

**Anonymous Robot Speech**

Madeline Lamo & Ryan Calo

We Robot 2018

**Introduction**

Bots talk to us every day. These often simple programs trace their lineage at least back to Joseph Weizenbaum who, in 1966, published a program known as Eliza. Eliza, named for the character in *My Fair Lady*, interacted credibly with people by posing Rogerian-style questions.<sup>1</sup> Today, virtually any platform capable of supporting communications—from Facebook to Twitter to phone messaging apps—plays host to thousands of bots of varying sophistication. Bots can be entertaining and helpful. They can constitute art. But bots also have the potential to cause harm in a wide variety of contexts by manipulating the people with whom they interact and by spreading misinformation.

Concerns about the role of bots in American life have, in recent months, led to increased calls for regulation. Oren Etzioni, CEO of the Allen Institute for Artificial Intelligence, offered in a September 2017 *New York Times* op-ed the rule that “an A.I. system must clearly disclose that it is not human.”<sup>2</sup> Billionaire businessman and technology investor Mark Cuban tweeted in January 2018 that Twitter and Facebook should “confirm a real name and real person behind every account” and ensure that there is “a single human behind every account.”<sup>3</sup> Senators Klobuchar, Warner, and McCain have drafted legislation known as the “Honest Ads Act” that would modify FEC regulations about use of social media, including bots, in the political context. As drafted, the bill would require all “digital platforms with over 1 million users” to “maintain a public file of all electioneering communications purchased by a person or group who spends more than \$10,000 aggregate dollars for online political advertisements.”<sup>4</sup> Senator Warner stated that he “wants Americans seeing an ad to ‘know whether the source of that ad was generated by foreign entities’”

---

<sup>1</sup> cite

<sup>2</sup> Oren Etzioni, “How to Regulate Artificial Intelligence,” *NEW YORK TIMES*, <https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html? r=0> (Sept. 1, 2017).

<sup>3</sup> Twitter, @mcuban, “It’s time for @twitter to confirm a real name and real person behind every account, and for @facebook to get far more stringent on the same. I don’t care what the user name is. But there needs to be a single human behind every account.” (10:49 a.m., January 28, 2018).

<sup>4</sup> <http://thehill.com/policy/technology/356065-senate-dems-want-curb-election-interference-with-new-bill>

## **EARLY DRAFT**

and that “users should know whether a story is trending because real people shared it or because bots or fake accounts engaged with it.”<sup>5</sup>

This paper considers the role that the First Amendment would play in any such regulations. Scholars who have considered the threshold question of First Amendment coverage of bot speech generally agree that constitutional free speech protections apply in this context. We tend to agree that bots are within the “scope” of First Amendment protection, to borrow Frederick Schauer’s famous terminology.<sup>6</sup> But scope is only a threshold question; coverage does not tell us whether any given government intervention will succeed or fail. Generally speaking, the government may with adequate justification require the disclosure of truthful information—such as the calorie count of food—especially if the speech is commercial in nature. The government can apply reasonable time, manner, or place restrictions to any speech. To require a bot to identify as a bot, rather than as any individual speaker, feel intuitively different from censoring bot speech or unmasking the anonymous proponent of an idea.

Our thesis is that restricting bot speech, including through coerced self-identification, may be trickier than it first appears. As we explore below in greater detail, courts may look with skepticism at a rule requiring all bots to reveal themselves in all circumstances. Does a concern over consumer or political manipulation, for instance, justify a requirement that artists tell us whether a person is behind their latest creation? Moreover, even interventions that appear First Amendment sensitive on their face may wind up impossible to enforce without abridging free speech or otherwise facilitating censorship.

The paper proceeds as follows. Part I gives background on the variety, utility, and danger of bots. Part II summarizes the case for First Amendment coverage of bot or robot speech. Part III develops the case law around anonymity in free speech. These Parts build up to Part IV, which analyzes whether proposals to require bots to identify as non-human are likely to survive constitutional scrutiny as envisioned by their proponents and as applied in actual practice. A final section concludes.

### **I. The utility, fun, and danger of bots**

Most of us interact with bots regularly, albeit to varying degrees. We talk to bots on the phone when we call a customer service line or receive a marketing call. Bots talk to us online through social media, where automated corporate accounts abound. Some social media bots are

---

<sup>5</sup> <https://www.cnn.com/2017/10/18/sens-warner-klobuchar-to-unveil-bill-on-russian-election-meddling.html>

<sup>6</sup> 117 Harv. L. Rev. 1765

## **EARLY DRAFT**

whimsical, fun, or artistic. But others cause trouble, most notably in the political context. As the artificial intelligence (“AI”) that underpins bots continues to improve, it becomes increasingly difficult for us to distinguish human from bot, both on the phone and online. This increases the potential harm a bot can cause. Here we briefly identify several categories of bots, as well as the positive qualities and potential mischief caused by each.

### *a. Commercial bots*

A new group of bots are under development to interact with consumers.<sup>7</sup> While perhaps newly visible, the idea of engaging consumers with bots is not new. In 2004, Ian Kerr wrote about ELLEgirlBuddy, an instant messenger bot designed to chat with teen girls online and encourage them to visit Ellegirl.com.<sup>8</sup> While ELLEgirlBuddy is now long retired,<sup>9</sup> commercial bots have evolved and proliferated. We regularly encounter them in the form of Interactive Voice Response (“IVR”) systems when we call our banks or other customer service lines. We receive sales marketing calls and struggle to discern whether we are speaking with a real human or a robot.<sup>10</sup> Many corporate entities use automated social media accounts, from Puma to Coca-Cola to the New England Patriots.<sup>11</sup> While the use of automated bot accounts may create an occasional PR nightmare for these companies (like when the official Coca-Cola Twitter account was tricked into tweeting out text from *Mein Kampf*),<sup>12</sup> the use of bots largely allows corporations to promote their brands online peacefully.

In some cases, being able to communicate with an automated agent by phone or a customer service chatbot online allows consumers to solve simple problems and answer questions quickly and easily. Bots enable people to communicate through text, which many consumers prefer over other means of communication.<sup>13</sup> Commercial chatbots are available at all hours of the day.<sup>14</sup> They never get cranky or frustrated, even when dealing with the most difficult customers.<sup>15</sup> Perhaps most importantly, they can significantly decrease the wait for customer service assistance.<sup>16</sup>

---

<sup>7</sup> [citation to discussion of contemporary chat bots]

<sup>8</sup> Ian Kerr, *The Californiacation of Commerce*, <http://www.uoltj.ca/articles/vol1.1-2/2003-2004.1.1-2.uoltj.Kerr.285-324.pdf> p. 313 (2004).

<sup>9</sup> *Id.*

<sup>10</sup> <https://io9.gizmodo.com/freakishly-realistic-telemarketing-robots-are-denying-t-1481050295>

<sup>11</sup> <https://digiday.com/marketing/5-biggest-bot-fails-brands-twitter/>

<sup>12</sup> *Id.*

<sup>13</sup> [https://www.aspect.com/globalassets/2016-aspect-consumer-experience-index-survey\\_index-results-final.pdf](https://www.aspect.com/globalassets/2016-aspect-consumer-experience-index-survey_index-results-final.pdf) (p. 6)

<sup>14</sup> Shep Hyken, “AI and Chatbots are Transforming the Customer Experience,” FORBES, <https://www.forbes.com/sites/shephyken/2017/07/15/ai-and-chatbots-are-transforming-the-customer-experience/#657ee5a841f7> (July 15, 2017).

<sup>15</sup> *Id.*

<sup>16</sup> *Id.* Of course, the quality may suffer.

## **EARLY DRAFT**

Yet commercial bots can also cause harm, primarily by tricking and confusing consumers.<sup>17</sup> Robocallers may deny their non-humanness,<sup>18</sup> call targeted individuals repeatedly, and even claim to be a representative of the IRS or another powerful entity that even a tech-savvy individual might feel too anxious to hang up on.<sup>19</sup> Vulnerable populations such as the elderly are particularly susceptible to scamming by robocallers.<sup>20</sup> The Federal Trade Commission recognizes the threat that robocalls pose to consumers and has passed regulations against such practices. It has recently won several lawsuits against companies using robocalls.<sup>21</sup> FTC representatives recently testified before the Senate Special Committee on Aging to discuss the specific threat robocalls pose to the elderly.<sup>22</sup>

Bots can also skew the marketplace, for instance, by creating confusion in product reviews. Online retailers commonly allow purchasers to leave reviews of products, where they can provide helpful information about quality, fit, and other details of use to potential buyers. These reviews are often accompanied by a rating of the product, often out of five stars. Fake reviews can be used to drive up a product's rating or drive down a competitor product's rating, and bots are an effective way to create large numbers of fake reviews in a short amount of time. While major online retailers such as Amazon try to fight bot reviews, they struggle to do so.<sup>23</sup> This can mislead consumers and encourage them to purchase terrible products with fraudulent positive rankings and reviews.

### *b. Artistic or entertainment bots*

In a uniquely creative and enjoyable corner of the internet, one finds bots as an art form, such as the creations of programmer-artist Darius Kazemi.<sup>24</sup> These can vary dramatically in their format. They can be funny, such as Kazemi's @TwoHeadlines account, which combines two current headlines from Google News to create combinations such as "The nuclear agreement is 'the

---

<sup>17</sup> Cf. cite to Hartzog, *Unfair and Deceptive Robots*.

<sup>18</sup> <https://io9.gizmodo.com/freakishly-realistic-telemarketing-robots-are-denying-t-1481050295>

<sup>19</sup> Hanna Landman, "Phone Scams: Preying on the Elderly," Avacare Medical Blog, <https://avacaremedical.com/blog/phone-scams-preying-elderly.html> (June 12, 2017); "IRS Urges Public to Stay Alert for Scam Phone Calls" <https://www.irs.gov/newsroom/irs-urges-public-to-stay-alert-for-scam-phone-calls> (Oct. 21, 2015).

<sup>20</sup> Landman, *supra* note X; "Prepared Statement of the Federal Trade Commission Before the United States Senate Special Committee on Aging" [https://www.ftc.gov/system/files/documents/public\\_statements/1256863/p034412\\_commission\\_testimony\\_re\\_robot\\_calls\\_senate\\_10-4-17.pdf](https://www.ftc.gov/system/files/documents/public_statements/1256863/p034412_commission_testimony_re_robot_calls_senate_10-4-17.pdf) (Oct. 4, 2017).

<sup>21</sup> See, e.g., <https://www.ftc.gov/news-events/press-releases/2016/07/ftc-providing-4-million-full-refunds-people-tricked-buying-bogus>.

<sup>22</sup> FTC Special Committee statement, *supra* note X.

<sup>23</sup> Shareen Pathak, "Amazon Reviews Have a Bot Problem," DIGIDAY, <https://digiday.com/marketing/amazon-reviews-bot-problem/> (Sept. 18, 2017).

<sup>24</sup> <https://www.bostonglobe.com/ideas/2014/01/24/the-botmaker-who-sees-through-internet/V7Qn7HU8TPPI7MSM2Tvbsj/story.html>

## **EARLY DRAFT**

worst deal ever’ — for Dale Earnhardt Jr.”<sup>25</sup> They can also be informative (@netflix\_bot tweets out updates about what is currently streaming on Netflix), create art (@greatartbot tweets out a new piece of computer-generated artwork every four hours; @pixelsorter resorts the pixels in images users send it to create beautiful, soothing images), and even identify the poetry that humans unintentionally tweet (@accidental575: “I am a robot / that finds haikus on Twitter / made by accident”; @pentameton: “With algorithms subtle and discreet / I seek iambic writings to retweet”). Some of these bots achieve their programmers’ artistic aims best when users cannot tell whether an account is automated or human-run. While artistic bots may create genuine confusion,<sup>26</sup> and while bot admittedly threatened a fashion show in Amsterdam with violence,<sup>27</sup> they typically do not bring about actual harm. They are largely a fun, imaginative format that combines art, humor, and technology.

### *c. Political bots*

Arguably the most troubling use of bots on social media arises in the political context. The use of bots in the political arena is a more recently recognized phenomenon than in the commercial context, so the attendant risks are less well-understood. Though the full scope of their influence is still unknown, recent investigations indicate that social media bots were used extensively by a Russian government-linked organization to influence the 2016 American presidential election.<sup>28</sup> Research by the Oxford Internet Institute shows that pro-Trump Twitter bots were four times as active as pro-Clinton bots around the time of the first presidential debate.<sup>29</sup> This margin increased to a five-to-one pro-Trump to pro-Clinton bot activity ratio by election day.<sup>30</sup> Some even speculate that the long-term goal of this interference was to undermine democracy more broadly.<sup>31</sup> While some of these bots share seemingly original content, others primarily magnified existing content by “retweeting” posts, following prominent accounts, and hijacking hashtags to make them trend.<sup>32</sup>

---

<sup>25</sup> Twitter, @TwoHeadlines, “The nuclear agreement is 'the worst deal ever' — for Dale Earnhardt Jr.” (9:08 a.m., February 19, 2018).

<sup>26</sup> See, e.g., Adrian Chen, “How I Found the Human Being Behind @Horse\_ebooks, the Internet’s Favorite Spambot,” GAWKER, <http://gawker.com/5887697/how-i-found-the-human-being-behind-horsebooks-the-internets-favorite-spambot> (Feb. 23, 2012).

<sup>27</sup> [add cite]

<sup>28</sup> [https://blog.twitter.com/official/en\\_us/topics/company/2018/2016-election-update.html](https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html)

<sup>29</sup> <http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2016/11/Data-Memo-US-Election.pdf>

<sup>30</sup> *Id.*

<sup>31</sup> Seema Metha, “Rep. Adam Schiff Says Alleged Russian Meddling in Election Was an Effort to Destroy American Democracy,” LOS ANGELES TIMES, <http://www.latimes.com/politics/essential/la-pol-ca-essential-politics-updates-schiff-argues-russian-intervention-in-1496173190-htmlstory.html> (May 30, 2017).

<sup>32</sup> <https://www.bloomberg.com/news/articles/2018-01-26/twitter-says-russian-linked-bots-retweeted-trump-470-000-times>; Lutz Finger, “Do Evil – The Business of Social Media Bots,” FORBES,

## **EARLY DRAFT**

While the full effect of this type of bot use has not yet been quantified, it seems clear that political bots may be used to skew trends, to make certain ideas and individuals appear more popular than they would be otherwise, and to stir up dissent and discord.

Political bots have the potential to create greater harm than commercial bots, and without the silver lining of increasing efficiency or otherwise improving society. But it would be unfair and incorrect to paint them as a uniquely evil force. Bots are in many ways an extension of other forms of media. Technology and the media in their many forms have long played a critical role in the political context. The Federalist papers, published anonymously in New York newspapers in the 1780s, helped sway popular opinion in favor of ratifying the Constitution. Franklin D. Roosevelt used the radio to speak directly into the homes of Americans with his fireside chats, later described as “a revolutionary experiment with a nascent media platform.”<sup>33</sup> John F. Kennedy won the support of the American electorate by appearing “robust and confident” in televised debates.<sup>34</sup> In 2011, social media entered the political fray by way of the Arab Spring. In Tunisia, Egypt, and elsewhere in the Middle East and North Africa, social media enabled activists to share their messages and organize demonstrations against powerful authoritarian governments. Bots may be a natural result of ever-improving technology. Like other media, they possess distinct qualities that make them a uniquely powerful communication tool in the political context. These unique qualities also allow them to cause harm in a variety of dramatic and often invisible ways.

Nevertheless, the unique affordances of bots present novel social media concerns. By automating “trolling,” i.e., the practice of criticizing or threatening certain speakers such as women and people of color in response to their views, bots can exacerbate highly problematic trends of online hate speech and abuse.<sup>35</sup> Bots can do this individually, by attacking one or more people, or at scale: bots are capable of sinking a useful hashtag but overusing it and flooding the hashtag with useless or countermanding information. For example, after the February 2018 school shooting in Parkland, Florida, Russian-controlled bots joined many social media users in tweeting

---

<https://www.forbes.com/sites/lutzfinger/2015/02/17/do-evil-the-business-of-social-media-bots/#149dabcefb58> (Feb. 17, 2015). See also Samuel C. Wooley & Douglas R. Guilbeault, *Computational Propaganda in the United States of America: Manufacturing Consensus Online*, Working Paper No. 2017.5, <http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf> (2017), at p. 10–11).

<sup>33</sup> Adrienne LaFrance, “Donald Trump is Testing Twitter’s Harassment Policy,” ATLANTIC, <https://www.theatlantic.com/politics/archive/2017/07/the-president-of-the-united-states-is-testing-twitters-harassment-policy/532497/> (July 2, 2017).

<sup>34</sup> Kenneth T. Walsh, “JFK: First TV President,” US NEWS, <https://www.usnews.com/news/blogs/ken-walshs-washington/2013/11/20/jfk-first-tv-president> (Nov. 20, 2013).

<sup>35</sup> Cite Daniel Citron’s books on online hate speech.

## **EARLY DRAFT**

#guncontrolnow—but accompanied the hashtag with messages, links, and images suggesting that stricter gun control laws would not have prevented the tragedy.<sup>36</sup>

Conversely, bots can engage in false amplifications. When bots coalesce around a certain hashtag, account, or news story, they can help that topic “trend” on social media.<sup>37</sup> For example, Russian-linked bots retweeted Donald Trump approximately ten times more than they retweeted Hillary Clinton in the months preceding the 2016 election,<sup>38</sup> thereby dramatically increasing the overall amount of attention given to those tweets. Or they can flood an administrative agency with duplicative comments.<sup>39</sup> Some describe this as “manufacturing consensus,”<sup>40</sup> or making a certain fringe viewpoint appear so popular that it seems legitimate and newsworthy.<sup>41</sup> Oxford Internet Institute director Philip Howard argues that “[i]f you use enough . . . bots and people, and cleverly link them together, you are what’s legitimate. You are creating truth.”<sup>42</sup>

Relatedly, bots can increase the number of followers someone has on social media. While this in itself does not seem problematic, it can deceive other social media users into thinking that someone is more powerful, important, or influential than they really are.<sup>43</sup> In the political context, this is particularly problematic, as a high follower count may suggest that a particular individual is particularly important or that her views are popular and widely-accepted. Twitter allows automated accounts generally, but has a policy against “aggressive following.”<sup>44</sup> Occasional bot “purges” may cause popular social media accounts to lose thousands or even millions of followers.<sup>45</sup>

Finally, and most visibly, bots can support coordinated campaigns of disinformation. Although seldom the only driver, hosts of bots can help spread false or misleading news or else stoke national strife during a crisis or other salient news event. It is this potential that has led federal

---

<sup>36</sup> <https://www.wired.com/story/pro-gun-russian-bots-flood-twitter-after-parkland-shooting/>

<sup>37</sup> Finger, *supra* note X.

<sup>38</sup> <https://www.bloomberg.com/news/articles/2018-01-26/twitter-says-russian-linked-bots-retweeted-trump-470-000-times>

<sup>39</sup> [cite]

<sup>40</sup> Farhad Manjoo, “How Twitter is Being Gamed to Feed Misinformation,” NEW YORK TIMES, <https://www.nytimes.com/2017/05/31/technology/how-twitter-is-being-gamed-to-feed-misinformation.html?mtrref=www.google.com&gwh=A3596219022CDBD0424BFC1D013501E9&gwt=pay> (May 31, 2017).

<sup>41</sup> See also Alice Marwick and Rebecca Lewis, *Media Manipulation and Disinformation Online* at 38, [https://datasociety.net/pubs/oh/DataAndSociety\\_MediaManipulationAndDisinformationOnline.pdf](https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf).

<sup>42</sup> Samuel Earle, “Trolls, Bots, and Fake News: The Mysterious World of Social Media Manipulation,” NEWSWEEK, <http://www.newsweek.com/trolls-bots-and-fake-news-dark-and-mysterious-world-social-media-manipulation-682155> (Oct. 14, 2017).

<sup>43</sup> Finger, *supra* note X.

<sup>44</sup> The Twitter Rules, <https://help.twitter.com/en/rules-and-policies/twitter-rules>.

<sup>45</sup> Kerry Flynn, “Twitter Influencers Suspect a ‘Bot Purge,’” MASHABLE, <https://mashable.com/2018/01/29/twitter-bots-purge-influencers-accounts/#p8stq6jiPiqF> (Jan. 29, 2018).

## **EARLY DRAFT**

lawmakers to grill social media executives in hearings in recent months and ultimately to propose the requirements around disclosure we allude to in the Introduction.

### **II. First Amendment Coverage of Bot Speech**

A threshold question to whether regulations of bots would survive a free speech challenge is whether automated or “robot” speech is protected under the First Amendment at all. What is and is not protected as “speech” under the First Amendment is a complex question, and one that requires us to set aside the lay understanding of what constitutes “speech.” As Frederick Schauer famously observes, “[t]he speech with which the First Amendment is even slightly concerned is but a small subset of the speech that pervades every part of our lives.”<sup>46</sup> While the justifications vary, there is a rough consensus among experts that automated speech such as online bots or robo-callers are among the subset that falls within the Constitution’s protection.

The Supreme Court has enumerated certain categories of communicative acts that receive varying degrees of protection under the First Amendment. At its heart, this is largely a normative determination by the Court about what the First Amendment *ought* to protect. For example, pornographic writing would not be protected under the First Amendment, despite the fact that it takes the form of written words on a page.<sup>47</sup> Expressive conduct such as burning a flag, however, would qualify for First Amendment protection, despite the fact that it lacks a verbal or written component.<sup>48</sup> The categories of protected and unprotected speech are complex and often difficult to define. While some have argued that First Amendment protection should only be extended to speech that is “explicitly political,”<sup>49</sup> the Supreme Court has declined to draw such a bright line.<sup>50</sup>

Evolution of technology does not automatically change the scope of the First Amendment protection. The Supreme Court clearly stated that First Amendment protection should not vary by speech medium, including new media that grow out of developing technology: “Whatever the challenges of applying the Constitution to ever-advancing technology, ‘the basic principles of freedom of speech and the press, like the First Amendment’s command, do not vary’ when a new

---

<sup>46</sup> Frederick Schauer, *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 Harv. L. Rev. 1765, 1777 (2004).

<sup>47</sup> *Miller*

<sup>48</sup> *Texas v. Johnson*

<sup>49</sup> Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 Ind. L.J. 1, 20 (1971).

<sup>50</sup> Tim Wu, *Debate: Machine Speech*, 161 U. Pa. L. Rev. 1495, 1507 (2013).

## **EARLY DRAFT**

and different medium for communication appears.”<sup>51</sup> The Supreme Court extended First Amendment protection to video games in *Brown v. Entertainment Merchants Association*, and a lower court recognized First Amendment protection for search engine results in *Zhang v. Baidu.com Inc.*<sup>52</sup> This indicates that the Supreme Court might be willing to treat robot speech comparably to human speech, so long as other constitutional and statutory requirements are satisfied.

Furthermore, greater attenuation between human and speech does not change the scope of First Amendment protection. Just because a statement is ultimately “made” by a robot does not mean that it is not the product of human creation. Tim Wu notes that “[l]ike a book, canvas, or pamphlet, the program is the medium the author uses to communicate his ideas to the world”<sup>53</sup> in the context of algorithm-generated communicative outputs. The degree of attenuation between a human creator and her final speech output can vary widely, and a greater degree of attenuation does not decrease the scope of First Amendment protection.<sup>54</sup> Thus, the fact that a Twitter bot creator may not know what her creation will tweet next does not place the bot outside the protection of the First Amendment.

Additionally, the fact that robot speech is generally unoriginal and just some kind of repetition or re-splicing of old communications does not place it outside the scope of First Amendment protection. Ultimately, “the fact that the person or entity claiming to be engaged in speech does not create the underlying content is irrelevant for purposes of First Amendment coverage.”<sup>55</sup> Therefore, even bots that do not generate any kind of original content could receive protection under the First Amendment.

Finally, the First Amendment protects the right of listeners and readers to receive information. Even when a great degree of attenuation exists between human creator and the final speech product, the First Amendment may still protect the communication, because it protects not only the right to speak but also the right to receive information.<sup>56</sup> Would-be listeners or readers can assert their own First Amendment rights, even when the censored speaker lacks First Amendment rights of her own.<sup>57</sup> Despite the current critiques of social media bots for their role in skewing the

---

<sup>51</sup> *Brown v. Entertainment Merchants Ass’n*, 564 U.S. 786, 790 (2011) (quoting *Joseph Burstyn, Inc. v. Wilson*, 343 U.S. 495, 503 (1952)).

<sup>52</sup> 10 F. Supp. 3d 433 (S.D.N.Y. 2014).

<sup>53</sup> *Id.* at 1504.

<sup>54</sup> Stuart Benjamin, *Debate: Algorithms and Speech*, 161 U. Pa. L. Rev. 1445, 1464–65 (2013).

<sup>55</sup> *Id.* at 1463 n.64.

<sup>56</sup> *Pico*

<sup>57</sup> *Kleindeinst v. Mandel* (academics asserted their own First Amendment rights wishing to hear lectures from foreign communist professor whose visa application was denied).

## **EARLY DRAFT**

American political dialogue online, there are in fact many enjoyable bots that make the internet a brighter, funnier, and more interesting place.<sup>58</sup> Thus, the First Amendment could, for example, protect the rights of internet users who wish to read tweets from their favorite bots.

More detailed discussions of the potential scope of First Amendment coverage of bot speech can be found elsewhere.<sup>59</sup> The general consensus in burgeoning literature seems to be that the First Amendment could very well apply. Some, such as Tim Wu, take a narrow, functionalist view.<sup>60</sup> Others such as Helen Norton, Toni Massaro, and Margot Kaminski assert that it will apply to bot speech and algorithmically-generated speech more broadly.<sup>61</sup> A forthcoming book by David Skover and Ronald Collins argues extensively and persuasively for speech protection for robots on the theory that the First Amendment is and has always been largely predicted on audience interests, which benefit from speech irrespective of the speaker.<sup>62</sup> We believe, and will assume here, that the First Amendment would be generally applicable to bot speech. The thrust of this paper is not the applicability of the First Amendment in the context of bot speech, but rather an examination of ways in which the constitutionality of bot speech regulation might differ from traditional speech regulation jurisprudence under the First Amendment.

### **III. Right to Anonymity**

We assume speech by bots is covered by the First Amendment. The proposals to regulate bot speech that motivate this paper are not, however, aimed at censorship per se. If they “abridge” speech, they do so by requiring a *category* of speaker to identify itself as such. The proposals do not even require the bot to identify precisely who is speaking, only that a person is not. It may seem tenuous, therefore, to argue that a rule aimed only at requiring calls or social media accounts by bots to acknowledge no human is behind them even rises to the level of a restriction. As we explain in the next Part, this is a closer question than it first appears. The very ambiguity around who is speaking may form an integral part of the message. Moreover, it may prove impossible to enforce a bot-disclaimer without identifying an otherwise anonymous speaker or providing the scaffolding for censorship. Thus, by way of additional background, we discuss here the origin and contours of the right to speak anonymously.

---

<sup>58</sup> See, e.g., <http://nymag.com/selectall/2015/11/12-weirdest-funniest-smartest-twitter-bots.html>; <https://qz.com/279139/the-17-best-bots-on-twitter/>.

<sup>59</sup> String cite.

<sup>60</sup> Tim Wu article

<sup>61</sup> Siri-ously and 2.0

<sup>62</sup> Robotica.

## **EARLY DRAFT**

### *a. The Supreme Court on the right to anonymity*

Discussion of the right to anonymity in American jurisprudence inevitably begins with a reference to the famed Federalist papers.<sup>63</sup> But the Supreme Court did not directly speak on the issue of anonymity protection until the mid-1900s. In *NAACP v. Alabama ex rel. Patterson*,<sup>64</sup> the Court held that Alabama could not require the local NAACP chapter to provide it with a list of names of its members. This decision was derived from First Amendment freedom of association protection rather than freedom of speech protection.<sup>65</sup> The right to anonymity was first formally linked to free speech protection two years later in *Talley v. California*,<sup>66</sup> in which the Supreme Court struck down a municipal ordinance that prohibited the distribution of handbills that did not include the name and address of person issuing them. The Court reasoned that an identification requirement would “tend to restrict freedom to distribute information” and thus, by extension, would inhibit freedom of expression.<sup>67</sup> Accordingly, the Court found that protection of the right to speak anonymously constituted an integral component of the right to express minority political views: “Anonymous pamphlets, leaflets, brochures and even books have played an important role in the progress of mankind. Persecuted groups and sects from time to time throughout history have been able to criticize oppressive practices and laws either anonymously or not at all.”<sup>68</sup>

The Court reaffirmed this staunch protection of the right to speak anonymously thirty-five years later in *McIntyre v. Ohio Elections Commission*<sup>69</sup> in which the Court struck down an Ohio law prohibiting the distribution of campaign literature that did not contain the name and address of the person issuing it. Drawing a parallel to the well-founded right to vote anonymously, the Court articulated that *Talley* had established “a respected tradition of anonymity in the advocacy of political causes”<sup>70</sup> that serves as “a shield from the tyranny of the majority.”<sup>71</sup> As Margot Kaminski notes, the *McIntyre* decision departed from the emphasis on minority political dissent found in *Talley* and *NAACP* and shifted towards a broader protection of anonymity as an element of expression

---

<sup>63</sup> *McIntyre v. Ohio Elections Comm’n*, 514 U.S. 334, 343 n.6 (1995).

<sup>64</sup> 357 U.S. 449 (1958).

<sup>65</sup> Need pincite.

<sup>66</sup> 362 U.S. 60 (1960).

<sup>67</sup> *Id.* at 64.

<sup>68</sup> *Id.*

<sup>69</sup> 514 U.S. 334 (1995).

<sup>70</sup> *Id.* at 343.

<sup>71</sup> *Id.* at 357.

## **EARLY DRAFT**

generally:<sup>72</sup> “Anonymity is a means of expressing oneself, and an author has the freedom to decide whether or not to disclose his or her true identity. An author may choose to be anonymous because of fear of retaliation, concern about social ostracism, or a desire to protect his or her privacy; the Court implied that the precise reason does not in fact matter.”<sup>73</sup>

### *b. Anti-mask laws and anonymity*

Kaminski has written extensively on the relationship between anti-mask statutes and the right to anonymity and we rely on that analysis here.<sup>74</sup> She identifies several different categories of anti-mask statutes across many jurisdictions, ranging from strict liability statutes that criminalize all mask wearing in public (with limited exceptions for Halloween and other innocuous activities) to statutes that penalize the wearing of a mask during the commission of a crime.<sup>75</sup> Kaminski argues that “[t]he variation in anti-mask statutes suggests that legislatures, like courts, struggle with determining when anonymity is functional and when it is expressive.”<sup>76</sup> The Supreme Court has not spoken specifically on the matter, and lower courts have interpreted the anonymity protections enumerated in *Talley* and *McIntyre* varying in the anti-mask context. Some decisions find that they established an independent right to anonymity under the First Amendment,<sup>77</sup> while others characterize the right as a component of the right to free expression.<sup>78</sup> Other courts declined to apply the First Amendment altogether in the context of anti-mask laws, most notably the Second Circuit in its *Kirek* decision.<sup>79</sup>

From her analysis of these cases, Kaminski concludes that courts differ sharply regarding the requirements of functionality and expressiveness in a First Amendment challenge to an anti-mask law. In some courts, “[b]ecause mask-wearing permits free speech and association, it is protected under *NAACP*.”<sup>80</sup> Elsewhere, however, “the functional element of anonymity is used . . . to justify not applying the First Amendment at all.”<sup>81</sup>

---

<sup>72</sup> Margot Kaminski, *Real Masks and Real Name Policies*, 23 Fordham Intell. Prop. Media & Ent. L.J. 815, 834 (2013).

<sup>73</sup> *Id.* at 834–35.

<sup>74</sup> See generally Kaminski, *supra* note X.

<sup>75</sup> *Id.* at 848–49.

<sup>76</sup> *Id.* at 850.

<sup>77</sup> See, e.g., *Am. Knights of the Ku Klux Klan v. City of Goshen*, 50 F. Supp. 2d 835 (N.D. Ind. 1999).

<sup>78</sup> See, e.g., *Gbafari v. Mun. Ct. for S.F. Jud. Dist. of S.F.*, 87 Cal. App. 3d 255, 265 (Cal. Ct. App. 1978).

<sup>79</sup> *Church of the Am. Knights of the Ku Klux Klan v. Kerek*, 356 F.3d 197, 211 (2d Cir. 2004).

<sup>80</sup> Kaminski, *supra* note X, at 875.

<sup>81</sup> *Id.* at 875–76.

## **EARLY DRAFT**

### *c. When unmasking laws survive First Amendment review*

The anonymity protected established in these cases is powerful but not absolute. Under the broad First Amendment protection of the right to anonymity in “core political speech” (as in *Talley* and *McIntyre*), a law must survive strict (or “exacting”) scrutiny.<sup>82</sup> The government bears the burden of proving that a challenged law is narrowly tailored in furtherance of a compelling state interest.<sup>83</sup> Such review is often thought to be “strict in theory, but fatal in fact,”<sup>84</sup> sounding a death knell for any law subject to it. However, lower scrutiny is applied in the context of commercial speech. Accordingly, anonymous speakers may be validly unmasked in several important contexts: first, in the context of commercial speech, which receives less protection; and second, in two particular areas in which courts have recognized government interests that justify the unmasking of anonymous speakers: electoral speech and at certain points in litigation.

#### i. Anonymity in commercial speech

Anonymity is seldom at issue in commercial speech, as companies naturally want consumers to recognize their brand and product names. However, agency regulations may require commercial products to conspicuously bear “the name and place of business of the manufacturer, packer, or distributor”<sup>85</sup> The justification behind permitting such disclosure requirements is the idea that more information is generally good for consumers:

Mandated disclosure of accurate, factual, commercial information does not offend the core First Amendment values of promoting efficient exchange of information or protecting individual liberty interests. Such disclosure furthers, rather than hinders, the First Amendment goal of the discover of truth and contributes to the efficiency of the ‘marketplace of ideas.’ Protection of the robust and free flow of accurate information is the principal First Amendment justification for protection commercial speech, and requiring disclosure of truthful information promotes that goal.<sup>86</sup>

In the interest of increasing the flow of information in the commercial marketplace, the First Amendment permits greater disclosure requirements. This includes requirements that commercial products bear the name of their manufacturers, packers, and distributors. It does not permit unlimited government regulation, however. The Sixth Circuit and the D.C.

---

<sup>82</sup> *McIntyre*, 514 U.S. at 334–35.

<sup>83</sup> *Id.*; *First Nat’l Bank of Boston v. Bellotti*, 435 U.S. 765, 786 (1978).

<sup>84</sup> *Fullilove v. Klutznick*, 448 U.S. 448, 518 (1980) (Marshall, J., concurring).

<sup>85</sup> 21 C.F.R. 101.5 (food labeling); 21 C.F.R. 201.1 (drug labeling).

<sup>86</sup> *New York State Restaurant Ass’n v. New York City Bd. of Health*, 556 F.3d 114, 132 (2d Cir. 2009) (citing *Nat’l Elec. Mfrs. Ass’n v. Sorrell*, 272 F.3d 104, 113–14 (2d Cir. 2001)) (upholding city law requiring restaurants to post nutrition information on menus).

## **EARLY DRAFT**

Circuit disagreed on whether requiring cigarette companies to include graphic visual warnings on cigarette packaging constituted a permissible “mere information” disclosure<sup>87</sup> or impermissibly “were aimed at changing behavior and hence pressed the cigarette industry into regulatory service.”<sup>88</sup>

### ii. Anonymity around elections

In *McIntyre*, the Court noted that a narrow identification requirement might be justified on the basis of certain government interests.<sup>89</sup> Although anonymity protection is at its highest in the context of political speech, the Supreme Court has recognized that the government interest in preserving the integrity of the electoral process is so compelling that it occasionally satisfies unmasking requirements specifically in the political speech context. For example, the Court upheld a Washington law requiring the state to release the names of signatories to ballot referendum petitions upon request under the Public Records Act.<sup>90</sup> The Court recognized that signing a referendum petition is expressive, as it communicates the message that the signatory supports the referendum, or at the very least thinks it should be put to a vote.<sup>91</sup> However, the Court upheld the law on the basis of the fact that it did not suppress speech but rather required disclosure of more information coupled with the fact that the disclosure was intended to strengthen the integrity of the electoral process.<sup>92</sup> That same year, the Court upheld the Bipartisan Campaign Reform Act’s disclaimer and disclosure requirements for campaign advertisements in *Citizens United v. FEC*.<sup>93</sup> In doing so, the Court emphasized the public’s “informational interest” and the importance of “making informed choices in the political marketplace.”<sup>94</sup>

The limits of what must be disclosed in order to preserve the integrity of the electoral process are unclear. The line between signatures on a pamphlet and signatures on a referendum ballot initiative is quite thin, and it is difficult to discern how the Court would rule on other election-related disclosure requirements. Several justices on the *Reed* Court argued that *Reed* and *McIntyre* should have reached the same result, including Justice Scalia, who would have reached the opposite

---

<sup>87</sup> *Disc. Tobacco City & Lottery, Inc. v. United States*, 674 F.3d 509, 556–61 (6th Cir. 2012).

<sup>88</sup> Ryan Calo, *Code, Nudge, or Notice?*, 99 Iowa L. Rev. 773, 777 (2013). *R.J. Reynolds Tobacco Co. v. FDA*, 696 F.3d 1205, 1221–22 (D.C. Cir. 2012).

<sup>89</sup> *McIntyre*, 514 U.S. at 353 (1995).

<sup>90</sup> *John Doe No. 1 v. Reed*, 561 U.S. 186 (2010).

<sup>91</sup> *Id.* at 195.

<sup>92</sup> *Id.* at 197. Note that this was a facial challenge; the Court did not rule on the validity of the Public Records Act disclosure requirement as applied in the context of a marriage amendment referendum. Kaminski, *supra* note X, at 839.

<sup>93</sup> 558 U.S. 310 (2010).

<sup>94</sup> *Id.* at 367.

## **EARLY DRAFT**

result in *McIntyre*,<sup>95</sup> and Justice Thomas, who believed that neither law was justified under strict scrutiny.<sup>96</sup>

### iii. Anonymity in litigation [under construction]

Questions of anonymity come up in the context of litigation, where a plaintiff alleges that an unknown defendant perpetrated some harm or where a litigant wishes to proceed anonymous due to the sensitive nature of the litigation. Consequently, there is a relatively robust body of case law around the question of when a plaintiff can unmask an otherwise anonymous agent or require a known defendant to identify a third party. [lays out the standard by which courts will reveal defendants and plaintiffs]

In *Branzburg v. Hayes*,<sup>97</sup> the Supreme Court held that requiring journalists to reveal their sources when subpoenaed by a Grand Jury did not violate the First Amendment. The Ninth Circuit's recent ruling in *United States v. Glassdoor, Inc.*<sup>98</sup> extended this reasoning to find that anonymous employee reviewers on Glassdoor.com could be subject to court-ordered unmasking in the context of an ongoing government investigation into workplace fraud. While the Ninth Circuit's *Glassdoor* ruling has already been the subject of extensive criticism by First Amendment advocates for its failure to take the unique qualities of online speech into account,<sup>99</sup> it remains to be seen whether the Supreme Court will intervene. Thus, for the time being, it seems that anonymous speakers may be unmasked in the context of Grand Jury subpoenas.

## **IV. Constitutionality of Proposed Regulations**

We now turn to an assessment of whether existing calls to identify bots as non-human would be blessed by American courts. While the general idea of requiring bots to disclose the fact that they are automated may seem like a promising content-neutral speech regulation, there are several ways in which such a regulation could fail under existing law. First, the requirement might be

---

<sup>95</sup> *Id.* at 219–20 (Scalia, J., concurring).

<sup>96</sup> *Id.* at 239 (Thomas, J., dissenting).

<sup>97</sup> 408 U.S. 665 (1972).

<sup>98</sup> <https://assets.documentcloud.org/documents/4222206/17-16221.pdf>

<sup>99</sup> *See, e.g.*, Lisa L. Hayes, “Anonymous Speech Online Dealt a Blow in U.S. v. Glassdoor Opinion,” Center for Democracy & Technology, <https://cdt.org/blog/anonymous-speech-online-dealt-a-blow-in-us-v-glassdoor-opinion/> (Nov. 8, 2017) (“the ability to speak anonymously can encourage speakers to engage more openly with one another. This is especially true with respect to online speech, given the internet’s unique technical characteristics. Online communications necessarily depend on intermediaries, such as internet service providers and messaging platforms, to facilitate their carriage, accessibility, and storage. As a result, internet users are particularly vulnerable to having their speech published, decontextualized, and examined in ways that they do not anticipate when initially posting a comment. If a person fears that statements she makes online will be forever linked to her professional or legal identity, she is likely to refrain from voicing at least some thoughts due to concerns about potential repercussions and reprisal”).

## **EARLY DRAFT**

difficult to justify in all the contexts in which it would apply. And second, even if across-the-board justification were possible, the regulation might be hard to carry out in practice without effectively unmasking protected human speakers. The likelihood of a requirement that would unmask the people behind bots directly could survive constitutional review is even slimmer due to the serious potential infringement on the right to anonymity.

### *a. Problems of justification*

Say Congress adopted, for example, Etzioni’s proposal and passed a law requiring an automated speaker to disclose the fact that it is not human. Such a law appears valid on its face as a content- and viewpoint-neutral restriction of time, place, and manner regulation. However, such a regulation might raise constitutional problems nonetheless. First, any justification proffered for a disclosure requirement would not likely meet constitutional muster in all contexts, thus rendering a universal disclosure requirement problematic. Concern over election interference can hardly justify such disclosure by a poetry bot, for example, and regulation in the interest of consumer welfare has no bearing on non-commercial speech. Second, a core principle of speech regulation generally is that there must be ample alternative means for the speaker to convey her message, and it seems that in the context of artistic bots, alternative channels may not be available. Finally, execution of a requirement that a bot disclose its bot-ness may interfere with the right to anonymity by requiring human speakers to unmask themselves.

Time, place, and manner regulations are almost always grouped together in First Amendment decisions, but they are three distinct ideas. Regulating time and place is easily comprehensible, and—if applied fairly—feels intuitively like reasonable government action. Given the choice between a noisy parade on our street at 3 a.m. or at 3 p.m., most of us would prefer the afternoon parade. And given the choice between a march down a public highway that blocks rush hour traffic or a march through a plaza that only reroutes foot traffic, most commuters would prefer the latter. It is less clear, however, what exactly a “manner” regulation entails. The Court ruled that a New York City ordinance requiring that concerts in Central Park use lower-volume sound systems provided by the city was a constitutional “place and manner” regulation in *Ward v. Rock Against Racism*.<sup>100</sup> In doing so, the Court did not define “manner,” but it did note that the ordinance “[did] not attempt to ban any particular manner or type of expression.”<sup>101</sup> In *Clark v. Community for Creative Non-Violence*, the Court held that barring protestors from sleeping on federal land in Washington,

---

<sup>100</sup> 491 US 781, 803 (1989).

<sup>101</sup> *Id.* at 802.

## **EARLY DRAFT**

D.C. to raise awareness of the problem of homelessness was constitutional as “a reasonable regulation of the manner in which a demonstration may be carried out.”<sup>102</sup>

Together, these cases suggest that regulating the manner of speech may result in speakers not always being able to convey their message through their preferred means of doing so. The government may regulate the time, place, or manner of speech only as long as the regulation (1) is content-neutral; (2) is narrowly tailored to serve a significant governmental interest; and (3) leaves open “ample alternative channels” to communicate the information.<sup>103</sup> Content neutrality is indeed a cornerstone of First Amendment protection. “[A]bove all else, the First Amendment means that government has no power to restrict expression because of its message, its ideas, its subject matter, or its content.”<sup>104</sup> Mandating automation disclosure would also be content-neutral, as it would apply regardless of the speaker’s message. Moreover, such a requirement generally does not restrict expression in any way—it actually increases the amount of information. The requirement of government neutrality would thus most likely be satisfied.

The requirement that the regulation be justified by a significant government interest, however, is not straightforward. A significant government interest could be anything from reducing crime<sup>105</sup> to national security<sup>106</sup> to protecting citizens from the unwelcome and excessive noise of a rock concert.<sup>107</sup> But despite the breadth and variety of their justification, government interventions tend to be context-specific. Existing FTC regulations suggest that an automation disclosure requirement for commercial bots would be permissible. The FTC requires celebrities and “influencers” on social media to disclose material connections with a company when they endorse a product, such as the fact that the company is paying them.<sup>108</sup> By requiring social media users to disclose the fact that they receive a financial benefit for their posts, the FTC aims to promote “the basic truth-in-advertising principle that endorsements must be honest and not misleading.”<sup>109</sup> Requiring disclosure of the fact that a speaker is automated in the commercial context seems similarly reasonable in light of the way we communicate in the digital era. These FTC requirements have little bearing in the context of other bot speech.

---

<sup>102</sup> *Clark v. Cmty. for Creative Non-Violence*, 468 U.S. 288, 297 (1984).

<sup>103</sup> *Id.* at 293 (citations omitted).

<sup>104</sup> *Police Dept. of Chicago v. Mosley*, 408 U.S. 92, 95 (1972).

<sup>105</sup> *City of Los Angeles v. Alameda Books, Inc.*, 535 U.S. 425, 429

<sup>106</sup> *Wayte v. United States*, 470 U.S. 598, 611 (1985).

<sup>107</sup> *Ward v. Rock Against Racism*, 491 U.S. 781, 796 (1989).

## **EARLY DRAFT**

In light of the widespread concern about foreign interference in the 2016 presidential election through social media bots, an automation disclosure requirement could perhaps be justified by a significant government interest with regard to political bots *in particular settings*. The line between “political” bots and private individuals expressing political views is a hazy one. Thus, while preserving free and fair elections is certainly a compelling reason for requiring political bots to disclose their bot-ness when engaged specifically in electioneering, a different justification than preserving elections would be necessary in all other political contexts.

At a minimum, any omnibus attempt to require all bots to identify themselves in all contexts would need to be a sort of Frankenstein monster of government interests, where the government enumerates one or more significant rationale *for each context* it governs. Any novel context for the use of bots that emerged would, further, require revisiting by Congress. But let us say this were possible and satisfied the courts. There are certain contexts where it is hard to imagine what government interest would justify even a non-human disclosure requirement.

Art bots furnish a good example. Artists may express themselves through the intentional haziness of social media accounts that make us ask, “is it a bot or not?” It is difficult to imagine what government interest would justify the obstruction of that creativity. This inquiry is closely tethered to the final prong of time, place, and manner analysis, which asks whether the speaker has alternate channels of communication available to convey her message. In the artistic context, it seems unlikely that an algorithmic artist whose work hinges on the uncertainty of whether her account is human-run could effectively communicate her message through alternative channels of communication. This requirement is less problematic for commercial bots, as there are ample alternative means of communicating commercial advertisements. Moreover, a disclosure requirement seems unlikely to detract from a commercial speaker’s message in the first place. That is not the case, however, for artistic bots, whose premise often rests on the ambiguity of their place on the robot-human spectrum.

### *b. Problems with enforcement*

Assume the government could justify a blanket disclosure requirement, or a series of context-specific rules. Implementation issues might arise that functionally covert the regulation from a time/manner/place or mere information requirement to an instance of unmasking. Imagine if a social media platform, in an attempt to comply with such a regulation, incorrectly labeled an account that did not include detailed personal information as a bot. How would the user go about getting the label removed? The user would likely need to do more than simply check a box that says

## **EARLY DRAFT**

“I am not a robot”—a bot could do that! In order to confirm her humanness, the social media user might be required to provide the platform with information about herself to such a degree that she would no longer be effectively anonymous. This is also so if the government or another citizen suspects that a non-labeled account is automated. Absent a government issued “human-ID,” the person behind the accused account would need to unmask herself to avoid sanction or censure.

This dynamic plays into greater platform-driven censorship more generally. For if every account must identify itself as non-human, it becomes trivial for platforms to rid themselves of autonomous speech entirely. Knowing extensive details about the people using their platforms gives social media platforms even greater power than they already have to censor. Censorship on social media is already problematic in many ways.<sup>110</sup> Empowering this on a broader scale greatly increases the potential risk of restricting free speech and brings us to a discussion of Cuban’s proposal that social media platforms confirm that a single human is behind every account.

### *c. Mandatory disclosure of the man behind the robot curtain*

Many would argue that requiring an automated speaker to disclose that it is not human does not go far enough. Requiring further information would exacerbate the above problems and present new ones. The Honest Ads Act, for example, would subject election advertisements on social media to the same disclosure requirements as television and radio advertisements.<sup>111</sup> While such a law might marginally increase transparency in political discourse on social media, it would likely fall short of its goal. As a more forceful alternative, Congress might consider a proposal like Cuban’s that would require all automated speakers to clearly disclose who is funding and operating them. Under such a law, social media users would not only know that an account is a bot rather than a real person, but would see which entities are operating the bots they encounter.

Ultimately, requiring automated speakers such as Twitter bots and robocallers to disclose the identity of who funds and operates them would likely be constitutional in the context of commercial speakers. This would mirror existing requirements about nutrition information, manufacturer information, and other commercial information seen in many commercial speech contexts. If, however, the disclosure requirements were so extensive that they were intended not to provide

---

<sup>110</sup> cite to Kate Klonick’s paper

<sup>111</sup> Ali Breland, “New Bill on Election Meddling Aims to Increase Transparency From Social Media,” THE HILL, <http://thehill.com/policy/technology/356065-senate-dems-want-curb-election-interference-with-new-bill> (Oct. 18, 2017).

## **EARLY DRAFT**

consumers with information but rather to shape behavior, then they might cross into the impermissible region of compelled speech.<sup>112</sup>

Whether such a requirement would be constitutional in the context of political speech depends on where courts delineate the boundaries of “electoral speech.” Two powerful opposing forces are at work in this arena. On one hand, political speech receives more protection than any other. Historical protection of anonymity is derived from a recognition that oppressed groups often must seek refuge in anonymity in order to make their voices heard. This suggests that the right to speak anonymously is at its strongest in the political context. On the other hand, the government’s interest in protecting the integrity of elections is incredibly strong. To prevent deception and confusion of voters, disclosure requirements may be justified. This suggests that unmasking requirements would be more easily justified in the political context than elsewhere. Contemporary public concern specific to the role of bots in the 2016 election supports this viewpoint.

The constitutionality of an unmasking requirement outside the commercial and electoral contexts seems unlikely. The government would be hard-pressed to articulate a sufficiently compelling interest that would justify such a broad regulation. Narrow tailoring is an essential element of such regulations, and a general unmasking requirement for all bots would likely not be narrowly tailored enough to advance any interests the government might raise.

## **Conclusion**

As calls for regulation of bot speech grow louder, particular with regard to political bots on social media, legislators and courts will have to decide how to approach regulation. Government actors may not heed Etzioni’s and Cuban’s calls for regulation, and the Honest Ads Act may die in committee or a losing vote. This piece takes these and likely future proposals seriously, however, and considers whether they would withstand First Amendment review. We find that while a superficial analysis may not see a problem with requiring bots be identified as such, the doctrinal and pragmatic consequences of a universal law could well offend free speech doctrine and values.

It is worth noting that even carefully drafted legislation that would survive First Amendment scrutiny may not accomplish all that it is intended to in this arena. Much of the harm bots cause, especially in the political context, is felt only in the aggregate. A single bot tweet is not the problem; it is tweets by thousands of bots that shift public discourse and legitimize reactionary viewpoints that endanger the integrity of our elections. It is unclear how mandatory automation disclosures or

---

<sup>112</sup> See generally Calo, *supra* note X.

## **EARLY DRAFT**

unmasking requirements would combat the aggregate effect of bot activity online. Furthermore, it is possible that technology itself will solve these problems more quickly and effectively than the law could. Innovators at universities around the country are working to develop bot-detector browser extensions that could help concerned social media users identify bots.<sup>113</sup>

Law will have its role to play. Emerging technology tends to expose the cracks in First Amendment jurisprudence,<sup>114</sup> and proposed regulation of artificially intelligent “speakers” such as social media bots and robo-callers has the potential to take the First Amendment into uncharted territory.

---

<sup>113</sup> Lauren Smiley, “The College Kids Doing What Twitter Won’t,” WIRED, <https://www.wired.com/story/the-college-kids-doing-what-twitter-wont/> (Nov. 1, 2017).

<sup>114</sup> See, e.g., *Brown v. Entertainment Merchants Ass’n*.