

¹⁰² Ibid.

¹⁰³ J. M. Berger, "How ISIS Games Twitter: The Militant Group That Conquered Northern Iraq Is Deploying a Sophisticated Social-Media Strategy," *The Atlantic*, 2014, <https://www.theatlantic.com/international/archive/2014/06/isis-iraq-twitter-social-media-strategy/372856/>

¹⁰⁴ Julia McQuaid, Jonathan Schroden, Pamela G. Faber, P. Kathleen Hammerberg, Alexander Powell, Zack Gold, David Knoll, and William Rosenau, CNA, 2017, *Independent Assessment of U.S. Government Efforts against Al-Qaeda*, CNA DRM-2017-U-015710-Final.pdf

Vera Zakem, Megan McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaigns*, CNA, 2018, https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf.

for CT forces in these countries, whether acting directly or indirectly by training partner nation security forces.

- As the global competition between the US and the likes of China and Russia heats up, these countries are likely to seek increased influence in non-aligned countries worldwide. Given the relatively weak media environments in many of these countries, the use of social media—and social media bots and botnets—could become a primary avenue for competitive activities.
- SOF entities such as the JMWC—if properly resourced and given the right authorities—could serve as frontline entities in detecting and combatting these types of adversary operations, both against VEOs and state actors. The JWMC and other SOF entities, such as the US Army Special Operations Command’s two POGs, if properly modernized, resourced, and authorized, could partner with State Department entities such as the Global Engagement Center to provide training for partner nation forces on how to establish their own social media botnets for defensive purposes (as the employment of those nets would be regulated by partner nation authorities). It may even be possible to leverage existing authorities, such as the 127(e) or 1202 programs, to provide this type of training to partner nation forces.¹⁰⁵
- SOF could also work in partnership with US embassy public diplomacy and media development efforts to employ bots for prosocial activities in developing countries, such as promoting public health and education initiatives or providing information in the wake of crisis events (e.g., natural disasters).

A technological arms race

We mentioned the issue of an “arms race” between bots and bot-detectors in our “Identifying Bots and Botnets” section. As bots have become more sophisticated, expert academics and social media network administrators struggle to differentiate bots from human “bot-like activity.”¹⁰⁶ As AI and ML advance, botmasters will likely be able to out-innovate bot hunters; the algorithms that detect bots are written to identify known characteristics, so adaptations in behavior will keep bots one step ahead.¹⁰⁷

¹⁰⁵ The language of 127(e) authorizes the DOD to use surrogate forces to counter violent extremist organizations; the language of 1202 authorizes DOD to employ surrogates for the purpose of irregular warfare against state adversaries.

¹⁰⁶ Kai-Cheng Yang et al., “Arming the Public with Artificial Intelligence to Counter Social Bots,” *Human Behavior and Emerging Technologies* 1.1, Feb. 6, 2019, arXiv:1901.00912v2 [cs.CY].

¹⁰⁷ Ibid.

AI may also increase the ability of programmers to deploy large numbers of “quality bots,” without the characteristics that would easily identify the account as a suspected bot. During a recent CNA national security seminar on “The Tools of Digital Information Operations,” COGSEC’s Tim Hwang pointed out that, today, a programmer has to decide between deploying a handful of quality bots or a large number of bots that would probably be quickly detected. With developments in AI, Hwang noted, there will be no such trade-off in the future.

Implications

- This trend has the possibility of negating the earlier impacts of increasing government regulation, because even if governments try to place more regulations on the use of botnets, the enforcement of those regulations would require successful and unambiguous detection of them and their use. This means that the space may yet exist for actors such as SOF to employ botnets even if a trend toward government attempts to restrict that space move ahead. In other words, SOF should not disregard the potential use of botnets as a tool for information operations even if the previous trend continues or accelerates. Additionally, SOF may consider investing in AI/ML technologies to improve the quality of social media bots to get ahead of the government regulation trend.
- The increased sophistication of bots could also pose challenges to the ability to detect their use, for example in developing, failing, and failed countries as mentioned above. If SOF believe they are able and desire to play the types of roles we suggested in those areas, they may need to invest also in improving technologies to detect the use of social media bots and botnets.
- The ability of US adversaries to employ social media bots to distract US forces (including SOF) in a MILDEC campaign will be enhanced by further improvements in the quality of bots themselves, unless the US can keep up in terms of technologies to detect their use. The potential force protection impacts of this increased MILDEC capability are likely to be most acutely felt by SOF, which typically operate as small teams in contested environments.
- Of course, US military forces (including SOF) could also deploy botnets for MILDEC against US adversaries, if they were given requisite resources and authorities to do so.

An increase in active users

The diversification and increase in the users of social media bots and botnets is likely to occur in a number of registers.

First, as mentioned in the previous section, advances in technology have increased, and will continue to increase, the number of people able to successfully deploy high-quality social

media bots and botnets. This practice will occur not only because technology will improve the quality of the bots and botnets but also because technology will ease the deployment of these tools, thus removing a barrier (i.e., technical skill) to use.

Second, we noted above that developing countries may be especially vulnerable to social media bot and botnet campaigns. It is important to keep in mind, though, that these developing countries are not merely passive actors. Over the coming years, a growing number of individuals from these countries will likely become active users of social media bots, successfully deploying them to pursue a variety of domestic, regional, and global objectives.

Third, although Russia has long-dominated the discussion about the malicious use of social media bots, it is by no means the only nation state engaged in this activity. A 2015 report found that “more than 40 countries deployed political bots” in an effort to “mimic social media users and manipulate public opinion.”¹⁰⁸ The 2019 “Freedom of the Net” report by Freedom House came to a similar conclusion in listing the key findings from its analysis of 2018 online activity:

More governments enlist bots and fake accounts to manipulate social media. Political leaders employed individuals to surreptitiously shape online opinions and harass opponents in 38 of the 65 countries covered in this report—another new high.¹⁰⁹

Implications

- As access to high-quality bots spreads, it may become increasingly difficult to detect bots and their activities. The increased sophistication of bots, combined with a trend toward their democratization, further indicates that SOF should consider investments in bot detection technologies.
- A global increase in social media bot and botnet use will likely mean a corresponding increase in the difficulty of countering bot campaigns. This could be because the noise of botnet activity worldwide makes it more difficult to detect any single campaign or because the nature of the campaigns could themselves become more sophisticated (e.g., employing large numbers of coordinated or uncoordinated botnets as a “Voltron”-like superbots).
- The global increase in social media bot and botnet use could also be exploited by USG actors (including SOF), who also might be able to hide their activities in the noise.

¹⁰⁸ Leanna Garfield, “5 Countries That Use Bots to Spread Political Propaganda,” *Business Insider*, Dec. 16, 2015. <https://www.businessinsider.com/political-bots-by-governments-around-the-world-2015-12#mexico-1>.

¹⁰⁹ “Press Release: Freedom on the Net 2019 Reveals Crisis on Popular Platforms,” Freedom House, last modified Feb. 2020. Accessed Sept. 8, 2020. <https://freedomhouse.org/report/freedom-on-the-net/2019/the-crisis-of-social-media/press-release>.

- An increase in the ease of use and proliferation of social media botnets could eventually lead to the creation of a global social media I&W network for SOF activities (e.g., botnets that could look for indicators of SOF activity and immediately amplify them for the sake of exposure). To combat such a possibility, SOF may consider investing in technology to create force protection bots—botnets that might employ flooding or fracturing techniques to counter the amplification of SOF activity indicators. The US could also seek to create its own amplification botnets for adversary activities of various kinds.

Opportunities and Risks for SOF

While we have outlined the likely implications of near- to mid-term changes in the social media bot environment, the reality is that social media bots are currently a tool of incredible utility that remains untapped by SOF and the USG. Below we discuss both opportunities and risks for the SOF of social media bot use.

Reduced need for cultural expertise

First, it is not always necessary to have in-depth cultural knowledge to deploy social media bots or botnets effectively. It has long been recognized that formulating an effective message is part of the ability to influence digitally. This is as true for prosocial and neutral messages as it is for malicious ones. Russia has proven particularly adept at crafting messages that will resonate with its desired audiences. One of its primary tactics is to target specific groups to play on existing grievances and exacerbate tensions. As described in CNA's report *Exploring the Utility of Memes for US Government Influence Campaigns*, "significant cultural, contextual, and experiential knowledge is required" for effective messaging, "as is a granular understanding of the intended audience."¹¹⁰ Russia uses its knowledge of target audiences to identify its opponents' vulnerabilities and formulate a plan for exploiting them.

Although it is true that the effective use of social media bots may require cultural expertise in some instances (unlike for the use of memes), it is not universally true. In some cases, social media bots and botnets fall outside this rule and provide a means of influencing conversations with no cultural expertise. For example, a trending hashtag or social media account could be flooded with content that is completely meaningless (e.g., photos of giraffes) or fractured by simply changing one letter in the spelling of the hashtag. In each instance, it is not necessary to have a sophisticated understanding of the target audience to achieve the desired effect.

Rapidly deployable capability

Second, social media bots and botnets, especially when the architecture is already in place, can be deployed incredibly quickly. Most analyses on how to respond to disinformation campaigns or social media botnet attacks tend to focus on the long-term goals of eliminating the capability.

¹¹⁰ Vera Zakem, Megan McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaign*, CNA, 2018, https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf.

After the 2016 US presidential election, for example, a wide range of US actors began to develop plans for responding to similar future crises. As late 2019 reporting notes, Cyber Command has been developing “capabilities that could be deployed against Russian entities if they attempt to interfere in the 2020 elections...[and that] build on an operation from 2018 when Cyber Command used emails, pop-ups, and texts to target Russian trolls and took the Internet Research Agency’s servers offline.”¹¹¹

A disadvantage of this type of technical approach is the amount of preparation required to facilitate an effective response. A brief description of the 2018 operation suggests, for example, that it was minimally necessary to know the email addresses and phone numbers of the Russian actors. Although use of social media bots is by no means a panacea, and there are significant challenges to confront in terms of authorities, it is noteworthy that in some cases social media bots could offer SOF an exceptionally quick answer to a detected threat.

For example, SOF might flood or fracture a social media campaign that posed a threat to an ongoing operation by attempting to reveal the location of the activity. In some ways, it would not matter if the effort to interfere with the operation was being coordinated by a state adversary or if it was an organic movement among the local population. In both cases, SOF could respond quickly and effectively, and with far less skill or confirmed intelligence than a more aggressive technical response might require. Social media bots and botnets thus offer a means to respond without the need for executing a full technical takedown of an adversarial botnet.

¹¹¹ Ellen Nakashima, “U.S. Cybercom Contemplates Information Warfare to Counter Russian Interference in 2020 Election,” *Washington Post*, Dec. 25, 2019. https://www.washingtonpost.com/national-security/us-cybercom-contemplates-information-warfare-to-counter-russian-interference-in-the-2020-election/2019/12/25/21bb246e-20e8-11ea-bed5-880264cc91a9_story.html.

Wide range of applications

Third, it is not difficult to identify instances in which SOF might use social media bots or botnets in each of the categories of the taxonomy we identified (see Table 1).

Table 1. Social media bots and opportunities for SOF

Bot Activities	Examples of SOF use
Distributing	SOF might partner with a US embassy to overtly share information about a natural disaster
Amplifying	SOF might partner with a US embassy to overtly amplify information about a public health campaign
Distorting	SOF might covertly attempt to increase discord by posting culturally relevant incendiary content in a foreign-language social media campaign with US national security implications
Hijacking	SOF might covertly or overtly respond to an identified Russian disinformation campaign by hijacking the relevant hashtag and turning the campaign into something innocuous or prosocial
Flooding	SOF might covertly overwhelm a social media account that threatened an ongoing direct action by live-tweeting details of the event
Fracturing	SOF might covertly break a social media campaign into multiple parts in an effort to diffuse the impact of messaging that threatened US forces by revealing their locations

Source: CNA.

Reputational risk

There are, of course, risks inherent in some of this activity. Discussions of the use of social media bots and botnets overwhelmingly frame both the activity and the perpetrators as malicious. This is especially evident in analyses of Russian social media activity, which is shaped by a number of geopolitical issues—long-standing tensions, a GPC backdrop, and the obviously nefarious motivations and objectives of the Russian government—regardless of whether the Russian botnets being analyzed are linked to the Russian government. It is consequently likely that a non-negligible percentage of US-deployed social media bots and botnets will be framed as malicious as they will be interpreted against a background of US hegemony and cultural aggression. In some cases, the risk might be outweighed by the benefits. For example, flooding a new hashtag with irrelevant information to bury information about the location of an ongoing operation might be worth the potential backlash. In other cases, the cost/benefit analysis might not support the deployment of social media bots or botnets.

However, as this report makes clear, social media bots and botnets can also be deployed to prosocial ends and can be deployed openly. As one social media expert we spoke with noted, there are instances in which “declared automation”—that is, automated accounts that self-identify as such—can perform socially beneficial services.¹¹² For example, the US might use its resources to fund and deploy a social media bot sharing information about earthquake activity with a population that otherwise does not have access to reliable information.

¹¹² Interview with private industry SME, Feb. 27, 2020.

Conclusion

Social media bots and botnets have been a consistent feature of mainstream news coverage for nearly four years. Many of the reports written about this tool, though, have focused on the broader issue of disinformation or the technological challenge of stopping nefarious actors. This report, by contrast, has highlighted six ways in which social media bots function to influence online discourse: distributing, amplifying, distorting, hijacking, flooding, and fracturing. Our goal in doing so was to turn the descriptive (i.e., a description of activities that have been observed to date) into the prescriptive (i.e., a toolkit that could be operationalized without significant technical skill or intelligence on the source of the activity).

The future of social media bots is largely unknowable, but in assessing the near- to mid-term landscape—with a particular focus on the ways in which these tools can influence conversations on social media—it seems clear that there any path forward is ripe with both challenges and opportunities for US government actors. Furthermore, this report highlights that the successful deployment of social media bots and botnets is not contingent on deep cultural knowledge or technical expertise. Rather, these are relatively simple mechanisms that have a tremendous capacity to influence discourse, and, if given adequate consideration and deployed correctly, could be a powerful asset in the toolkit of the USG.

Appendix A: The Legal Landscape and Platform Policies

The laws and platform policies governing the use of social media bots could have a great impact on US public discourse, national security, and democracy as the country continues to grapple with both foreign and domestic disinformation. However, the landscape of the relevant laws, regulations, and policies is difficult to piece together, and little literature exists bringing all applicable provisions and policies together in one place. In this section, we briefly describe the current limited landscape of legal restrictions on bots in the US and in the international system and the ways the platforms themselves address the use of automation on their sites. For a more detailed discussion of this topic, please refer to this report's companion paper, *Social Media Bots: Laws, Regulations, and Platform Policies*.

The legal landscape

Currently, the US does not have a clear plan for combatting internet-based information influence campaigns, evidenced partly by the dearth of federal laws related to the subject.¹¹³ A number of challenges, including constraints from the First Amendment, hinder attempts to pass bot- and botnet-related legislation. Though there have been attempts to pass bot-related legislation in Congress, including California Senator Diane Feinstein's efforts to pass the Bot Disclosure and Accountability Act, no bills have yet become law.¹¹⁴ In fact, the only bot-related law to pass in the US has been at the state level, when California's Bolstering Online

¹¹³ Sophie Kodner, "Covert Bots: The Cyber-Nuisances Threatening our Newsfeeds and Our Democracy," *New Perspectives in Foreign Policy*, no. 13 (2017), p. 26.

¹¹⁴ Bot Disclosure and Accountability Act of 2018, S. 3127, 115th Cong., 2018; Bot Disclosure and Accountability Act of 2019, S. 2125, 116th Cong., 2019.

Transparency, or B.O.T., Act became law on July 1, 2019.¹¹⁵ New Jersey and Washington subsequently introduced similar bills, although they had not passed as of September 2020.¹¹⁶

The EU has been at the forefront of international efforts to oppose disinformation and has developed a Code of Practice urging social media platforms to implement self-regulation measures. Many of the largest sites, including Facebook and Twitter, signed on in October 2018.¹¹⁷ Bot-relevant provisions in the Code include those aimed at closing fake accounts and labeling bot interactions.¹¹⁸ In May 2020, the European Commission released its yearly findings on the Code's implementation.¹¹⁹ The report stated that, while there have been positive developments, much progress needs to be made, specifically with bot labeling and removal.¹²⁰

Platform policies

Because governments face a number of challenges in their attempts to pass legislation and regulations on social media bots, much of the burden instead falls on the social media companies. However, the platforms also face unique dilemmas when thinking through effective bot regulation, including the need to answer to shareholders who many see bot activity as valid engagement that leads to higher share prices. The web of platform policies related to bots is difficult to untangle, as the sites often publish applicable provisions in disparate locations, including within their blog posts, community standards, developer policies, and newsrooms. In addition, because the language used can vary by platform, it is often difficult to tell when a

¹¹⁵ Bolstering Online Transparency Act, Cal. Code BPC § 17940 (2019); Renee Diresta, "A New Law Makes Bots Identify Themselves—That's the Problem," *Wired*, Jul. 24, 2019. <https://www.wired.com/story/law-makes-bots-identify-themselves/>.

¹¹⁶ New Jersey Assembly, No. 4563, 218th Legislature, Oct. 15, 2018; Jonathan Lai, "Is That Tweet from a Human?" *Philadelphia Inquirer*, Dec. 28, 2018, <https://www.inquirer.com/politics/nj-bot-bill-disclosure-proposal-law-social-media-free-speech-20181228.html>.

¹¹⁷ "Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Tackling Online Disinformation: A European Approach," *Eur-Lex*, Apr. 26, 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236>; "Code of Practice on Disinformation," European Commission, Sept. 26, 2018. <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>; "Code of Practice on Disinformation one year on: online platforms submit self-assessment reports," European Commission, Oct. 28, 2019.

¹¹⁸ "A Europe That Protects: The EU Steps Up Action Against Disinformation," European Commission, Dec. 4, 2018. https://europa.eu/rapid/press-release_IP-18-6647_en.htm.

¹¹⁹ Iva Plasilova et al., *Study for the "Assessment of the implementation of the Code of Practice on Disinformation,"* European Commission, 2020, 3. <https://ec.europa.eu/digital-single-market/en/news/study-assessment-implementation-code-practice-disinformation>.

¹²⁰ *Ibid.*, 3, 48.

new policy supersedes an older policy, and most platforms do not have policies specifically addressing bots. Despite these difficulties, social media platform policy provisions applicable to bots can be roughly broken out into four categories: automation, fake accounts and misrepresentation, spam, and artificial amplification.

- **Automation policies:** Some platforms have general policies for running automated software on their sites.¹²¹ Because the platforms recognize bots can also provide valuable services or have benign purposes, these policies typically prohibit bots from platform manipulation, while allowing the non-manipulative use of bots.¹²²
- **Fake account and misrepresentation policies:** Many policies prohibit the creation of inauthentic profiles and misrepresentation of identity, including the impersonation of others.¹²³ While real people create some fake accounts, bots also create and/or operate accounts that fall into this category. Thus, these policies cover activity by both bots and real people. However, not all types of fake accounts are malicious, and the platforms often expressly state that some types of impersonation, including satirical and fan accounts, are permissible. The bots behind these types of accounts must generally be expressly labeled, though, to fit within the permitted uses.
- **Spam policies:** The definition of “spam” varies by platform and is sometimes conflated to mean “inauthentic engagement” on the whole, but, for the purposes of our research, we used the dictionary definition of “unsolicited usually commercial messages [...] sent to a large number of recipients or posted in a large number of places.”¹²⁴ Bots are useful for this type of bulk messaging, and many platforms forbid any type of automation for this type of high-volume communication.¹²⁵

¹²¹ Colin Crowell, “Our Approach to Bots and Misinformation,” Twitter Blog, Jun. 14, 2017, https://blog.twitter.com/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.html.

¹²² Digital Forensic Research Lab, “Is This Granny a Bot? The Challenges of Detecting Automation,” Medium (blog), Nov. 6, 2019. <https://medium.com/dfrlab/is-this-granny-a-bot-the-challenges-of-detecting-automation-115a2082b410>

¹²³ “Misrepresentation,” Facebook Community Standards; “Terms of Use,” Instagram. <https://help.instagram.com/581066165581870>; “WhatsApp Legal Info,” WhatsApp, <https://www.whatsapp.com/legal/>.

¹²⁴ “spam,” Merriam-Webster. <https://www.merriam-webster.com/dictionary/spam>.

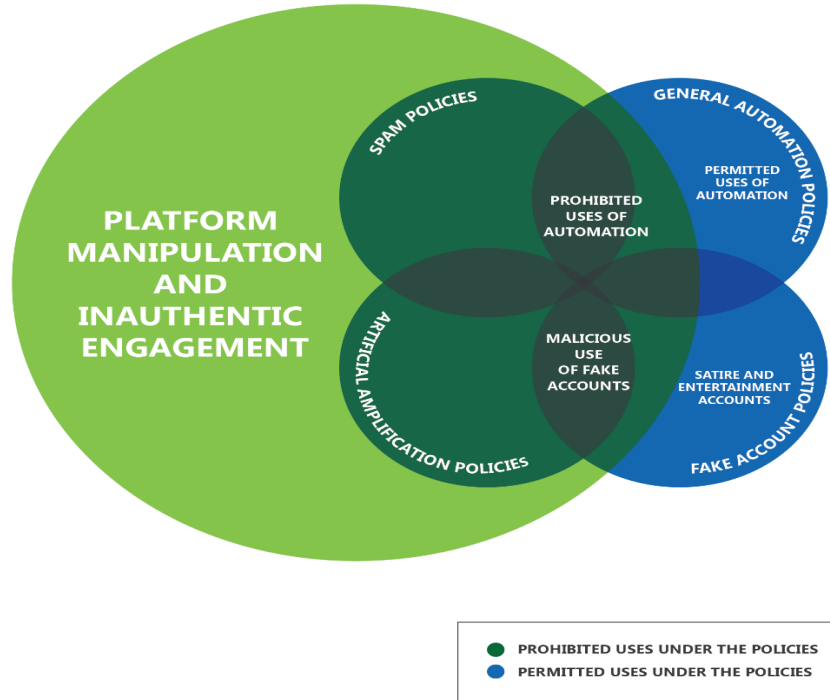
¹²⁵ Emily Taylor, Stacie Walsh, and Samantha Bradshaw, *Industry Responses to the Malicious Use of Social Media*, NATO Stratcom Centre of Excellence, 2018, 12. <https://www.stratcomcoe.org/industry-responses-malicious-use-social-media>; Yoel Roth, “Automation and the Use of Multiple Accounts,” Twitter Developer Blog, Feb. 21, 2018. https://blog.twitter.com/developer/en_us/topics/tips/2018/automation-and-the-use-of-multiple-accounts.html.

- **Artificial amplification policies:** For our research, “artificial amplification” means the use of inauthentic means, whether automation- or human-generated, to make posts, profiles, etc. seem more popular than they actually are. This is often done to make a topic trend or to displace an already trending topic with a new one. Bots are useful for boosting popularity because they provide a relatively easy means of creating a high volume of inauthentic engagement or a large number of fake followers. Platforms wholly prohibit this kind of manipulation of their sites and address it in a variety of ways, including through policies aimed at the falsification of popularity, the manipulation of trending topics, and the generation of fake views.¹²⁶

Figure 37 depicts the way these policies often overlap in setting out prohibited bot behaviors. It also shows that, while the behavior covered within spam and artificial amplification policies is wholly banned as impermissible platform manipulation, general automation and fake account policies allow for some types of behavior related to social media bots, while forbidding other types of behavior.

¹²⁶ “Terms of Use”; “Terms of Service,” YouTube, Dec. 10, 2019. <https://www.youtube.com/static?template=terms>; “Automation and the Use of Multiple Accounts,” Twitter Developer Blog, Feb. 21, 2018, https://blog.twitter.com/developer/en_us/topics/tips/2018/automation-and-the-use-of-multiple-accounts.html; “Automation rules,” Twitter Help Center, Nov. 3, 2017. <https://help.twitter.com/en/rules-and-policies/twitter-automation>; Michael Keller, “The Flourishing Business of Fake YouTube Views,” *New York Times*, Aug. 11, 2018. <https://www.nytimes.com/interactive/2018/08/11/technology/youtube-fake-view-sellers.html>.

Figure 37. Social media platform policies addressing bots



Source: CNA.

The consequences for violating a policy vary depending on platform, but several, including Facebook and Twitter, say they look at the specific violation and its severity, as well as the user’s history on the platform.¹²⁷ While the first violation may receive only a warning, and sometimes platforms might just take down the offending tweet or piece of content, they also typically reserve the right to ban individuals and accounts from their platforms.¹²⁸ At their most severe, policies allow companies to sue for violation of their terms.¹²⁹

¹²⁷ “Misrepresentation”; *Twitter Progress Report: Code of Practice against Disinformation*, 2019. <https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>, 15.

¹²⁸ “Misrepresentation”; *Facebook report on the implementation of the Code of Practice for Disinformation*, 2019. <https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>; “Terms of Use”; *How WhatsApp Fights Bulk Messaging and Automated Behavior*, 2019. https://chatbot.com.hr/wp-content/uploads/2019/07/WA_StoppingAbuse_Whitepaper_020418_Update.pdf.

¹²⁹ WhatsApp, <https://faq.whatsapp.com/en/android/26000259/>; Craig Silverman and Alex Kantrowitz, “Facebook The Plaintiff: Why The Company Is Suddenly Suing So Many Bad Actors,” *Buzzfeed*, Dec. 11, 2019, https://www.buzzfeednews.com/article/craigsilverman/facebook-is-suing-to-send-a-message-to-scammers-and?utm_campaign=The%20Interface&utm_medium=email&utm_source=Revue%20newsletter.

Figures

Figure 1.	Taxonomy of social media bot and botnet activities.....	ii
Figure 2.	Research approach.....	4
Figure 3.	Bot slang glossary.....	8
Figure 4.	@censusAmericans Twitter bot	11
Figure 5.	Weather Bot Twitter feed.....	11
Figure 6.	Chatbots.....	12
Figure 7.	Social media chatbots	13
Figure 8.	Social media spambot activity.....	15
Figure 9.	Botnet architecture.....	16
Figure 10.	Tweet displayed during cable news broadcast	21
Figure 11.	Indicators to identify a bot account on social media	23
Figure 12.	Spectrum of social media accounts	24
Figure 13.	Example of a cyborg click farm	26
Figure 14.	Types of bot activity	28
Figure 15.	Activity under analysis in this report.....	29
Figure 16.	Images of the national bot and OEFTracker	31
Figure 17.	Types of examples included in this section	32
Figure 18.	Example of a landscape generated by the bot @softlandscapes, a neutral example of a distribution bot.....	33
Figure 19.	Example of tweet generated by the bot @mothgenerator	34
Figure 20.	The Twitter biography of @ParityBOT, a prosocial example of a distribution bot.....	35
Figure 21.	Example of a spam bot spreading potentially malicious links.....	36
Figure 22.	Twitter account posting a fake CNN homepage showing the story of the Columbian Chemicals plant explosion.....	37
Figure 23.	Types of examples included in this section	38
Figure 24.	BBC cartoon of bot use by the Gulf countries in their diplomatic dispute.....	40
Figure 25.	Types of examples included in this section	41
Figure 26.	A tweet sent after the Kentucky gubernatorial elections alleging voter fraud.....	42
Figure 27.	Types of examples included in this section	43
Figure 28.	A tweet hijacking the #LGBTfacts hashtag from its original anti- homosexual purpose	44
Figure 29.	Types of examples included in this section	45

Figure 30.	K-Pop fan videos posted in an effort to flood the hashtag #whitelivesmatter.	46
Figure 31.	Ben Nimmo tweet flooded with responses	47
Figure 32.	Ben Nimmo tweet alerting Twitter to the botnet.....	48
Figure 33.	Activist tweet using #RompeElMiedo to spread information about police activity.....	49
Figure 34.	Incorrectly spelled Independence Day tweets	52
Figure 35.	Types of examples included in this section	53
Figure 36.	Twitter recommends incorrectly spelled hashtag.....	54
Figure 37.	Social media platform policies addressing bots.....	72

Abbreviations

AI	artificial intelligence
CT	counterterrorism
DARPA	Defense Advanced Research Projects Agency
DDOS	distributed denial-of-service
DFRL	Digital Forensic Research Lab (Atlantic Council)
DOD	Department of Defense
EU	European Union
GPC	great power competition
I&W	indications and warning
JMWC	Joint MISO WebOps Center
MILDEC	military deception
ML	machine learning
POG	psychological operations group
SECDEF	US secretary of defense
SME	subject matter expert
SOCOM	Special Operations Command
SOF	special operations forces
USG	US government
VEO	violent extremist organization

References

- Abbas, Siraj. "History and Future of Chatbots." Medium (blog), Apr. 25, 2017. <https://medium.com/@sirajea/history-and-future-of-chatbots-a1c2521f56e7>.
- Alothali, Eiman, et al. "Detecting Social Bots on Twitter: A Literature Review." *2018 International Conference on Innovations in Information Technology (IIT)*. IEEE, 2018. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8605995&tag=1>.
- "Automation Rules." Twitter Help Center. Nov. 3, 2017. <https://help.twitter.com/en/rules-and-policies/twitter-automation>.
- Azaria, Jonathan. "The Challenges of DIY Botnet Detection – and How to Overcome Them." Imperva (blog), Feb. 4, 2019. <https://www.imperva.com/blog/the-challenges-of-diy-botnet-detection-and-how-to-overcome-them/>.
- Bambauer, Derek. "How Section 230 Reform Endangers Internet Free Speech." Brookings. Jul. 1, 2020. <https://www.brookings.edu/techstream/how-section-230-reform-endangers-internet-free-speech/>.
- Berger, J.M. "How ISIS Games Twitter: The Militant Group That Conquered Northern Iraq Is Deploying a Sophisticated Social-Media Strategy." *The Atlantic*, 2014. <https://www.theatlantic.com/international/archive/2014/06/isis-iraq-twitter-social-media-strategy/372856/>.
- Bolstering Online Transparency Act. Cal. Code BPC § 17940 (2019).
- Borthwick, John. "Media Hacking." Medium (blog). Mar. 7, 2015. <https://web.archive.org/web/20150309221009/https://medium.com/in-beta/media-hacking-3b1e350d619c>.
- Boshmaf, Yazan, Ildar Muslukhov, Konstantin Beznosov and Matei Ripeanu. "The Socialbot Network: When Bots Socialize for Fame and Money." in *Proceedings of the 27th Annual Computer Security Applications Conference (ACSAC'11)*, Dec. 2011. <http://lersse-dl.ece.ubc.ca/record/272:-LERSSE-RefConfPaper-2011-008>.
- Bot Disclosure and Accountability Act of 2018. S. 3127, 115th Cong., 2018. <https://www.congress.gov/bill/115th-congress/senate-bill/3127/text>.
- Bot Disclosure and Accountability Act of 2019. S. 2125 116th Cong., 2019. <https://www.congress.gov/bill/116th-congress/senate-bill/2125?q=%7B%22search%22%3A%5B%22bot%22%5D%7D&s=1&r=2>.
- Brinkmann, Martin. "YouTube's New Commenting System Aims to Push Google+, Nothing More." ghacks.net, Sept. 25, 2013. <https://www.ghacks.net/2013/09/25/youtubes-new-commenting-system-aims-push-google-nothing/>.
- Bromwich, Jonah Engel. "Bots of the Internet, Reveal Yourselves!" *New York Times*, Jul. 16, 2018. <https://www.nytimes.com/2018/07/16/style/how-to-regulate-bots.html>.

Broniatowski, David A. et al. "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate." *American Journal of Public Health* 108, no. 10 (Oct. 1, 2018): 1378–1384. <https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2018.304567>.

Burnett, Sara. "Crackdown on 'Bots' Sweeps Up People Who Tweet Often." *AP News*, Aug. 4, 2018. <https://apnews.com/06efed5ede4d461fb2eac5b2c89e3c11>.

censusAmericans (@censusAmericans). Accessed Sept. 18, 2020. <http://www.twitter.com/censusAmericans>.

Chambers, Tim. "#FamiliesBelongTogether Robotic Attack This Week." Medium (blog), July 1, 2018. <https://medium.com/@tchambers/familiesbelongtogether-robotic-attack-this-week-c8f1425b3297>.

Chen, Adrian. "The Agency." *New York Times*. Jun. 2, 2015. <https://www.nytimes.com/2015/06/07/magazine/the-agency.html>.

Clarke, Richard D. "Statement of General Richard D. Clarke, U.S. Army Commander United States Special Operations Command." House Armed Services Committee Intelligence, Emerging Threats and Capabilities Subcommittee, Apr. 9, 2019. https://armedservices.house.gov/_cache/files/7/9/7970f176-0def-4a2d-beb3-a7d5d69e513b/9C80F888EEE40D8E82ABFF5336C012C3.hhr-116-as26-wstate-clarker-20190409.pdf.

"Code of Practice on Disinformation." European Commission. Sept. 26, 2018. <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>.

"Code of Practice on Disinformation One Year On: Online Platforms Submit Self-Assessment Reports." European Commission. Oct. 28, 2019.

Coleman, Alistair. "Analysis: Spammers and Terrorist Groups Exploit Egyptian Protest Hashtags." *BBC Monitoring*, Sept. 23, 2019. <https://monitoring.bbc.co.uk/product/c2013ul2>.

Collins, Ben and Shoshana Wodinsky. "Twitter Pulls Down Bot Network That Pushed Pro-Saudi Talking Points About Disappeared Journalist." *NBC News*, Oct. 18, 2018. <https://www.nbcnews.com/tech/tech-news/exclusive-twitter-pulls-down-bot-network-pushing-pro-saudi-talking-n921871>.

"Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—Tackling Online Disinformation: A European Approach." Eur-Lex. Apr. 26, 2018. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236>.

Confessore, Nicholas et al. "The Follower Factory." *New York Times*, Jan. 27, 2018. <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>.

Cox, Joseph. "This Bot Tweets Photos and Names of People Who Bought 'Drugs' on Venmo." Motherboard Tech by Vice (blog), Jul. 19, 2018. https://www.vice.com/en_us/article/qvmkvx/twitter-bot-venmo-buying-drugs-photo-names.

Crowell, Colin. "Our Approach to Bots and Misinformation." Twitter Blog. Jun. 14, 2017. https://blog.twitter.com/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.html.

- Diaz, Johnny. "Hurricane Dorian? On Social Media, Dorian Is Getting Misspelled All Over The Place." *SunSentinel*, Sept. 2, 2019. <https://www.sun-sentinel.com/news/weather/hurricane/fl-ne-hurricane-dorian-dorain-trending-social-media-20190902-3zdgrlaktjbanbmslca6nfuh5q-story.html>.
- Digital Forensic Research Lab, "#BotSpot: Bots Boost NFL Divide." Medium (blog), Sept. 30, 2017. <https://medium.com/dfrlab/botspot-bots-boost-nfl-divides-abec2e025ddb>.
- Digital Forensic Research Lab, "#BotSpot: The Intimidators." Medium (blog), Aug. 30, 2017. <https://medium.com/dfrlab/botspot-the-intimidators-135244bfe46b>.
- Digital Forensic Research Lab, "#BotSpot: Twelve Ways to Spot a Bot." Medium (blog), Aug. 28, 2017. <https://medium.com/dfrlab/botspot-twelve-ways-to-spot-a-bot-aedc7d9c110c>.
- Diresta, Renee. "A New Law Makes Bots Identify Themselves—That's the Problem." *Wired*, Jul. 24, 2019. <https://www.wired.com/story/law-makes-bots-identify-themselves/>.
- Distil Networks, "Bad Bot Report 2019: The Bot Arms Race Continues," Blue Cube Security. Accessed Sept. 8, 2020, <https://www.bluecubesecurity.com/wp-content/uploads/bad-bot-report-2019LR.pdf>.
- Dotto, Carlotta, and Seb Cubbon. "How to Spot A Bot (or Not): The Main Indicators of Online Automation, Co-Ordination and Inauthentic Activity." *First Draft News*, Nov. 28, 2019. <https://firstdraftnews.org/latest/how-to-spot-a-bot-or-not-the-main-indicators-of-online-automation-co-ordination-and-inauthentic-activity/>.
- D'Urso, Joey. "The Real People Pretending to Be 'Boris Bots' on Facebook." *BBC News*, Nov. 1, 2019, <https://www.bbc.com/news/blogs-trending-50218615>.
- Edgett, Sean J. "Testimony before the U.S. Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism." Oct. 31, 2017. <https://www.judiciary.senate.gov/imo/media/doc/10-31-17%20Edgett%20Testimony.pdf>.
- "A Europe that Protects: The EU Steps Up Action Against Disinformation." European Commission. Dec. 4, 2018. https://europa.eu/rapid/press-release_IP-18-6647_en.htm.
- Emerging Technology from the arXiv. "How DARPA Took On the Twitter Bot Menace with One Hand Behind Its Back." *MIT Technology Review*, Jan. 28, 2016. <https://www.technologyreview.com/s/546256/how-darpa-took-on-the-twitter-bot-menace-with-one-hand-behind-its-back/>.
- "Executive Order on Preventing Online Censorship." White House. May 28, 2020. <https://www.whitehouse.gov/presidential-actions/executive-order-preventing-online-censorship/>.
- Ferrara, Emilio, Onor Varol, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. "The Rise of Social Bots." *Communications of the ACM*, 59 (7), 96–104, Jun. 2016. Available at SSRN. <https://ssrn.com/abstract=2982515>.
- Finley, Kling. "Pro-Government Twitter Bots Try to Hush Mexican Activists." *Wired*, Aug. 23, 2015. <https://www.wired.com/2015/08/pro-government-twitter-bots-try-hush-mexican-activists/>.

- Gallagher, Erin. *Mexican Botnet Dirty Wars*. Presented at the Chaos Communication Camp 2015, Zehdenick, Germany, 2015. Retrieved from https://media.ccc.de/v/camp2015-6795-mexican_botnet_dirty_wars#video.
- Garfield, Leanna. "5 Countries That Use Bots to Spread Political Propaganda." *Business Insider*, Dec. 16, 2015. <https://www.businessinsider.com/political-bots-by-governments-around-the-world-2015-12#mexico-1>.
- Ghuman, Ajit. "Chatbots for Customer Service: Why You Need to Rethink Your Strategy." *helpshift* (blog), May 8, 2019. <https://www.helpshift.com/blog/chatbots-for-customer-service-new-approach/>.
- Gilmer, Marcus. "Twitter Can't Even Celebrate Independence Day Without Misspelling the Hashtag." *Mashable.com*, Jul. 4, 2018. <https://mashable.com/article/independence-day-twitter-hashtag-misspelled/>.
- Goldenstein, Taylor. "#Sanbernadino Has Been Shared Over 333,000 Times Even Though It's Misspelled." *Los Angeles Times*, Dec. 2, 2015. <https://www.latimes.com/local/la-me-san-bernadino-misspelled-hashtag-20151202-htlmlstory.html>.
- Gorwa, Robert, and Douglas Guilbeault. *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*. 2018. <https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.184>.
- Gupta, Krisna. "Define Botnets and Their Types?" *The Tech Win* (blog), Dec. 13, 2018. <https://thetechwin.wordpress.com/2018/12/13/define-botnets-and-their-types/>.
- Hern, Alex, and Luke Harding. "Russian-Led Troll Network Based in West Africa Uncovered." *The Guardian*, March 13, 2020. <https://www.theguardian.com/technology/2020/mar/13/facebook-uncovers-russian-led-troll-network-based-in-west-africa>.
- How WhatsApp Fights Bulk Messaging and Automated Behavior*. 2019. https://chatbot.com.hr/wp-content/uploads/2019/07/WA_StoppingAbuse_Whitepaper_020418_Update.pdf.
- Ilachinski, Andrew. *AI, Robots, and Swarms: Issues, Questions, and Recommended Studies*. CNA, 2017. DRM-2017-U-014796-Final.
- "Inauthentic Behavior." Facebook Community Standards. https://www.facebook.com/communitystandards/inauthentic_behavior.
- Indiana University Observatory on Social Media. "Botometer." <https://botometer.iuni.iu.edu/#/>.
- Is the US still at war in Afghanistan? (@OEF tracker). Accessed Sept. 14, 2020. <https://twitter.com/OEFTracker>.
- "Is This Granny a Bot? The Challenges of Detecting Automation." *Medium* (blog). Nov. 6, 2019. <https://medium.com/dfrlab/is-this-granny-a-bot-the-challenges-of-detecting-automation-115a2082b410>.
- Keller, Michael. "The Flourishing Business of Fake YouTube Views." *New York Times*, Aug. 11, 2018. <https://www.nytimes.com/interactive/2018/08/11/technology/youtube-fake-view-sellers.html>.

- Hinesley, Kara. "Lifeline's New Twitter DM Chatbot Helps Friends and Family #BeALifeline." *Twitter Blog*, Oct. 17, 2018. https://blog.twitter.com/en_au/topics/company/2018/Lifeline-launches-Twitter-DM-chatbot-to-help-BeALifeline.html.
- Kodner, Sophie. "Covert Bots: The Cyber-Nuisances Threatening our Newsfeeds and Our Democracy." *New Perspectives in Foreign Policy*, no. 13 (2017).
- Krebs, Brian. "Twitter Bots Use Likes, RTs for Intimidation." *KrebsonSecurity* (blog), Aug. 30, 2017. <https://krebsonsecurity.com/2017/08/twitter-bots-use-likes-rt-for-intimidation/>.
- Lai, Jonathan. "Is That Tweet From a Human?" *Philadelphia Inquirer*, Dec. 28, 2018. <https://www.inquirer.com/politics/nj-bot-bill-disclosure-proposal-law-social-media-free-speech-20181228.html>.
- Lukito, Josephine, and Chris Wells. "Most Major Outlets Have Used Russian Tweets as Sources for Partisan Opinion: Study." *Columbia Journalism Review* 8, Mar. 8, 2018. <https://www.cjr.org/analysis/tweets-russia-news.php>.
- Makuch, Ben. "We Talked to Phone Farmers Who Use Ad Fraud to Earn Beer Money." *Motherboard Tech by Vice* (blog), Aug. 5, 2019. https://www.vice.com/en_us/article/7xg5gq/we-talked-to-phone-farmers-who-use-ad-fraud-to-earn-beer-money.
- Marechal, Nathalie. "When Bots Tweet: Toward a Normative Framework for Bots on Social Networking Sites." *International Journal of Communication* 10 (2016), Feature 5022-5031, 2016. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2016/10/marechal.pdf>.
- Marlow, Thomas, Sean Miller, and J. Timmons Roberts. "Twitter Discourses on Climate Change: Exploring Topics and the Presence of Bots." *SocArXiv*, Feb. 26, 2020. doi:10.31235/osf.io/h6ktn.
- Martineau, Paris. "What Is a Bot?" *Wired*, Nov. 16, 2018. <https://www.wired.com/story/the-know-it-alls-what-is-a-bot/>.
- Marwick, Alice and Rebecca Lewis. "Media Manipulation and Disinformation Online." *Data & Society Research Institute, datasociety.net*, May 15, 2017. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf
- Matyszczyk, Chris. "Anti-Gay Twitter Hashtag Hijacked by Wit." *CNet*. Jan. 24, 2012. <https://www.cnet.com/news/anti-gay-twitter-hashtag-hijacked-by-wit/>.
- McBride, Megan and Zack Gold with contributions by Jonathan Schroden and Lauren Frey. *Cryptocurrency: Implications for Special Operations Forces*. CNA, 2018. https://www.cna.org/CNA_files/PDF/CRM-2019-U-020186-Final.pdf.
- McGraw, Meridith. Twitter post. 6:59AM, Nov. 22, 2019. <https://twitter.com/meridithmcgraw/status/1197847059539382272>.
- McQuaid, Julia, Jonathan Schroden, Pamela G. Faber, P. Kathleen Hammerberg, Alexander Powell, Zack Gold, David Knoll, and William Rosenau. *Independent Assessment of U.S. Government Efforts against Al-Qaeda*. CNA, 2017. CNA DRM-2017-U-015710-Final.pdf.
- Mello Jr., John P. "Headless Web Traffic Threatens Internet Economy." *E-Commerce Times*, Mar. 25, 2014. <https://www.ecommercetimes.com/story/80194.html>.

Milman, Oliver. "Revealed: Quarter of All Tweets About Climate Crisis Produced by Bots." *The Guardian*, February 21, 2020. <https://www.theguardian.com/technology/2020/feb/21/climate-tweets-twitter-bots-analysis>.

"Misinformation Is a Threat to Democracy in the Developing World." Council on Foreign Relations. Jan. 29, 2019. <https://www.cfr.org/blog/misinformation-threat-democracy-developing-world>.

"Misrepresentation." Facebook Community Standards. <https://www.facebook.com/communitystandards/misrepresentation>

"Misrepresentation"; *Facebook Report on the Implementation of the Code of Practice for Disinformation*. 2019. <https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>.

moth generator (@mothgenerator). Accessed Sept. 14, 2020. <https://twitter.com/mothgenerator>.

Mustafaraj, Eni and P. Takis Metaxas. "From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search." In *Proceedings of The World Congress on Engineering and Computer Science (WCECS 2010)*, 2010. <https://repository.wellesley.edu/object/ir122>.

Nakashima, Ellen. "U.S. Cyber Command Operation Disrupted Internet Access of Russian Troll Factory on Day of 2018 Midterms." *Washington Post*, Feb. 27, 2019. https://www.washingtonpost.com/world/national-security/us-cyber-command-operation-disrupted-internet-access-of-russian-troll-factory-on-day-of-2018-midterms/2019/02/26/1827fc9e-36d6-11e9-af5b-b51b7ff322e9_story.html.

the national bot (@the_nationalbot). Accessed Sept. 8, 2020. https://twitter.com/the_nationalbot.

Nelson, Damon Paul and Matthew Young Pollard. Intelligence Authorization Act for Fiscal Years 2018, 2019, and 2020, S. 1589, 116th Cong., 1st sess., May 22, 2019.

New Jersey Assembly. No. 4563. 218th Legislature. 2018. https://www.njleg.state.nj.us/2018/Bills/A5000/4563_I1.htm.

Nimmo, Ben. "Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It." *StopFake.org*, May 19, 2015. <https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/>.

O'Carroll, Eion. "From Russia, 'With Hastags? How Social Bots Dilute Online Speech." *The Christian Science Monitor*, Jul. 18, 2018. <https://www.csmonitor.com/Technology/2018/0718/From-Russia-with-hashtags-How-social-bots-dilute-online-speech>.

Oentaryo, Richard J. et al. "On Profiling Bots in Social Media." In *Proceedings of Social Informatics: 8th International Conference, SocInfo 2016*, Bellevue, WA, 2016, Research Collection School Of Information Systems. Available at: https://ink.library.smu.edu.sg/sis_research/3648.

Palmer, Danny. "Hackers Don't Just Want Your Credit Cards, Now They Want the Pattern of Your Life." *ZDNet*, Apr. 16, 2016. <https://www.zdnet.com/article/hackers-dont-just-want-your-credit-cards-now-they-want-the-pattern-of-your-life/>.

ParityBOT (@ParityBOT). Accessed Sept. 8, 2020. <https://twitter.com/ParityBOT>.

Pinnell, Owen. "The Online War Between Qatar and Saudi Arabia." BBC. Jun. 3, 2018. <https://www.bbc.com/news/blogs-trending-44294826>.

Plasilova, Iva, Jordan Hill, Malin Carlberg, Marion Goubet, and Richard Procee. *Study for the "Assessment of the Implementation of the Code of Practice on Disinformation."* European Commission. 2020. <https://ec.europa.eu/digital-single-market/en/news/study-assessment-implementation-code-practice-disinformation>.

"Platform Manipulation and Spam Policy." Twitter Help Center. <https://help.twitter.com/en/rules-and-policies/platform-manipulation>.

Pomerleau, Mark. "New Authorities Mean Lots of New Missions at Cyber Command." *Fifth Domain*. May 8, 2019. <https://www.fifthdomain.com/dod/cybercom/2019/05/08/new-authorities-mean-lots-of-new-missions-at-cyber-command/>.

Porup, J.M. "How Mexican Twitter Bots Shut Down Dissent." *Motherboard Tech by Vice* (blog), Aug. 24, 2015. https://www.vice.com/en_us/article/z4maww/how-mexican-twitter-bots-shut-down-dissent.

"Press Release: Freedom on the Net 2019 Reveals Crisis on Popular Platforms." Freedom House, last modified Feb. 2020. Accessed Sept. 8, 2020. <https://freedomhouse.org/report/freedom-on-the-net/2019/the-crisis-of-social-media/press-release>.

Puddephatt, Steven. "Bots x Humans: a Solution Is Needed." *Infosecurity Magazine* (web), Oct. 11, 2019. <https://www.infosecurity-magazine.com/opinions/bots-humans-solution/>.

Rauchfleisch, Adrian and Jonas Kaiser. "The False Positive Problem of Automatic Bot Detection in Social Science Research." *Berkman Klein Center Research Publication No. 2020-3*, Mar. 2020. Available at SSRN: <https://ssrn.com/abstract=3565233>.

Ritzen, Yarno. "The fake Twitter accounts influencing the Gulf crisis." *Al-Jazeera*. Jul. 21, 2019. <https://www.aljazeera.com/news/2019/07/fake-twitter-accounts-influencing-gulf-crisis-190717052607770.html>.

Roberts, Siobhan. "Who's a Bot? Who's Not?" *New York Times*, Jun. 16, 2020. <https://www.nytimes.com/2020/06/16/science/social-media-bots-kazemi.html>.

Roth, Yoel. "Automation and the Use of Multiple Accounts." *Twitter Developer Blog*. Feb. 21, 2018. https://blog.twitter.com/developer/en_us/topics/tips/2018/automation-and-the-use-of-multiple-accounts.html.

Roth, Yoel, and Nick Pickles. "Bot or Not? The Facts About Platform Manipulation on Twitter." *Twitter Blog*. May 18, 2020. https://blog.twitter.com/en_us/topics/company/2020/bot-or-not.html.

"Russian Bots Rigged Voice Kids TV Talent Show Result." *BBC News Europe*, May 16, 2019. <https://www.bbc.com/news/world-europe-48293196>.

Saltzman, Marc. "Simple Tips for Mastering Apple's Siri and Other Digital Voice-Enabled Assistants." *AARP*, Oct. 21, 2019. <https://www.aarp.org/home-family/personal-technology/info-2019/how-to-use-siri.html>.

Samuels, Elyse, and Monica Akhtar. "Are 'Bots' Manipulating the 2020 Conversation? Here's What's Changed Since 2016." *Washington Post*, Nov. 20, 2019. <https://www.washingtonpost.com/politics/2019/11/20/are-bots-manipulating-conversation-heres-whats-changed-since/>.

- “Senator Hawley Introduces Legislation to Amend Section 230 Immunity for Big Tech Companies.” Josh Hawley. Jun. 19, 2019. <https://www.hawley.senate.gov/senator-hawley-introduces-legislation-amend-section-230-immunity-big-tech-companies>.
- Silverman, Craig, and Alex Kantrowitz. “Facebook The Plaintiff: Why the Company Is Suddenly Suing So Many Bad Actors.” BuzzFeed. Dec. 11, 2019. https://www.buzzfeednews.com/article/craigsilverman/facebook-is-suing-to-send-a-message-to-scammers-and?utm_campaign=The%20Interface&utm_medium=email&utm_source=Revue%20newsletter.
- soft landscapes (@softlandscapes). Accessed Sept. 8, 2020. <https://twitter.com/softlandscapes>.
- Sonka, Joe. “Thousands of Twitter 'Bots' Targeted Kentucky With Fake News on Election Night.” *USA Today*, Nov. 11, 2019. <https://www.usatoday.com/story/news/politics/2019/11/11/kentucky-elections-2019-thousands-twitter-bots-spread-fake-facts/2564439001/>.
- “spam.” Merriam-Webster. <https://www.merriam-webster.com/dictionary/spam>.
- “The Surprising News Strategy of Pro-Russia Bots.” *BBC News*, Sept. 12, 2017. <https://www.bbc.com/news/blogs-trending-41203789>.
- Taylor, Emily, Stacie Walsh, and Samantha Bradshaw. *Industry Responses to the Malicious Use of Social Media*. NATO Stratcom Centre of Excellence. 2018. <https://www.stratcomcoe.org/industry-responses-malicious-use-social-media>.
- “Terms of Service.” YouTube. Dec. 10, 2019. <https://www.youtube.com/static?template=terms>.
- “Terms of Use.” Instagram. <https://help.instagram.com/581066165581870>.
- Twitter Progress Report: Code of Practice against Disinformation*. 2019. <https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>.
- Varol, Onur, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. “Online Human-Bot Interactions: Detection, Estimation, and Characterization,” In *Proceedings of the 11th International Conference on Web and Social Media (ICWSM '17)*, last revised Mar. 27, 2017. arXiv:1703.03107v2 [cs.SI], <https://arxiv.org/pdf/1703.03107.pdf>.
- Vincent, James. “K-Pop Stans Are Flooding Right-Wing Hashtags Like #Bluelivesmatter And #MAGA.” *The Verge*, Vox Media, Jun. 3, 2020. <https://www.theverge.com/2020/6/3/21278950/k-pop-stans-social-media-flooding-hashtags-bluelivesmatter-maga>.
- Vis, Farida. “The Rapid Spread of Misinformation Online.” *Outlook on the Global Agenda 2014: 28a–29b*, The Network of Global Agenda Councils. <http://reports.weforum.org/outlook-14/top-trends-category-page/10-the-rapid-spread-of-misinformation-online/>.
- Warner, Kelsey. “WHO Rolls Out WhatsApp Chatbot To Answer Questions and Debunk Myths about Coronavirus,” *The National*. Mar. 25, 2020. <https://www.thenational.ae/uae/health/who-rolls-out-whatsapp-chatbot-to-answer-questions-and-debunk-myths-about-coronavirus-1.997415>.

- Watts, Clint. "Testimony on Extremist Content and Russian Disinformation Online: Working with Tech to Find Solutions before the U.S. Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism." Oct. 31, 2017. <https://www.judiciary.senate.gov/imo/media/doc/10-31-17%20Watts%20Testimony.pdf>.
- Weather Bot (@tempextremes). Accessed Sept. 8, 2020. <https://twitter.com/tempextremes>.
- Weedon, Jen, William Nuland, and Alex Stamos. *Information Operations and Facebook*. Facebook. 2017. <https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf>.
- Weisman, Steve. "What Is a Distributed Denial of Service Attack (DDoS) and What Can You Do About Them?" NortonLifeLock (blog), Norton, modified Jul. 6, 2020. Accessed Sept. 8, 2020. <https://us.norton.com/internetsecurity-emerging-threats-what-is-a-ddos-attack-30sectech-by-norton.html>.
- "What Are Software Bots?" *ThinkAutomation*. Accessed Sept. 8, 2020. <https://www.thinkautomation.com/bots-and-ai/what-are-software-bots/>.
- "WhatsApp FAQ." WhatsApp. <https://faq.whatsapp.com/en/android/26000259/>.
- "WhatsApp Legal Info." WhatsApp. <https://www.whatsapp.com/legal/>.
- Woolley, Samuel C., and Philip Howard. "Computational Propaganda Worldwide: Executive Summary," Working Paper No. 2017.11, Oxford, UK: Project on Computational Propaganda. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf>.
- Yang, Kai-Cheng, et al. "Arming the Public with Artificial Intelligence to Counter Social Bots." *Human Behavior and Emerging Technologies* 1.1 (2019): 48-61. arXiv:1901.00912v2 [cs.CY].
- Zakem, Vera, Megan McBride, and Kate Hammerberg. *Exploring the Utility of Memes for U.S. Government Influence Campaigns*. CNA. 2018. https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf.
- Zuckerberg, Mark. "Mark Zuckerberg: The Internet Needs New Rules. Let's Start in These Four Areas." *Washington Post*, Mar. 30, 2019. https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html.

This report was written by CNA's Strategy, Policy, Plans, and Programs Division (SP3).

SP3 provides strategic and political-military analysis informed by regional expertise to support operational and policy-level decision-makers across the Department of the Navy, the Office of the Secretary of Defense, the unified combatant commands, the intelligence community, and domestic agencies. The division leverages social science research methods, field research, regional expertise, primary language skills, Track 1.5 partnerships, and policy and operational experience to support senior decision-makers.

CNA is a not-for-profit research organization that serves the public interest by providing in-depth analysis and result-oriented solutions to help government leaders choose the best course of action in setting policy and managing operations.



Dedicated to the Safety and Security of the Nation

DRM-2020-U-028199-Final

3003 Washington Boulevard, Arlington, VA 22201

www.cna.org • 703-824-2000