10

Social Bots for Peace: Combating Automated Control with Automated Civic Engagement?

Samuel Woolley and Mark Kumleben

Introduction

In the lead up to the 2020 US presidential election, a series of suspicious social media accounts made a series of inflammatory political posts across Facebook, Twitter, and Instagram. The content they shared 'focused almost exclusively on racial issues in the US' but sometimes touched upon LGBTQ issues and political content aimed at other US social groups (Ward et al., 2020). Specifically, the accounts worked to engage the Black Lives Matter community. One, @The_black_secret, 'posted a video of a racial incident with the comment "Blacks have a right to defend

S. Woolley (✉)
University of Texas at Austin, Austin, TX, USA
e-mail: woolley@utexas.edu

M. Kumleben
Digital Intelligence Lab, Institute for the Future, Palo Alto, CA, USA
e-mail: mark@kumleben.com

M. A. Naseem, A. Arshad-Ayaz (eds.), *Social Media as a Space for Peace Education*,
Palgrave Studies in Educational Media,
https://doi.org/10.1007/978-3-030-50949-1_10

themselves against Racism" that drew more than 5000 reactions and more than 2000 shares'.

The accounts purported to be run by people living around the United States, with location tags in California, Indiana, North Carolina, and several other states. But an in-depth investigation by CNN and Clemson University—with corroboration from Facebook and other social media firms—revealed that the accounts were not, in fact, tweeting from within the United States (Gleicher, 2020). These accounts were based at troll farms in Ghana and Nigeria. Facebook later revealed that these African propaganda firms were supported by 'individuals associated with past activity by the Russian Internet Research Agency' (IRA). The digital content from just 200 accounts created by the Ghanaian trolls 'reached hundreds of thousands, perhaps millions, of people worldwide' (Ward et al., 2020). US Senator Mark Warner, in response to the CNN story, said:

> I've said for years now that it would be foolish to believe there was only the one well-publicized IRA facility in St. Petersburg. This new reporting is a reminder of the continuing threat we face from Russia and its continuing efforts to divide and manipulate us on social media. (Ward et al.)

Warner's comments echo a deep public skepticism that has arisen around the political use of social media. The activities of the African trolls, leveraging novel tactics and technology to sow manipulative information during the 2020 contest, are an evolved version of the types of online behaviors that have caused this skepticism.

Just four years earlier, Warner and other US government officials witnessed broadscale attempts, social and technologically, to 'hack' the 2016 US election (Sanger & Edmondson, 2019). The Russian government, alongside groups including Cambridge Analytica, leveraged platforms like Facebook and Twitter in efforts to alter the flow of information during a pivotal election. Russia's IRA and many of the other players using underhanded tactics to spread information on social media leveraged computational propaganda—they leveraged automated and algorithmically managed tools over social media in attempts to manipulate public opinion (DiResta et al., 2018; Woolley & Howard, 2016).

One of the primary tools used by the Russian IRA, but also by countless other organizations and actors involved in 'information operations' over social media, is the political bot. A bot, simply put, is a piece of automated software used to do an online task people would otherwise have to do. These automated pieces of code are useful at scaling tasks—including sifting through massive troves of data—achieving excavation and analyses that would otherwise take a person or group months or years to finish. The IRA and other computational propagandists most often use socially-oriented bots, which can be built to mimic real people, in efforts to amplify particular content while suppressing others. Some have argued that the use of computation, automation, social media, and other emergent tools for these means and ends amounts to a new form of information warfare (Singer & Brooking, 2018).

But not all bots capable of operating over social media are used to pose as real people in efforts to deceive. Not all social automation online is leveraged by governments hoping to sow polarization and confusion amongst their adversaries. There are also examples of bots being harnessed for all manner of civically beneficial uses. Journalists have used bots to automate various aspects of reporting over Twitter (Lokot & Diakopoulos, 2015), watchdog groups have built social media bots that track and report on politicians' edits to Wikipedia (Ford, Puschmann, & Dubois, 2016), social commentators have created bots to automatically generate critique (Woolley et al., 2016). In these and other, circumstances social bots can also be used as social scaffolding or civic prostheses—deployed as tools for supporting newmaking, political monitoring, or—more broadly—forms of democratic engagement (Hwang, Woolley, & Borel, 2016).

How might bots be used for peace? For peace education? Do the negative political uses of bots overshadow their possible usefulness for diplomacy, mediation, and other modes of avoiding violent conflict? Are the ways in which bots have previously deployed for the purposes of civic engagement useful or transferable to parallel uses of social automation for peace? Because socially-oriented bots are such useful tools for amplifying or augmenting conversations about a given topic, and because more general versions of bots can scale data analysis work, it is very much possible that these automated tools have potential for peace education.

This paper delves into the political uses of social media bots in order to illuminate the ways these tools might be harnessed for the purposes of peace education. First, we explore the panoply of negative political instances of social media bot usage by discussing recent debates and trends in the academic literature on computational propaganda. By doing so, we hope to illuminate paths that those hoping to leverage bots for more democratically beneficial ends should avoid. We do not, for instance, recommend that bots be used to spread 'peace spam'. We do, however, explore this research in the hope that it will educate readers about the various potentials—good, bad, and ugly—of social media bots. Furthermore, this initial section underscores the kinds of challenges faced by people hoping to leverage social media and bots, for peace education.

In the subsequent section, we explore two cases in which social media bots have already been used for purposes related to peace education. We use insights from these cases to then discuss the broader ways in which social media bots might be used for peace education and to argue for the need for more research on beneficial bots. Finally, we conclude with a section reflecting on the possibilities of automation and social media for peace education, tying research on computational propaganda and computational work for the benefit of democracy to the larger goals of peace education in the digital era.

Bots and Computational Propaganda

Social bots, automated programs that emulate human communication and interaction online, have become widespread in recent years (DiResta et al., 2018; Woolley, 2016). Although they have existed online for decades, these automated social actors are now commonly used interactive tools in customer service, journalism, healthcare, and entertainment. On social media, social bots have been used for the purposes of humor, art, commentary, and sports. Since events like the 2016 US presidential election, the 2020 US election, and the 2020 Covid-19 crisis, in which social bots were used over Twitter, Facebook, and other platforms to spread conspiracy and disinformation, they have become

well known for their role in political communication and manipulation campaigns (Badawy, Ferrara, & Lerman, 2018; Ko, 2020; Woolley & Guilbeault, 2017).

Political bots, social bots built explicitly to engage on political topics over social media, are a continued cause for concern during elections, security crises, and disasters (Ferrara, Varol, Davis, Menczer, & Flammini, 2016; Metaxas & Mustafaraj, 2012; Woolley & Howard, 2016). They are among the core tools of organizations and actors hoping to manipulate public opinion, to alter trending content, to change search results, and promote other forms of digitally prioritized content, during pivotal moments in which people are seeking information to make important decisions about public and private life. Bots are one mechanism used in computational propaganda campaigns, automated and algorithmically driven efforts to alter information flows and impact behavior (Woolley & Howard, 2018).

The use of social bots to manipulate political conversation is now a global phenomenon. A global analysis of state-sponsored computational propaganda and digital disinformation efforts (Bradshaw & Howard, 2019) found that over 70 countries now use automated tactics as part of their online communication strategy. Computational propaganda is among the most significant political consequences of the latest innovations in social media. Social media companies are still working to respond to the problem of political bot usage and other forms of computational propaganda. In 2014, Twitter acknowledged that as many as 23 million active users on its platform were bots (Motti, 2014). With the company's internal attempts to combat these efforts, alongside bot-makers' efforts to launch new automated profiles, these numbers have waxed and waned. The use of political bots on that platform, however, has continued—with Twitter bots playing a role in conversations about contentious events in 2018, 2019, and 2020 (Avaaz & ISD, 2019; Barojan, 2018; Murdock, 2020) We hypothesize that, in years ahead, if a country is caught up in a crisis or highly competitive election, then those in control or vying for control will continue use bots to manipulate public opinion.

As a worldwide phenomenon, computational propaganda has affected politics in every region of the globe. The Oxford Internet Institute has uncovered organized manipulation by parties, governments, and foreign

actors in democracies and authoritarian regimes alike (Bradshaw & Howard, 2019). Twitter bots were spotted particularly early in Mexico (Alfonso, 2012), earning themselves the nickname 'Penabots' for their tendency to support the presidential campaign of Enrique Pena Nieto. From there, professional bot-masters like Andres Sepulveda spread the practice across Latin America, with Sepulveda alone claiming to have run bot networks targeting conversations during political events in nine countries in the region (Robertson, 2016). Brazilian researchers have found that bots have become a regular fixture on the Brazilian internet, active not only around election times but also around critical political events such as the impeachment of President Rousseff (Arnaudo, 2018).

Eastern Europe has also been hard-hit by computational propaganda, in part due to Russian interest in the unstable region. Ukraine has been targeted by an organized, highly disciplined computational propaganda campaign involving both automated accounts and human trolls (Bugorkova, 2015)—a dangerous threat to a deeply divided country, of a type which should particularly concern peace educators. The Polish online community faces similar difficulties, riven by internal 'troll wars' and attacked from without by Russian computational propaganda (Gorwa, 2017). In Asia, South Korea's elections have been attacked by members of their own intelligence service (Sang-hun, 2013), and China has added computational propaganda to their horde of online propagandists (@DFRLab, 2019). Chinese efforts do not limit themselves to explicitly political issues, but have also included automated propaganda around the COVID-19 pandemic (Bechis, 2020).

Many countries targeted by computational propaganda are of particular concern to peace activists, because of political, ethnic, or religious tensions which could escalate into violence. Computational propaganda is increasingly common in Nigeria (Funke, 2018), South Africa (Zille, 2018), and Kenya (Portland Communications, 2018)—countries with histories of religious, ethnic, or election violence. Peace educators must pay particular attention to those countries at risk of violence, in which computationally amplified disinformation could spark serious incidents such as communal lynchings or election riots.

Political bots, suppress free expression and civic innovation via the demobilization of activist groups and the suffocation of democratic free

speech. With this in mind, political bots and their effects must be better understood for the sake of mediated civic engagement, digital literacy, and democracy.

Bots and Peace Education

Bots are an expansive category of computer programs, and social bots potentially reflect the full variety of our online lives. There is no inherent reason that bots need be deceptive or harmful, and many bot-makers build bots for politically neutral or socially beneficial purposes. These bot-makers themselves come from a wide variety of backgrounds and motivation, far beyond the stereotypical shadowy propagandists or sleazy con-artists. An ethnographic survey of 40 bot-makers found makers from 'a diverse range of professional, demographic, and cultural backgrounds' (Woolley, Shorey, & Howard, 2018, p. 8). Anybody with a message to promote can do so using social bots, but many responsible bot-makers carefully consider the ethical implications of the manner in which they do so. Even relatively simple bots can engage in complex and unpredictable emergent interactions, with potentially negative consequences (Tsvetkova, García-Gavilanes, Floridi, & Yasseri, 2017). Furthermore, activists and researchers alike have responsibilities to the people their bots affect, and must work to develop best practices which will minimize potential harm.

Not all bots need be political. Human beings engage themselves in many fields of activity, from the commercial to the artistic, without explicitly considering politics. Bots may facilitate any of these activities without acting in a problematic manner. Art bots have become relatively common on social media, both as distributors of existing art and creators of new content. Some bots randomly post images from the collections of institutions like the Metropolitan Museum of Art or the Museum of Modern Art (Sharma, 2020), or images of a particular genre. In an era of information overload, where museum collections appear limitless and appreciators of art may feel lost in endless catalogs, art bots provide an opportunity for serendipitous discovery. Bots can also create new art, particularly in mediums like poetry. The idea of automated or

computational poetry dates back far before the internet, finding its most well-realized expression in the work of Oulipo members such as Raymond Queneau (Gallix, 2013). Liam Cooke (2014) has built 'Poem.exe', which builds on Queneau's *A Hundred Thousand Billion Poems* to automatically generate combinatorial poetry and post the resulting works to Twitter (@Poem_exe). The N+7 bot (@n7bot) applies an Oulipian technique known as 'N+7' to various texts of public interest, including Donald Trump's tweets. As text and image synthesizers mature technologically, art bots will become an ever more popular channel for the creation and distribution of computational art.

The power of bots to alert or remind social media users is not limited to discovering art. Bots can be used to discover deals, such as the Hundred Zeros bot (@hundredzeros), which works to inform Twitter users of free eBooks for the Amazon Kindle. Many deal-hunting sites use bots to post bargains, and chatbots can even find recommendations tailored to a user's needs and budget (Escribano, 2016). News bots can monitor public data sources and inform followers of events and issues which would otherwise go unnoticed, from Congressional Wikipedia edits to stop-and-frisk incidents (Fader, 2016). The repetitive nature of these bots draws attention to the ongoing nature of the political problems they highlight, and can humanize what would otherwise be sterile statistics. Activists in many countries have created bots to monitor Wikipedia edits from political institutions like Congressional offices, providing an example of the ability of public-interest bots to scale globally (Cox, 2014). Bots can also inform the public of natural disasters (@earthquakesLA) and alert researchers of relevant social media activity to predict disease outbreaks (Scherer, 2014). It would be a mistake to dismiss the ability of bots to spread socially beneficial information—they are just as effective with helpful information as with deceptive content.

Bots are not limited to posting on their own, but may also interact with users proactively and directly. Reddit uses bots extensively to moderate communities and enforces rules automatically—the Reddit automoderator bot alone provides services for over 9500 smaller communities (reddit.com, n.d.), providing services such as minimum reputation requirements, spam removal, and counter-misinformation efforts. Listener bots are another potential source of positive engagement

(Woolley et al., 2018). These bots use hashtags or keywords to find users and engage with them in a targeted manner. Bots can be used to intervene against racist speech online (Munger, 2016), responding to racial slurs to remind users of the offense they cause. Listener bots are not limited to simple responses, such as the pre-formulated messages offered by the 'Drop the I' bot (Hogan, 2018). Rather, they automate a time-consuming step in the communications process—generating initial engagement. Listener bots can be used by activists to find users willing to engage, with a human user 'stepping into' the account once it has successfully communicated with another user, much in the manner of the semi-automated customer service chatbots used by some online companies. Direct, proactive interaction between social bots and users is an emerging field, and one whose ethics merit careful consideration (Krafft, Macy, & Pentland, 2017), but it offers great opportunities to change views and work towards healthier communication online.

Case Study: Build Up's Online Peace Initiatives

Our first case study is the work of Build Up, a non-profit organization devoted to peace innovation. Build Up has explored the use of bots to spread peace messages and engage in anti-polarization outreach, and the lessons of their work are useful for future peace education initiatives. For International Peace Day 2017, in collaboration with International Alert, Build Up launched its #robotsforpeace program to promote Peace Day (Build Up, 2017). Robots for Peace encouraged programmers and non-programmers alike to build their own Twitter bots to 'build a flock of robots' spreading peace messaging around the hashtag #peaceday. The program's website asked users to follow certain automation best practices for ethical reasons, including not posting spam and clearly identifying the account as a bot.

It appears that many users did create bots to spread #peaceday, and some are still active as of 2020, such as @giadapeacebot and @unite4peacebot. Although a laudable initiative, promoting both Peace Day itself and computer literacy among the peace community, Robots for Peace had certain limitations. The project itself was transient—its peacebots.org

website is now only available through online archives, and it appears to have received little traction in news media. Because of the decentralized nature of the project, it was difficult to measure the engagement and to ensure compliance with ethical best practices. For instance, some of the #peaceday bots do not identify themselves as bots, such as @saxby28 and @manutd83725831. Build Up acknowledged that the 'megaphone' approach, where bots simply spread a pro-peace message far and wide, has limitations, noting that 'there is a fine line between amplifying a message so it receives the attention we believe it deserves (as we are trying to do) and manufacturing consensus to a point where it loses credibility' (Laurrari, 2017).

Build Up did not stop, however, with Robots for Peace. Rather, they also developed a more complex and ambitious project called The Commons. This program would develop bots to engage Americans interested in the issue of political polarization, using targeted messages and automated conversations in order to connect interested subjects with a human facilitator (Build Up, 2018). On Twitter, The Commons mapped and targeted both liberal and conservative hashtags, responding to users posting those hashtags with questions about polarization. Those who responded positively were connected to humans. On Twitter, 50% of users who responded positively to a bot were willing to talk to a human, and while the proportion was lower on Facebook (using targeted ads), it was still more effective than other forms of cold approaches. A second run of The Commons systematically evaluated its messages to determine the effectiveness of various approaches, and was able to dial in on messages most likely to achieve positive responses (Build Up, 2019a).

By using a 'funnel of engagement' model, The Commons was able to deploy bots where they were most effective—at the top of the funnel, where the difficulty for peace educators is to cost-effectively contact large numbers of people and to identify which of them are interested in learning about peace issues (Build Up, 2017, p. 35). Bots appear to perform well compared to traditional approaches in raising awareness and gathering interest, but humans are still necessary to properly engage with subjects and move them to action. These humans can be helped by automated approaches, such as the dynamic identification of political hashtags, but both for ethical and practical reasons, should identify themselves as

separate from bot accounts. Build Up found that humans messaging from a separate personal account helped to overcome users' distrust of spam.

Generally, Build Up took care to use bots ethically in this project. They used identifiable bot avatars and did not accidentally spam or harass users. In addition to the ethical necessity of avoiding deception or spam, the use of targeted and light-touch automated messaging helps avoid 'bot fatigue', where users may become frustrated with peace bots or feel that they are being manipulated by peace educators (Laurrari, 2017). On the other side of the coin, though, Build Up also noted that some users respond positively to the novelty of bots, as users who agree with pro-peace positions may be excited to see peace educators using a new tool for good. This novelty will not last forever, particularly as commercial bots become more prevalent, but is a useful advantage for peace educators who intend to deploy bots in the near future.

The examples of Robots for Peace and The Commons should remind us of the importance of ethical considerations. When deploying bots, experimental ethics are not just worthwhile in themselves but will also help to win the trust of those contacted by bots. Bot initiatives should be carefully monitored both for ethics and efficacy, so that their messaging can be tailored towards maximum engagement as The Commons was. Bots are an interesting and innovative way to reach social media users, but to realize their full potential they should be deployed in a full-spectrum approach alongside targeted ads and human facilitators.

Case Study: Botivist

Activists for social causes face many hurdles and pressures in their work. Recruitment of new core activists and friendly volunteers is one major issue—recruitment is time-consuming, it can add financial costs, it's difficult to standardize, it can be tough to evaluate the effectiveness of a message, and public recruiting can even be physically dangerous (Savage, 2016). Bots could, in theory, help with all of these issues, and a group of researchers and activists have trialed that approach with Botivist. Botivist is a system to recruit volunteers over Twitter, using a bot which tags

relevant users and asks for their input. The Botivist team not only developed a working model, but applied it to a specific issue and evaluated different forms of messaging from Botivist (Savage, Monroy-Hernández, & Hollerer, 2016).

Though it can be adapted to many issues, Botivist began by tackling the issue of corruption in Latin America, tweeting in Spanish about the problem. Users who had recently tweeted a corruption-related term were tagged together in a tweet with one of four approaches: direct, gain, loss, and solidarity (all included a call to action, but the latter three included an extra sentence designed to persuade users). Unlike communication from humans, a direct call to action was the most effective, getting responses from 30% of users targeted. This relatively high rate of response, compared to other communication techniques, suggests that Botivist and similar programs may be very productive in helping activists deliver calls for action to potential volunteers.

The creators of Botivist were naturally concerned with the differences from a human approach a bot recruiter would require. Purely direct, rather than persuasive calls to action were most successful, suggesting that users appreciate bots which act like tools for humans, rather than imitating humans themselves. Most negative responses to Botivist used the term 'bot', and this was more common for approaches which included 'humanizing' sentiments alongside the direct call to action (Savage et al., 2016, p. 6). Researchers concluded that users are uncomfortable with bots which act 'too human', and were more likely to ask if the use of bots in activism was appropriate. For ethical reasons, Botivist accounts were clearly identified as bots, but it appears this honesty also has practical benefits in gaining user trust. One aspect of Botivist they also noted was the population of users who responded, as classified by the hashtags they previously tweeted. Users who regularly tweeted activism-related terms were more likely to respond, as well as a small cluster of users interested in online marketing terms. This latter, unexpected aspect tallies with Build Up's findings that some users are excited by the novelty of bots. It also bodes well for the acceptance of peace bots among younger generations who are more accustomed to an internet shaped at every level by online marketing.

Botivist's case implies some important suggestions for peace educators. A deceptive or manipulative approach, even besides ethical concerns, is likely to be rejected by savvy users. With social bots widely considered a problem, peace bots must be designed with ethical considerations at the forefront. However, Botivist shows that simple and transparent bots are still effective—perhaps more so than bots which attempt to act like humans. An account which appears to be a 'bot proxy' (Woolley, Shorey, & Howard, 2018) for human activists will be better received than a bot which pretends to be human itself. Furthermore, bots are most useful when they are targeted at those who are already interested in receiving their message. This makes them suited to delivering calls to action for a population who have existing awareness of social issues, moving further down the funnel of engagement than The Commons did. Furthermore, with the careful data collection and analysis performed by the Botivist team, activists can standardize and A/B test their messaging in a way which is difficult to do with human communicators. Botivist itself may be adaptable to many issues, but the use of bot messaging allows activists to discover the particular communication strategies best suited for their area of focus. It should be heartening to peace educators to see, through the experience of Botivist, that the most ethical strategies for automated activism are also the most effective.

Conclusion: Bots as a Tool for Peacebuilders

Automated systems and artificial intelligences will soon be the foremost drivers of technological progress. Social media and search algorithms already determine the news diet of billions, and propaganda bots show little sign of abating their subversive activity. With this in mind, peacebuilders must consider the potential of automation. The UN is already researching AI for peace (Masood & Waehlisch, 2019), but they cannot go it alone. Bots are a logical way for smaller organizations or individual peace educators to multiply their efforts—they may have 'democratized propaganda' (Woolley & Guilbeault, 2017), but they have also democratized peace education.

However, there are certain questions which advocates of peace bots must answer, and a shotgun approach to building bots may do more harm than good. Initial suggestions for peace bots focused on the possibility of mirroring propaganda bot campaigns, using bots to spread pro-democracy messages (Howard, 2012). However, this was criticized as likely ineffective in comparison to humans (York, 2013), and initiatives like Robots for Peace failed to gain traction. Peacebuilders have explicitly rejected 'replicating a strategy that could also be used to mindlessly share hate or misinformation' (Laurrari, 2017). Happily for peace educators, developing bot technology has now made better approaches possible.

If the 'megaphone' approach, where bots mindlessly boost pro-peace content, is ineffective, a more sophisticated effort must be necessary. Bots are not merely autonomous marketing tools, but can empower human activists to better spread human messages. In the case studies above, bots were most effective when they were used to relieve existing burdens on activists, particularly in raising awareness and initial engagement. Peace educators already know well the issues they face in public communication—those known areas of friction are the obvious place to start, from which further innovation will emerge naturally as peacebuilders become accustomed to using bots.

Rather than simply promoting existing content, peace bots can move interested citizens down the 'funnel of engagement' (Build Up, 2019b) to inspire pro-peace action. Furthermore, this type of bot is ethical by design, as certain critical ethical principles are also necessary for them to be most effective. In the case of Botivist, peace bots benefited from targeting users already interested in activism and from clearly acting like bots rather than imitating humans (Savage et al., 2016). Peace educators should consult ethical best practices like those published by JustPeace Labs (Easterday, 2018) as a matter of principle, but they can also do that with confidence that the public will appreciate their ethical stance and respond positively.

When considering the future of peace bots, there are certain existing trends and foreseeable problems we should take into account. The first is the most foundational of peace work: the importance of context. As the saying goes, all politics is local, and this is particularly true for peace education which aims to convince individuals of the value of peace.

Peacebuilders must be informed about local issues and fluent in local terms if they hope to appropriately target bots and spark conversations. The PeaceTech Lab has built many lexicons of online hate speech terms in various countries (PeaceTech Lab, n.d.), and researchers have suggested using natural language processing to continually detect dangerous speech (Masood & Waehlisch, 2019). Next, there is the importance of targeting the right individuals, to find those who are open to peace education. Considering that members of the online marketing community appear interested in programs like Botivist (Savage et al., 2016), it is likely possible to adapt existing marketing techniques or even to approach marketers for *pro bono* contributions.

Tech companies, too, could get involved, and a silver lining of their ad-targeting surveillance might be its use for peace. Finally, there is the issue of 'bot fatigue' (Laurrari, 2017) and the risk of 'peace spam'. Peace educators should coordinate to make responsible use of our information commons, avoiding unnecessary duplication of messaging. Tech companies, too, could get involved, providing more API access and spam-reducing features to peacebuilders. For instance, as the Botivist team suggested, tech companies could include features to warn bots if a user has recently been contacted by a similar bot. Organizations using bots should also consistently track and evaluate the performance of their messaging, since making communications more efficient will reduce bot spam.

Researchers, too, should take an interest in the relationship between bots and peacebuilding. There are many potentially fruitful avenues for research which would help promote good practices and fight manipulative bots. Plans for further study might incorporate looking into how certain cases of political bot usage in one country may have affected implementation and usage in other countries. Peace organizations would benefit from the lessons learned in other countries, like Botivist in Bolivia. Another project could lie in the building of a prediction model of bot usage in upcoming international elections.

This year there are numerous highly contested international elections scheduled. Several of these are taking place in countries with authoritarian regimes and emerging democracies. It would be interesting to work towards predicting political bot usage in these upcoming elections and determine what potential impact such use has on electoral outcomes.

Data collected could also be used to develop early warning systems, either for bot use itself or for violence fomented by online speech. Compared to other areas of international development, peacebuilding is poorly funded (Vernon, 2017), and researchers can help activists get more bang for their buck by studying the most efficient forms of automated communication. When resources are tight, we should pick the low-hanging fruit first.

Eventually, it may be possible for researchers and peacebuilders together to develop a 'white-label' peace communications bot system, as Botivist hopes to do, which could be easily adapted to local contexts and deployed dynamically based on the needs of peacebuilders. Adversarial peace bots are one possibility, which would seek out and rebut propaganda bots as they post disinformation. So are chatbots which could be summoned by social media users to help support pro-peace arguments. Once we have developed institutional knowledge and established best practices, vast possibilities for peace bots may open up. It is easier to destroy than to build, and propaganda botnets use cruder and cheaper techniques than are necessary for peace education, but that is all the more reason to develop sophisticated, efficient, and ethical peace bot systems.

References

Alfonso, F. (2012, May 21). Twitter bots silence critics of Mexico's leading presidential candidate. *The Daily Dot*. Retrieved from https://www.dailydot.com/news/pena-nieto-twitter-bots-mexico-election/

Arnaudo, D. (2018). Brazil: Political bot intervention during pivotal events. In S. C. Woolley & P. N. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (pp. 128–152). Oxford University Press.

Avaaz & ISD. (2019). *Disrupted: Evidence of widespread digital disruption of the 2019 European Parliament elections* (p. 4).

Badawy, A., Ferrara, E., & Lerman, K. (2018). Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 258–265. https://doi.org/10.1109/ASONAM.2018.8508646

Barojan, D. (2018, April 20). #BotSpot: Bots target Malaysian elections – DFRLab – Medium [Blog]. *Medium / DFRLab*. Retrieved from https://medium.com/dfrlab/botspot-bots-target-malaysian-elections-785a3c25645b

Bechis, F. (2020, March 31). How China unleashed Twitter bots to spread COVID-19 propaganda in Italy. *Formiche.net*. Retrieved from https://formiche.net/2020/03/china-unleashed-twitter-bots-covid19-propaganda-italy/

Bradshaw, S., & Howard, P. N. (2019). *The global disinformation order: 2019 global inventory of organised social media manipulation* (p. 24). Oxford University, Project on Computational Propaganda.

Bugorkova, O. (2015, March 19). Ukraine conflict: Inside Russia's "Kremlin troll army." *BBC News*. Retrieved from http://www.bbc.co.uk/news/world-europe-31962644

Build Up. (2017). *#Robotsforpeace – We're building a flock of robots to share messages of peace on Twitter on International Peace Day 2017.* Retrieved from https://web.archive.org/web/20190227112341/http://www.peacebots.org/

Build Up. (2018, February 28). Building the commons. *Medium*. Retrieved from https://medium.com/@howtobuildup/building-the-commons-dc60e6ee7b69

Build Up. (2019a). *The commons: An intervention to depolarize political conversation on Twitter and Facebook in the USA*. Retrieved from https://howtobuildup.org/wp-content/uploads/2019/12/TheCommons-2019-Report_final.pdf

Build Up. (2019b, December 14). Scaling the commons. *Medium*. Retrieved from https://medium.com/@howtobuildup/scaling-the-commons-969b15c98012

Cooke, L. (2014). *Poem.exe*. Retrieved April 7, 2020, from https://poemexe.com/

Cox, J. (2014, July 10). These bots Tweet when government officials edit Wikipedia. *Vice*. Retrieved from https://www.vice.com/en_us/article/pgaka8/these-bots-tweet-when-government-officials-edit-wikipedia

DFRLab. (2019, September 4). Twitter's Hong Kong archives: Chinese commercial bots at work. *Medium*. Retrieved from https://medium.com/dfrlab/twitters-hong-kong-archives-chinese-commercial-bots-at-work-f4c7ae8eea64

DiResta, R., Shaffer, D. K., Ruppel, B., Matney, R., Fox, R., Albright, D. J., et al. (2018). *The tactics & tropes of the internet research agency* (p. 101). New Knowledge.

Easterday, J. (2018, March 27). *Ethical guidelines for PeaceTech – JustPeace Labs*. Retrieved from https://justpeacelabs.org/ethical-guidelines-for-peacetech/

Escribano, J. (2016, November 7). How we wrote a Wirecutter-like bot with Reply.ai in just a few hours. *Medium*. Retrieved from https://chatbotslife.

com/how-we-wrote-a-wirecutter-like-bot-with-reply-ai-in-just-a-few-hours-59e7ddfa242c

Fader, L. (2016, February 26). A brief survey of journalistic Twitter bot projects. *Medium*. Retrieved from https://points.datasociety.net/a-brief-survey-of-journalistic-twitter-bot-projects-109204a8d585

Ferrara, E., Varol, O., Davis, C. A., Menczer, F., & Flammini, A. (2016). The rise of social bots | July 2016 | Communications of the ACM. *Communications of the ACM, 59*(7), 96–104.

Ford, H., Puschmann, C., & Dubois, D. (2016). Keeping Ottawa honest, one tweet at a time? Politicians, journalists, Wikipedians and their Twitter bots. *International Journal of Communication, 10*(Special Issue), 20.

Funke, D. (2018, November 30). Nigeria is the next battleground for election misinformation. *Poynter*. Retrieved from https://www.poynter.org/fact-checking/2018/nigeria-is-the-next-battleground-for-election-misinformation/

Gallix, A. (2013, July 12). Oulipo: Freeing literature by tightening its rules. *The Guardian*. Retrieved from https://www.theguardian.com/books/books-blog/2013/jul/12/oulipo-freeing-literature-tightening-rules

Gleicher, N. (2020, March 12). Removing coordinated inauthentic behavior from Russia. Facebook. Retrieved from https://about.fb.com/news/2020/03/removing-coordinated-inauthenticbehavior-from-russia/

Gorwa, R. (2017). Computational propaganda in Poland: Unpacking the ecosystem of social media manipulation. In S. Woolley & P. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (pp. 86–103). Oxford Internet Institute, University of Oxford.

Hogan, J. R. (2018, July 28). We built a Twitter bot that corrects people who say 'illegal immigrant'. *Splinter*. Retrieved from https://splinternews.com/we-built-a-twitter-bot-that-corrects-people-who-say-ill-1793849536

Howard, P. N. (2012, December 31). Let's build pro-democracy Twitter bots. *Slate Magazine*. Retrieved from https://slate.com/technology/2012/12/twitter-bots-for-democracy-could-combat-authoritarian-governments-50-cent-social-media-armies.html

Hwang, T., Woolley, S., & Borel, B. (2016, March 8). How politicians should and shouldn't use Twitter bots. *Slate*. Retrieved from http://www.slate.com/articles/technology/future_tense/2016/03/how_politicians_should_use_twitter_bots.html

Ko, R. (2020, March 1). Meet 'Sara', 'Sharon' and "Mel": Why people spreading coronavirus anxiety on Twitter might actually be bots. *The Conversation*. Retrieved from https://theconversation.com/meet-sara-sharon-and-mel-why-people-spreading-coronavirus-anxiety-on-twitter-might-actually-be-bots-134802

Krafft, P. M., Macy, M., & Pentland, A. (2017). Bots as virtual confederates: Design and ethics. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing – CSCW '17*, 183–190. doi:https://doi.org/10.1145/2998181.2998354

Laurrari, H. P. (2017, September 12). Automatic for the peaceful. *Medium*. Retrieved from https://medium.com/@howtobuildup/automatic-for-the-peaceful-52d24d1d23e

Lokot, T., & Diakopoulos, N. (2015). News bots: Automating news and information dissemination on Twitter. *Digital Journalism*, 1–18. https://doi.org/10.1080/21670811.2015.1081822

Masood, D., & Waehlisch, M. (2019, March 28). Future wars will be waged with robots. But so might future peace. *Euronews*. Retrieved from https://www.euronews.com/2019/03/28/future-wars-will-be-waged-with-robots-but-so-might-future-peace-view

Metaxas, P. T., & Mustafaraj, E. (2012). Social media and the elections. *Science, 338*(6106), 472–473. https://doi.org/10.1126/science.1230456

Motti, J. (2014, August 12). Twitter acknowledges 23 million active users are actually bots. *Tech Times*. Retrieved from http://www.techtimes.com/articles/12840/20140812/twitter-acknowledges-14-percent-users-bots-5-per-cent-spam-bots.htm

Munger, K. (2016, November 17). This researcher programmed bots to fight racism on Twitter. It worked. *Washington Post*. Retrieved from https://www.washingtonpost.com/news/monkey-cage/wp/2016/11/17/this-researcher-programmed-bots-to-fight-racism-on-twitter-it-worked/

Murdock, J. (2020, April 3). Twitter purges thousands of bot accounts that used coronavirus in political propaganda campaign. *Newsweek*. Retrieved from https://www.newsweek.com/twitter-purges-thousands-bot-accounts-coronavirus-political-propaganda-campaign-1495951

PeaceTech Lab. (n.d.). *PeaceTech Lab | Hate speech*. PeaceTech Lab | Putting the right tools in the right hands to build peace. Retrieved April 9, 2020, from https://www.peacetechlab.org/hate-speech

Portland Communications. (2018). How Africa Tweets 2018. *Portland*. Retrieved from https://portland-communications.com/publications/how-africa-tweets-2018/

reddit.com. (n.d.). *Overview for automoderator*. Retrieved April 9, 2020, from https://www.reddit.com/user/automoderator

Robertson, J. (2016, March 31). How to hack an election. *Bloomberg.com*. Retrieved from https://www.bloomberg.com/features/2016-how-to-hack-an-election/

Sanger, D. E., & Edmondson, C. (2019, July 25). Russia targeted election systems in all 50 states, report finds. *The New York Times*. Retrieved from https://www.nytimes.com/2019/07/25/us/politics/russian-hacking-elections.html

Sang-hun, C. (2013, November 21). Prosecutors detail attempt to sway South Korean election. *The New York Times*. Retrieved from http://www.nytimes.com/2013/11/22/world/asia/prosecutors-detail-bid-to-sway-south-korean-election.html

Savage, S. (2016, February 26). Activist bots: Helpful but missing human love? *Medium*. Retrieved from https://points.datasociety.net/unleashing-the-power-of-activist-bots-to-citizens-1fe888f60207

Savage, S., Monroy-Hernández, A., & Hollerer, T. (2016). *Botivist: Calling volunteers to action using online bots*. Microsoft. Retrieved from https://www.microsoft.com/en-us/research/publication/botivist-calling-volunteers-to-action-using-online-bots/

Scherer, M. (2014, August 6). Meet the bots that knew ebola was coming. *Time*. Retrieved from https://time.com/3086550/ebola-outbreak-africa-world-health-organization/

Sharma, R. (2020, January 6). 25 best Twitter bots you should follow in 2020. *Beebom*. Retrieved from https://beebom.com/best-twitter-bots/

Singer, P. W., & Brooking, E. T. (2018). *LikeWar: The weaponization of social media*. Eamon Dolan/Houghton Mifflin Harcourt.

Tsvetkova, M., García-Gavilanes, R., Floridi, L., & Yasseri, T. (2017). Even good bots fight: The case of Wikipedia. *PLOS ONE*, *12*(2), e0171774. doi:https://doi.org/10.1371/journal.pone.0171774

Vernon, P. (2017, September). Redressing the balance: Why we need more peacebuilding in an increasingly uncertain world. *International Alert*. Retrieved from https://www.international-alert.org/publications/redressing-the-balance

Ward, C., Polglase, K., Shukla, S., Mezzofiore, G., Lister, T. (2020, April 11). Russian election meddling is back – via Ghana and Nigeria – and in your feeds. CNN. Retrieved from https://www.cnn.com/2020/03/12/world/russia-ghana-troll-farms-2020-ward/index.html

Woolley, S. (2016). Automating Power: Social bot interference in global politics. First Monday.

Woolley, S., boyd, danah, Broussard, M., Elish, M., Fader, L., Hwang, T., et al. (2016, February 23). How to think about bots. *Motherboard*. Retrieved from http://motherboard.vice.com/read/how-to-think-about-bots

Woolley, S., & Guilbeault, D. (2017). Computational propaganda in United States: Manufacturing consensus online. In S. Woolley & P. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (pp. 185–211). Oxford Internet Institute, University of Oxford.

Woolley, S., & Howard, P. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press. Retrieved from https://global.oup.com/academic/product/computational-propaganda-9780190931414?cc=us&lang=en&

Woolley, S., Shorey, S., & Howard, P. (2018). The bot proxy: Designing automated self expression. In *A networked self and platforms, stories, connections* (pp. 59–76). Routledge. doi:https://doi.org/10.4324/9781315193434-5

Woolley, S. C., & Howard, P. N. (2016). Automation, algorithms, and politics| Political communication, computational propaganda, and autonomous agents – Introduction. *International Journal of Communication, 10*, 9.

York, J. C. (2013, January 4). No, we should not build pro-democracy Twitter bots. *Slate Magazine*. Retrieved from https://slate.com/technology/2013/01/pro-democracy-twitter-bots-won-t-work-here-s-why.html

Zille, H. (2018). From the inside: Fake accounts and bots and their vicious influence on SA's public discourse. *Daily Maverick*. Retrieved from https://www.dailymaverick.co.za/opinionista/2018-10-29-from-the-inside-fake-accounts-and-bots-and-their-vicious-influence-on-sas-public-discourse/