

U.C.L.A. Law Review

Regulating Bot Speech

Madeline Lamo & Ryan Calo

ABSTRACT

We live in a world of artificial speakers with real impact. So-called “bots” foment political strife, skew online discourse, and manipulate the marketplace. Concerns over bot speech have led prominent figures in the world of technology to call for regulations in response to the unique threats bots pose. Recently, legislators have begun to heed these calls, drafting laws that would require online bots to clearly indicate that they are not human.

This work is the first to consider how efforts to regulate bots might run afoul of the First Amendment. At first blush, requiring a bot to self-disclose raises little in the way of free speech concerns—it does not censor speech as such, nor does it unmask the identity of the person behind the automated account. However, a deeper analysis reveals several areas of First Amendment tension. Bot disclosure laws fit poorly with the state’s stated goals, risk unmasking anonymous speakers in the enforcement process, and create a scaffolding for censorship by private actors and other governments.

Ultimately bots represent a diverse and emerging medium of speech. Their use for mischief should not overshadow their novel capacity to inform, entertain, and critique. We conclude by urging society to proceed with caution in regulating bots, lest we inadvertently curtail a new, unfolding form of expression.

AUTHORS

Madeline Lamo is a Hazelton Fellow at the University of Washington Tech Policy Lab. Ryan Calo is a Lane Powell and D. Wayne Gittinger Associate Professor at the University of Washington School of Law. The authors would like to thank participants at the 7th Annual Robotics Law and Policy Workshop (We Robot) at Stanford Law School as well as the *New York Times* and Electronic Frontier Foundation for their engagement with our work.



TABLE OF CONTENTS

INTRODUCTION.....	990
I. BOTS: A TAXONOMY.....	993
A. Commercial Bots.....	995
B. Political Bots.....	997
C. Creative Bots.....	1001
II. THE COVERAGE QUESTION: IS BOT SPEECH “SPEECH”?.....	1003
III. THE PROTECTION QUESTION: CAN BOTS BE FORCED TO SELF-DISCLOSE?.....	1008
A. Justifying Disclosure.....	1009
1. Compelled Speech.....	1010
2. Time, Manner, or Place.....	1012
3. Means and Ends.....	1014
B. Unmasking Through Enforcement.....	1018
C. Automating Censorship.....	1024
IV. REGULATING BOT SPEECH: A CASE FOR CAUTION.....	1025
CONCLUSION.....	1028

INTRODUCTION

In May of 2018, Google released a powerful new assistive technology called Duplex.¹ Duplex is a software system that can call up your hairdresser or favorite restaurant and book you an appointment. What makes Duplex both exciting and controversial is that the system mimics a real person—down to the “ums” and pauses that characterize human speech.² The receptionist who books an appointment with Duplex is unlikely to know he is speaking with a machine unless Google so discloses.

The concept of a machine posing as a person is hardly new. Recent developments—technological, as well as political and economic—have elevated attention to automated agents, or “bots.” Social media in particular has proved a fertile ground for this phenomenon. The presence of millions and millions of automated accounts on Twitter, Facebook, and other platforms can be disconcerting and even dangerous.³ Bots can create an appearance of false consensus, make a candidate or idea seem more popular than the reality, and even hijack attempts at genuine dialogue and community building.⁴ There is evidence that bots created in Russia played a significant role in spreading disinformation during the 2016 presidential election:⁵ During the 2016 election, bot activity pushing Trump campaign hashtags was timed to coincide with critical events like

1. Drew Harwell, *A Google Program Can Pass as a Human on the Phone. Should It Be Required to Tell People It's a Machine?*, WASH. POST: THE SWITCH (May 8, 2018), https://www.washingtonpost.com/news/the-switch/wp/2018/05/08/a-google-program-can-pass-as-a-human-on-the-phone-should-it-be-required-to-tell-people-its-a-machine/?noredirect=on&utm_term=.6035140c540c [https://perma.cc/7TK7-YW9F].

2. *Id.*

3. See Craig Timber & Elizabeth Dvoskin, *Twitter Is Sweeping Out Fake Accounts Like Never Before, Putting User Growth at Risk*, WASH. POST (July 6, 2018), https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/?utm_term=.8abcf09cf2a [https://perma.cc/G6JG-3YD5].

4. See *infra* Subpart I.B.

5. Gerrit De Vynck & Selina Wang, *Russian Bots Retweeted Trump's Twitter 470,000 Times*, BLOOMBERG (Jan. 26, 2018, 3:41 PM), <https://www.bloomberg.com/news/articles/2018-01-26/twitter-says-russian-linked-bots-retweeted-trump-470-000-times> [https://perma.cc/9FAC-8KBH]; Chris Geidner, *Federal Prosecutors File the First Charges Alleging 2018 Russian Election Interference*, BUZZFEED NEWS (Oct. 19, 2018, 3:01 PM), <https://www.buzzfeednews.com/article/chrisgeidner/first-charges-alleging-2018-election-interference> [https://perma.cc/9QSS-RVXX]; see also ROBERT S. MUELLER, III, U.S. DEP'T OF JUSTICE, REPORT ON THE INVESTIGATION INTO RUSSIAN INTERFERENCE IN THE 2016 PRESIDENTIAL ELECTION 28–29 (2019).

debates and election day itself.⁶ Bots continue to foment political and cultural discord as of this writing.

These and other concerns have led to calls for the government to step in. Commentators, including the head of the Allen Institute for Artificial Intelligence writing in the *New York Times*, publicly urge regulators to adopt a requirement that all bots identify themselves as nonhuman.⁷ Some regulators are heeding that call. California passed a law making it unlawful for bots to engage in marketing or electioneering without disclosing that they are not human.⁸ The U.S. Senate is also weighing a possible blanket bot disclosure law.⁹

This Article examines how mandatory disclosure laws that prohibit bots from operating unless they identify themselves as nonhuman might fare under principles of free expression. The question is an interesting one, in part because a cursory First Amendment analysis obscures a deeper tension. Requiring a bot merely to acknowledge that it is a bot does not appear at first blush to implicate censorship or threaten the right to anonymous speech. However, requiring across-the-board disclosure in response to specific concerns about political and commercial dangers creates a risk of overbroad regulation. Furthermore, the unintended consequence of bot disclosure laws for speech and privacy could be significant—for example, where a person accused of running an illicit bot has to prove they are human by revealing their actual identity. Crafting a narrowly tailored, enforceable law requiring bot disclosure turns out to be much harder than proponents realize, and indeed threatens to curtail an emerging form of expression.

While a series of recent contributions has assessed whether bot speech is covered by the First Amendment,¹⁰ this Article is among the first to discuss the protections offered by the First Amendment in light of coverage.¹¹ This Article's

6. See generally BENCE KOLLANYI ET AL., COMPUTATIONAL PROPAGANDA PROJECT, BOTS AND AUTOMATION OVER TWITTER DURING U.S. ELECTION 3–4 (2016), <http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2016/11/Data-Memo-US-Election.pdf> [<https://perma.cc/TNY3-PZBM>].

7. Oren Etzioni, *How to Regulate Artificial Intelligence*, N.Y. TIMES (Sept. 1, 2017), <https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html>.

8. S.B. 1001, Cal. Leg., 2017–2018 Sess. (Cal. 2018). California's law goes into effect in July of 2019.

9. Bot Disclosure and Accountability Act, S. 3127, 115th Cong. (2018).

10. See, e.g., Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445 (2013); Toni M. Massaro & Helen Norton, *Siri-ously? Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169 (2016); Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495 (2013).

11. Frederick Schauer discusses the distinction between First Amendment coverage and protection at length in *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 HARV. L. REV. 1765 (2004). He characterizes coverage as the “boundaries” or “scope” of First Amendment application. *Id.* at 1768–69. Once the threshold determination of coverage has been made, Schauer then turns to the question of

analysis of real-world speech regulation suggests the potentially unique ways First Amendment law may come to interact with automated speakers—an interaction that is particularly interesting in light of the new forms of expression bots permit and the very ambiguity surrounding their nature.

The Article proceeds as follows. Part I delineates the inquiry by describing bots and their common uses, benefits, and harms. Like the Internet itself, bots enable “cheap speech”—communication that is easily and inexpensively transmitted to the masses.¹² They can also generate surprising speech—claims about the world that even the programmer did not anticipate. And they can create speech that fails to fall neatly into any particular category. To paraphrase a now classic saying: On the Internet, nobody knows you’re a bot.¹³ Bots’ inherent ambiguity not only creates potential for bot harm, but also allows for interesting new possibilities for communication.

Parts II and III analyze the free speech concerns arising from proposed regulations that force bots to identify themselves as nonhuman. We identify three areas of potential First Amendment infringement, including an inadequate fit between the government’s stated ends and its chosen regulatory means; the prospect that real people will have to reveal their identities to prove they are human; and the prospect that bot disclosure requirements will provide a scaffolding for private or foreign censorship of an emerging category of speech.

A final Part IV discusses the ramifications of bots, and early attempts to regulate them, for free speech more generally. We propose a series of principles aimed at creating space for this still emerging medium of speech, urging legislators to consider creative applications of existing law, to legislate in a piecemeal fashion, and to be mindful of the complexities inherent in the enforcement of a bot disclosure law.

what protection the First Amendment offers: “When the First Amendment does show up, the full arsenal of First Amendment rules, principles, standards, distinctions, presumptions, tools, factors, and three-part tests becomes available to determine whether the particular speech will actually wind up being protected.” *Id.* at 1769.

12. Eugene Volokh, *Cheap Speech and What It Will Do*, 104 YALE L.J. 1805, 1807 (1995).

13. See Glenn Fleishman, *Cartoon Captures Spirit of the Internet*, N.Y. TIMES (Dec. 14, 2000), <https://www.nytimes.com/2000/12/14/technology/cartoon-captures-spirit-of-the-internet.html> (“On the Internet, nobody knows you’re a dog.”).

I. BOTS: A TAXONOMY

A variation of “robot,” the term *bot* is nontrivial to define.¹⁴ Whatever their specific function or level of complexity, bots are software programs that run according to instructions.¹⁵ We use the term here to refer to automated agents that initiate communication online, by phone, or through other technologically mediated means. We thereby exclude a wide variety of automated activity, including the many software programs that “scrape” websites for information or that perform automated functions such as purchasing stocks or event tickets. A broad definition of bots also sweeps in automated assistants, such as Siri, Alexa, Cortana, or Google’s forthcoming Duplex, which act as interfaces between devices and users.¹⁶ Bot disclosure laws could potentially reach these services as well.¹⁷ The focus of our analysis, however, and that of recent calls for legislation, is on interactive agents engaging in spontaneous communication with the public.

Bots offer certain interesting new affordances to programmers and users. They are an explosive source of what Eugene Volokh referred to as “cheap speech,” that is, speech with low barriers and costs to the speaker.¹⁸ A bot could, for example, find every single reference on Twitter to the famous scene in *Star Wars* where Greedo confronts Han Solo and comment, accurately, that Han shot first.¹⁹ An individual person could not. Moreover, bots display emergent behavior, meaning that they regularly generate content that neither the programmer nor the user of the bot could have clearly anticipated in advance.

14. For a working definition of robot, see Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 529–32 (2015). See also *infra* Part III (discussing the challenges for legislation posed by definitions).

15. See Matt Francis, *4 Things You Absolutely Need to Know About Software Bots*, WORKINGMOUSE (June 27, 2017), <https://workingmouse.com.au/innovation/4-things-you-absolutely-need-to-know-about-software-bots> [https://perma.cc/C87Y-3DPL].

16. Cf. Robert Gorwa & Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, POL’Y & INTERNET 4 (2018), <https://arxiv.org/pdf/1801.06863.pdf>.

17. See Harwell, *supra* note 1 (discussing concerns about whether Duplex should have to disclose that it is a bot). No proposed legislation to date has addressed artificially intelligent assistants, however.

18. Volokh, *supra* note 12, at 1806–07. “Cheap speech” refers to forms of communication that can reach many people without the involvement of formerly necessary institutions and resources. *Id.* Writing in 1995, Volokh theorized that advancements in technology would greatly lower the logistical barriers to entry into various communicative marketplaces—music, news, books, and more, so even speakers without the wealth and power to secure printing presses and the like could disseminate their message widely. *Id.*

19. Cf. Kristopher Tapley, *How the Kasdans Settled the Biggest Debate in ‘Star Wars’ History With ‘Solo’*, VARIETY (May 26, 2018, 8:00 AM), <https://variety.com/2018/film/news/solo-a-star-wars-story-han-shot-first-lawrence-jonathan-kasdan-1202812392> [https://perma.cc/SK32-R5F4].

Although Microsoft did not program its Twitter bot, Tay, to use hateful language, the bot infamously devolved into a “hate-spewing Nazi” after interacting with other Twitter users for mere hours.²⁰

Some bots can, subject to caveats and constraints, pose as human beings, simulating a certain degree of interpersonal communication on a particular topic. As far back as the 1960s, Joseph Weizenbaum illustrated the tendency we have to anthropomorphize bots with his chatbot Eliza.²¹ Eliza was an automated “therapist” that engaged subjects (mostly graduate students) with Rogerian questions such as, “How does that make you feel?” The depth with which his students engaged Eliza so troubled Weizenbaum that he wrote a book critiquing social techniques in computing.²²

Bots, even narrowly defined, are many and varied. So are their effects. Commercial bots can alert consumers to better prices and new products and services, but they may also gather information about consumers, spam people with offers, or denigrate commercial competitors.²³ Political bots can widen participation but also manufacture a false sense of consensus and sow discord.²⁴ Other bots delight and entertain, even enlighten.²⁵ The sheer variety of bots, and the variety of their uses and effects, has repercussions for the constitutional analysis that follows in Part II. By way of illustration, here we discuss three categories of bots: commercial, political, and creative.

20. April Glaser, *Bots Need to Learn Some Manners, and It's On Us to Teach Them*, WIRED (Apr. 13, 2016, 2:55 PM), <https://www.wired.com/2016/04/bots-emergent-behavior-deception> [<https://perma.cc/B6ZU-9J8H>]; see also Calo, *supra* note 14, at 538–39; *infra* note 85 (giving the example of a Twitter bot that threatened an event in Amsterdam).

21. Joseph Weizenbaum, *ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine*, 9 COMPUTATIONAL LINGUISTICS 26 (1966).

22. JOSEPH WEIZENBAUM, *COMPUTER POWER AND HUMAN REASON: FROM JUDGMENT TO CALCULATION* 6–7 (1976). Weizenbaum reported seeing long transcripts of chats with Eliza that included students’ intimate personal details. *Id.* at 7.

23. See *infra* Subpart I.A.

24. Lutz Finger, *Do Evil—The Business of Social Media Bots*, FORBES (Feb. 17, 2015, 9:41 AM), <https://www.forbes.com/sites/lutzfinger/2015/02/17/do-evil-the-business-of-social-media-bots/#45ab6ca4fb58> [<https://perma.cc/E4LK-7694>]; Samuel C. Woolley & Douglas R. Guilbeault, *Computational Propaganda in the United States of America: Manufacturing Consensus Online* 8–9 (Computational Propaganda Res. Project, Working Paper No. 2017.5, 2017), <http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf> [<https://perma.cc/4TQM-MKP5>]; see also *infra* Subpart I.B.

25. See *infra* Subpart I.C.

A. Commercial Bots

Among the first sustained discussion of bots in legal discourse is Ian Kerr's 2004 critique of ELLEgirlBuddy, an instant messenger bot designed to chat with teen girls online and encourage them to visit Ellegirl.com.²⁶ ELLEgirlBuddy posed as a teenager on various instant-messaging services of the time.²⁷ While few mistook the bot for a real person, many interacted with the software extensively.²⁸ As Kerr observes, this permitted Ellegirl.com not only to drive traffic, but to glean insights about its advertising base of teenage girls by analyzing transcripts of their chats with ELLEgirlBuddy.²⁹

ELLEgirlBuddy is long retired, but commercial bots have evolved and proliferated. We regularly encounter them in the form of interactive voice 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.³⁰ Many corporate entities use automated social media accounts, from Puma to Coca-Cola to the New England Patriots.³¹ 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*),³² the use of bots largely allows corporations to promote their brands online without incident.

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 consumers to handle problems on their own, which many consumers prefer over other means of communication when engaging with a company.³³ Commercial chatbots are

26. Ian R. Kerr, *Bots, Babes and the Californication of Commerce*, 1 U. OTTAWA L. & TECH. J. 285, 313 (2004).

27. *Id.* at 313–14.

28. *Id.*

29. *Id.* at 313–15.

30. George Dvorsky, *Freakishly Realistic Telemarketing Robots Are Denying They're Robots*, GIZMODO: 109 (Dec. 11, 2013, 10:20 AM), <https://io9.gizmodo.com/freakishly-realistic-telemarketing-robots-are-denying-t-1481050295> [<https://perma.cc/K2BS-QJDE>].

31. See Tanya Dua, *5 of the Biggest Bot Fails by Brand on Twitter*, DIGIDAY (Feb. 9, 2015), <https://digiday.com/marketing/5-biggest-bot-fails-brands-twitter> [<https://perma.cc/EB9V-ATFH>].

32. *Id.*

33. ASPECT SOFTWARE, 2016 ASPECT CONSUMER EXPERIENCE INDEX 6, https://www.aspect.com/globalassets/2016-aspect-consumer-experience-index-survey_index-results-final.pdf [<https://perma.cc/VQ5H-UD4C>] (finding that two-thirds of consumers surveyed felt good about handling transactions without having to speak to a person).

available at all hours of the day.³⁴ They never get cranky or frustrated, even when dealing with the most difficult customers.³⁵ Perhaps most importantly, they can significantly decrease the wait for customer service assistance.³⁶

Yet commercial bots can also cause harm, primarily by tricking and confusing consumers.³⁷ Robocallers may deny that they are automated,³⁸ 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.³⁹ Vulnerable populations such as the elderly are particularly susceptible to scamming by robocallers.⁴⁰ The Federal Communications Commission recognizes the threat that robocalls pose to consumers and has passed regulations against such practices.⁴¹ The Federal Trade Commission (FTC) has won several lawsuits against companies with predatory robocall practices.⁴² In late 2017, FTC representatives testified before the U.S. Senate Special Committee on Aging to discuss the specific threat robocalls pose to the elderly.⁴³

34. Shep Hyken, *AI and Chatbots Are Transforming the Customer Experience*, FORBES (July 15, 2017, 9:02 AM), <https://www.forbes.com/sites/shephyken/2017/07/15/ai-and-chatbots-are-transforming-the-customer-experience/#657ee5a841f7> [<https://perma.cc/C3DG-7SA5>].

35. *Id.*

36. *Id.* Of course, the quality may suffer.

37. See generally Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785 (2015).

38. Dvorsky, *supra* note 30.

39. See, e.g., Hanna Landman, *Phone Scams: Preying on the Elderly*, AVACARE MED.: BLOG (June 12, 2017), <https://avacaremedical.com/blog/phone-scams-preying-elderly.html> [<https://perma.cc/JW6G-6TXC>]; *IRS Urges Public to Stay Alert for Scam Phone Calls*, IRS (Oct. 21, 2015), <https://www.irs.gov/newsroom/irs-urges-public-to-stay-alert-for-scam-phone-calls> [<https://perma.cc/K3YG-D4MC>].

40. Landman, *supra* note 39; Lois Greisman, Assoc. Dir. of the Div. of Mktg. Practices, Bureau of Consumer Prot., Prepared Statement of the Federal Trade Commission Before the United States Senate Special Committee on Aging (Oct. 4, 2017).

41. 47 C.F.R. § 64.1200–.1202 (2018).

42. See, e.g., *FTC Providing \$4 Million in Full Refunds to People Tricked Into Buying Bogus “Extended Auto Warranties”*, FTC (July 19, 2016), <https://www.ftc.gov/news-events/press-releases/2016/07/ftc-providing-4-million-full-refunds-people-tricked-buying-bogus> [<https://perma.cc/XBD6-32FG>]; *FTC and State of Florida Win Summary Judgment: Court Orders Ringleader of Debt-Relief Scam to Pay \$23 Million and Imposes Industry Bans*, FTC (Dec. 14, 2018), <https://www.ftc.gov/news-events/press-releases/2018/12/ftc-state-florida-win-summary-judgment-court-orders-ringleader> [<https://perma.cc/DP8G-2AUS>]; *Sales Lead Generators Fined and Barred From Violating FTC’s Telemarketing Sales Rule*, FTC (Nov. 1, 2016), <https://www.ftc.gov/news-events/press-releases/2016/11/sales-lead-generators-fined-barred-violating-ftcs-telemarketing> [<http://perma.cc/Q5ZE-9DGA>].

43. Greisman, *supra* note 40.

Bots can also skew the marketplace, for instance, by creating confusion in product reviews.⁴⁴ 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.⁴⁵ This can mislead consumers and encourage them to purchase terrible products with fraudulent positive rankings and reviews. While major online retailers such as Amazon try to fight bot reviews, they struggle to do so.⁴⁶

We anticipate that, as such technologies improve and proliferate, new distortions will emerge. As noted earlier, Google recently debuted Duplex, its new AI-powered personal assistant.⁴⁷ Unlike other digital voice assistants, Duplex can place outgoing calls and interact with humans on the other line, though only in certain narrow contexts such as scheduling restaurant reservations or hair appointments.⁴⁸ Duplex received significant attention for its sophistication level, as the demo calls Google released sounded impressively humanlike.⁴⁹ While it remains to be seen if Duplex fares as well in the real world as its demos suggested, it seems likely that such technology may skew marketplaces in ways yet unknown. Duplex opens up a door to new possibilities that other tech firms are likely to walk through.

B. 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.⁵⁰ Research by the Oxford Internet Institute

44. Shareen Pathak, *Amazon Reviews Have a Bot Problem*, DIGIDAY (Sept. 18, 2017), <https://digiday.com/marketing/amazon-reviews-bot-problem> [<https://perma.cc/DJK4-ZKWP>].

45. *Id.*

46. *Id.*

47. Harwell, *supra* note 1.

48. *Id.*

49. *See id.*

50. *Update on Twitter's Review of the 2016 U.S. Election*, TWITTER: BLOG (Jan. 19, 2018), https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html [<https://perma.cc/B6VY-VQDC>].

shows that pro-Trump Twitter bots were four times as active as pro-Clinton bots during the first presidential debate.⁵¹ This margin increased to a five-to-one pro-Trump to pro-Clinton bot activity ratio by election day.⁵² Some even argue that the long-term goal of this interference was to undermine democracy more broadly, rather than to advocate for or against a particular candidate.⁵³ While some of these bots shared seemingly original content, others primarily magnified existing content by “retweeting” posts, following prominent accounts, and posting frequently about certain topics in order to make them “trend.”⁵⁴ The full effect of this type of bot use has not yet been quantified, but it seems clear that political bots may be used to skew discourse, to make certain ideas and individuals appear more popular than they would be otherwise, and to stir up dissent and discord.

For all of their dangers and flaws, however, political bots are in many ways an extension of other forms of media and worthy of similar consideration. 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 U.S. 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.”⁵⁵ John F. Kennedy won the support of the American electorate by appearing “robust and confident” in the first televised presidential debates.⁵⁶ 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

51. KOLLANYI ET AL., *supra* note 6, at 4.

52. *Id.*

53. Seema Metha, *Rep. Adam Schiff Says Alleged Russian Meddling in Election Was an Effort to Destroy American Democracy*, L.A. TIMES: ESSENTIAL POL. (May 30, 2017, 1:42 PM), <http://www.latimes.com/politics/essential/la-pol-ca-essential-politics-updates-schiff-argues-russian-intervention-in-1496173190-htmstory.html> [<https://perma.cc/KFJ8-ELKP>].

54. See Gerrit De Vynck & Selina Wang, *Russian Bots Retweeted Trump’s Twitter 470,000 Times*, BLOOMBERG (Jan. 26, 2018, 3:41 PM), <https://www.bloomberg.com/news/articles/2018-01-26/twitter-says-russian-linked-bots-retweeted-trump-470-000-times>; Finger, *supra* note 24; see also Woolley & Guilbeault, *supra* note 24, at 10–11.

55. Adrienne LaFrance, *Donald Trump Is Testing Twitter’s Harassment Policy*, ATLANTIC (July 2, 2017), <https://www.theatlantic.com/politics/archive/2017/07/the-president-of-the-united-states-is-testing-twitters-harassment-policy/532497> [<https://perma.cc/AAR2-PUUZ>].

56. Kenneth T. Walsh, *JFK: First TV President*, U.S. NEWS (Nov. 20, 2013, 7:48 AM), <https://www.usnews.com/news/blogs/ken-walshs-washington/2013/11/20/jfk-first-tv-president>.

demonstrations against powerful authoritarian governments.⁵⁷ Bots may be a natural result of ever-evolving technology, and their unique qualities make them a uniquely powerful means of communication in the political context.

Bots' distinct qualities also allow them to engage in nefarious online activity that is difficult to quantify and control, however. Most visibly, bots can support coordinated campaigns of disinformation. Perhaps most famously, U.S. officials have accused the Russian government of using social media—including social media bots—to interfere in American elections and sow discord.⁵⁸ 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 lawmakers to grill social media executives at hearings in recent months and to propose the disclosure requirements we highlighted in the Introduction.⁵⁹

Bots can also engage in online harassment at an unprecedented scale. By automating “trolling,” that is, the practice of criticizing or threatening certain speakers (especially women and people of color) in response to their views, bots can exacerbate highly problematic trends of online hate speech and abuse.⁶⁰ Bots can harass or “troll” at scale. They can sink a useful hashtag by 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 #guncontrolnow—but accompanied the hashtag with messages, links, and images suggesting that stricter gun control laws would *not* have prevented the tragedy.⁶¹

57. Amy Mitchell et al., *The Role of Social Media in the Arab Uprisings*, PEW RES. CTR. (Nov. 28, 2012), <http://www.journalism.org/2012/11/28/role-social-media-arab-uprisings> [<https://perma.cc/Q5ZE-9DGA>].

58. Indictment ¶¶ 6–7, *United States v. Internet Research Agency LLC*, No. 1:18-cr-00032-DLF (D.D.C. Feb. 16, 2018); Taylor Hatmaker, *Special Counsel Robert Mueller Indicts Russian Bot Farm for Election Meddling*, TECHCRUNCH (Feb. 16, 2018), <https://techcrunch.com/2018/02/16/mueller-indictment-internet-research-agency-russia> [<https://perma.cc/96ET-9JQT>].

59. See Karoun Demirjian, *Top Senate Intel Democrat Proposes Measures to Counter Influence Campaigns on Social Media*, WASH. POST (July 30, 2018), https://www.washingtonpost.com/powerpost/top-senate-intel-democrat-proposes-measures-to-counter-influence-campaigns-on-social-media/2018/07/30/50de4786-9420-11e8-810c-5fa705927d54_story.html?noredirect=on&utm_term=.b0d40d97cd03 [<https://perma.cc/JF26-YUBS>]; see also Elizabeth Zima, *California Wants to Govern Bots and Police User Privacy on Social Media*, GOV'T TECH. (Feb. 23, 2018), <http://www.govtech.com/social/California-Wants-to-Govern-bots-and-Police-User-Privacy-on-Social-Media.html> [<https://perma.cc/DLK6-STCY>].

60. DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 52–55 (2014).

61. Erin Griffith, *Pro-Gun Russian Bots Flood Twitter After Parkland Shooting*, WIRED (Feb. 15, 2018, 2:00 PM), <https://www.wired.com/story/pro-gun-russian-bots-flood-twitter-after-parkland-shooting> [<https://perma.cc/A6G7-SDF4>].

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.⁶² For example, Russian-linked bots retweeted Donald Trump approximately ten times more than they retweeted Hillary Clinton in the months preceding the 2016 election, thereby dramatically increasing the overall amount of attention given to Trump’s tweets.⁶³ They can flood an administrative agency with duplicative comments, creating a sufficiently vast amount of content as to be effectively unreviewable, thereby rendering the notice-and-comment system nearly meaningless.⁶⁴ By “manufacturing consensus,”⁶⁵ bots can make fringe viewpoints appear legitimate and newsworthy.⁶⁶ 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.”⁶⁷

Relatedly, bots can increase the number of followers someone has on social media, deceiving other social media users into thinking that someone is more powerful, important, or influential than they really are.⁶⁸ In the political context, this is particularly problematic, as a high follower count may suggest that a certain individual is worth paying attention to, or that her views are popular and widely accepted. This particular set of problems is extensive enough that platforms are beginning to respond: Twitter allows automated accounts generally but has a policy against “aggressive following.”⁶⁹ Occasional bot “purges” by platforms such as Twitter may cause popular social media accounts to lose thousands or even millions of followers.⁷⁰

62. Finger, *supra* note 24.

63. De Vynck & Wang, *supra* note 54.

64. See Issie Lapowsky, *How Bots Broke the FCC’s Public Comment System*, WIRED (Nov. 28, 2017, 12:19 PM), <https://www.wired.com/story/bots-broke-fcc-public-comment-system> [<https://perma.cc/X226-4RTW>].

65. Farhad Manjoo, *How Twitter Is Being Gamed to Feed Misinformation*, N.Y. TIMES (May 31, 2017), <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>.

66. ALICE MARWICK & REBECCA LEWIS, DATA & SOC’Y, MEDIA MANIPULATION AND DISINFORMATION ONLINE 38 (2017) https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf [<https://perma.cc/X63S-Y5NW>].

67. Sam Earle, *Trolls, Bots, and Fake News: The Mysterious World of Social Media Manipulation*, NEWSWEEK (Oct. 14, 2017, 8:40 AM), <http://www.newsweek.com/trolls-bots-and-fake-news-dark-and-mysterious-world-social-media-manipulation-682155> [<https://perma.cc/HCP2-MXM7>].

68. Finger, *supra* note 24.

69. *The Twitter Rules*, TWITTER, <https://help.twitter.com/en/rules-and-policies/twitter-rules> [<https://perma.cc/DWY4-8QFS>].

70. Kerry Flynn, *Twitter Influencers Suspect a ‘Bot Purge’*, MASHABLE (Jan. 29, 2018), <https://mashable.com/2018/01/29/twitter-bots-purge-influencers-accounts/#p8stq6jiPiqF> [<https://perma.cc/TK8K-AEKB>].

C. Creative 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.⁷¹ 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 worst deal ever'—for Dale Earnhardt Jr."⁷² They can also be informative (@earthquakebot tweets when an earthquake of 5.0 or greater occurs),⁷³ create art (@greatartbot tweets out a new piece of computer-generated artwork every four hours⁷⁴ and now-defunct @pixelsorter resorted the pixels in images users sent it to create beautiful, soothing images),⁷⁵ and even identify the poetry that humans unintentionally tweet (@accidental575: "I am a robot / that finds haikus on Twitter / made by accident";⁷⁶ @pentametron: "With algorithms subtle and discreet / I seek iambic writings to retweet").⁷⁷

Another inventive use of bots comes in the area of academic research.⁷⁸ Innovative researchers have begun using bots to gather information about online activity.⁷⁹ In conducting this research, it may be essential to develop bots that appear convincingly human. For example, political scientist Kevin Munger used

-
71. Leon Neyfakh, *The Botmaker Who Sees Through the Internet*, BOS. GLOBE (Jan. 24, 2014), <https://www.bostonglobe.com/ideas/2014/01/24/the-botmaker-who-sees-through-internet/V7Qn7HU8TPPI7MSM2Tvbs/> [https://perma.cc/P8VN-6QCS].
 72. Two Headlines (@TwoHeadlines), TWITTER (Feb. 19, 2018, 9:08 AM), <https://twitter.com/TwoHeadlines/status/965634116396965893> [https://perma.cc/3SDT-GQRB].
 73. Earthquake Robot (@earthquakebot), TWITTER, <https://twitter.com/earthquakeBot> [https://perma.cc/4RSN-XE8S].
 74. Great Artist (@greatartbot), TWITTER, <https://twitter.com/greatartbot> [https://perma.cc/HF8F-H2TE].
 75. Pixel Sorter (@pixelsorter), TWITTER, <https://web.archive.org/web/20180505045456/twitter.com/pixelsorter> [https://perma.cc/6P6B-HGLQ].
 76. Accidental Haiku (@accidental575), TWITTER, <https://twitter.com/accidental575> [https://perma.cc/AF5W-Y6NC].
 77. Pentametron (@pentametron), TWITTER, <https://twitter.com/pentametron> [https://perma.cc/N7V6-6ER3].
 78. We owe this insight to Tadayoshi Kohno, a computer scientist who reminded us that researchers of all kinds use bots to gather information and test hypotheses online. See also *infra* notes 80–83 and accompanying text.
 79. See Chris Bail, *Building Apps & Bots for Social Science Research*, SUMMER INST. COMPUTATIONAL SOC. SCI. (June 17, 2018), <https://compsocialscience.github.io/summer-institute/2018/materials/day2-digital-trace-data/building-apps-bots/rmarkdown/Building%20Apps%20and%20Bots%20for%20Social%20Science%20Research.nb.html> [https://perma.cc/VZ9Z-LGTS].

humanlike Twitter bots to evaluate the effectiveness of chastisement as a response to racial harassment online.⁸⁰ Munger created Twitter bots, some of which masqueraded as black men and others as white men, with varying follower counts in each category.⁸¹ The bots responded to harassing tweets that included the n-word with “@[subject] Hey man, just remember that there are real people who are hurt when you harass them with that kind of language.”⁸² By employing this technique at scale, Munger was able to gather data indicating that these rebukes were most effective—that is, they were followed by the sharpest decrease in use of the n-word—when they came from apparent white male accounts with a high number of followers.⁸³ Munger’s study exemplifies how bots can be used to gather data about online activity and, more broadly, the ways that creative bots can have a positive social influence.

While creative bots may create genuine confusion⁸⁴ and even chaos,⁸⁵ they typically represent a harmless, imaginative format that provides artists, researchers, and others with a new tool for expression and inquiry. The fact of automation permits the botmaker to achieve an audience reach and creative scale that might be hard to accomplish otherwise. Importantly, some bots achieve their programmers’ artistic or research-driven aims best when users either believe the account is human-run or cannot tell whether an account is automated. The very ambiguity around whether the interaction constitutes genuine interpersonal connection, overt deception, or something else, generates new possibilities for storytelling and data collection.

80. Sam Machkovech, *Twitter Bots Can Reduce Racist Slurs—If People Think the Bots Are White*, ARS TECHNICA (Nov. 15, 2016, 1:57 PM), <https://arstechnica.com/science/2016/11/twitter-bots-can-reduce-racist-slurs-if-people-think-the-bots-are-white> [https://perma.cc/2FCW-EMFD]; see also Kevin Munger, *Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment*, 39 POL. BEHAV. 629, 629 (2017).

81. See Machkovech, *supra*, note 80.

82. *Id.*

83. *Id.*

84. See, e.g., Adrian Chen, *How I Found the Human Being Behind @Horse_Ebooks, the Internet’s Favorite Spambot*, GAWKER (Feb. 23, 2012, 4:15 PM), <http://gawker.com/5887697/how-i-found-the-human-being-behind-horseebots-the-internets-favorite-spambot> [https://perma.cc/U7JL-SKAJ].

85. In at least one instance, bots have issued a credible threat of violence that resulted in a police investigation. A Twitter bot once surprised its operator and creator by threatening an Amsterdam fashion show. For this story and others, see Greg Miller, *A Brief History of Robot Law*, ATLANTIC (Mar. 17, 2016), <https://www.theatlantic.com/technology/archive/2016/03/a-brief-history-of-robot-law/474156> [https://perma.cc/3KKC-CSVZ].

II. THE COVERAGE QUESTION: IS BOT SPEECH “SPEECH”?

Bots are used by a variety of people for a variety of reasons. We focus here on a single commonality: the use of bots to communicate online. Due to their communicative function, our read of the doctrine suggests that any attempt to limit the use of bots will at least implicate free speech. Such has been the emerging consensus among First Amendment scholars who have examined automated speech over the past few years.

This question of “coverage,” that is, whether automated speech by bots even falls within the ambit of the Constitution as protected speech, is only a threshold question.⁸⁶ If bot speech is covered by the First Amendment—and we agree it is—then “the full arsenal of First Amendment rules, principles, standards, distinctions, presumptions, tools, factors, and three-part tests becomes available to determine whether the particular speech will actually wind up being protected.”⁸⁷ This Part discusses the literature around the First Amendment coverage of bot speech. Part III then poses a novel question regarding the effect of First Amendment protection: If bot speech is covered, does the First Amendment permit the popular intervention of requiring bots to identify themselves as nonhuman? We conclude both that bot speech is covered and that the scope of coverage places concrete limits on anticipated bot disclosure laws.

We turn first to the threshold question of whether automated or “robot” speech is protected under the First Amendment at all. 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.”⁸⁸ While their rationales and justifications vary, there is a rough consensus among experts that automated speech such as that generated by online bots or robocallers is among the subset that falls within the Constitution’s protection.

The U.S. Supreme Court has enumerated certain categories of communicative acts that receive varying degrees of protection under the First Amendment. At its heart, this is 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.⁸⁹ Expressive conduct such as burning a flag, however, would qualify for First Amendment protection, despite the fact that it

86. See Schauer, *supra* note 11, at 1769.

87. *Id.*

88. *Id.* at 1777.

89. *Miller v. California*, 413 U.S. 15, 24 (1973).

lacks a verbal or written component.⁹⁰ 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,”⁹¹ the Supreme Court has declined to draw such a bright line.⁹² Whether bot speech receives First Amendment protection is a complex and multifaceted analysis.

That bot speech is new is not disqualifying. The Supreme Court clearly stated that First Amendment protection should not vary by speech medium, including new media that grows 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 and different medium for communication appears.”⁹³ The Supreme Court recognized First Amendment protections for the Internet in 1997 and, a decade later, to depictions of violence in video games.⁹⁴ This suggests 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.⁹⁵

Nor is bot speech categorically excludable from First Amendment coverage because it fits into a historically recognized category of so-called low-value speech. In *United States v. Stevens*,⁹⁶ decided in 2010, the Court refused to add depictions of animal cruelty to its short list of “historic and traditional categories”⁹⁷ of unprotected speech, “the prevention and punishment of which have never been thought to raise any constitutional problem.”⁹⁸ The Court also cautioned legislatures that they have no “freewheeling authority to declare new categories of speech outside the scope of the First Amendment.”⁹⁹ It did not close the door

90. *Texas v. Johnson*, 491 U.S. 397, 406 (1989).

91. Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 *IND. L.J.* 1, 20 (1971).

92. Wu, *supra* note 10, at 1507.

93. *Brown v. Entm’t Merchs. Ass’n*, 564 U.S. 786, 789–90 (2011) (quoting *Joseph Burstyn, Inc. v. Wilson*, 343 U.S. 495, 503 (1952)).

94. See *ACLU v. Reno*, 521 U.S. 844 (1997); *Brown*, 564 U.S. 786. A lower court also recognized First Amendment protection for search engine results in *Jian Zhang v. Baidu.com Inc.*, 10 F. Supp. 3d 433 (S.D.N.Y. 2014).

95. *Cf.* Wu, *supra* note 10, at 1514 (“While video games are a type of computer program, *Brown* does not declare that all computer programs fall under the coverage of the First Amendment. Yet the evolution of games and film before them are an important reminder that the technologies here described, even if not inclusions today, may someday be so declared.”).

96. 559 U.S. 460 (2010).

97. *Id.* at 468 (citing *Simon & Schuster, Inc. v. Members of N.Y. State Crime Victims Bd.*, 502 U.S. 105, 127 (1991) (Kennedy, J., concurring)).

98. *Id.* at 469 (citing *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942)).

99. *Id.* at 472.

entirely to further judicial identification of unprotected categories, however, noting that “[m]aybe there are some categories of speech that have been historically unprotected, but have not yet been specifically identified or discussed as such in our case law.”¹⁰⁰ As noted by Alexander Tsesis, the Court’s decision the following year in *Entertainment Merchants Association*—in which it found that nonthreatening depictions of violence were not a historically recognized category of low-value speech—further established that “the judiciary will not be deferential to lawmakers’ assessments about what speech is low-value and therefore unworthy of full constitutional protection.”¹⁰¹

The fact that robot speech is generally unoriginal and just some kind of repetition or re-splicing of old communications also fails to place it outside the scope of First Amendment protection. Ultimately, Stuart Benjamin writes, “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.”¹⁰² There is no requirement that speech be original, creative, or well-reasoned in order to qualify for First Amendment protection. Therefore, even bots that do not generate any kind of original content might receive protection under the First Amendment.

Furthermore, greater attenuation between a human bot creator and her bot’s speech should 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”¹⁰³ 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 should not decrease the scope of First Amendment protection.¹⁰⁴ Thus, the fact that a Twitter bot creator may not know what her creation will tweet next should not place the bot outside the protection of the First Amendment.

Finally, the First Amendment protects not only the speaker’s right to speak but the right of those who wish to read or listen to bot speech. Even when a great degree of attenuation exists between the human creator and the final speech product, the First Amendment may still protect the communication, because it

100. *Id.*

101. Alexander Tsesis, *The Categorical Free Speech Doctrine and Contextualization*, 65 EMORY L.J. 495, 500 (2015).

102. Benjamin, *supra* note 10, at 1463 n.64.

103. Wu, *supra* note 10, at 1507.

104. Benjamin, *supra* note 10, at 1464–65.

protects not only the right to speak but also the right to receive information.¹⁰⁵ Would-be listeners or readers can assert their own First Amendment rights even when the censored speaker lacks First Amendment rights of her own.¹⁰⁶ Despite the current critiques of social media bots for their role in skewing the American political dialogue online, there are in fact many enjoyable bots that make the internet a brighter, funnier, and more interesting place.¹⁰⁷ Thus, the First Amendment could, for example, protect the rights of internet users who wish to read tweets from their favorite bots.

The general consensus in the burgeoning literature seems to be that the First Amendment should apply, for a mixture of the reasons described above. Some scholars, such as Tim Wu, take a narrow, functionalist view.¹⁰⁸ Wu distinguishes between *functional communications* and actual *speech* created by machines, arguing that only the latter falls within the ambit of First Amendment protection.¹⁰⁹ Functional communications include things such as car alarms and mapping software, tools designed only “to assist the user with a task, not to express to him any ideas or influence his worldview.”¹¹⁰ By contrast, protected speech encompasses “blog posts, tweets, online photo streams, and probably slightly shorter or more symbolic expressions such as Yelp or Amazon reviews written by humans.”¹¹¹ This broad second category would seem to include the political, commercial, and artistic bot speech that we consider in this Article.

Others such as Stuart Benjamin, Helen Norton, and Toni Massaro assert that the First Amendment will apply to bot speech and algorithmically generated speech more broadly.¹¹² Noting that “there is a human mind behind all the algorithms,” Benjamin states that “the fact that an algorithm is involved does not mean that a machine is doing the talking.”¹¹³ Only when machines demonstrate

105. *Bd. of Educ. v. Pico*, 457 U.S. 853, 867 (1982).

106. *Kleindienst v. Mandel*, 408 U.S. 753, 764 (1972) (academics asserted their own First Amendment rights wishing to hear lectures from foreign communist professor whose visa application was denied).

107. See, e.g., Lainna Fader, *12 Weird, Excellent Twitter Bots Chosen by Twitter's Best Bot-Makers*, NY MAG. (Nov. 9, 2015), <http://nymag.com/selectall/2015/11/12-weirdest-funniest-smartest-twitter-bots.html> [<https://perma.cc/NE2V-JEUU>]; Zachary M. Seward, *The 17 Best Bots on Twitter*, QUARTZ (Oct. 10, 2014), <https://qz.com/279139/the-17-best-bots-on-twitter> [<https://perma.cc/2PVD-YHSJ>].

108. See generally Wu, *supra* note 10.

109. *Id.* at 1521–24.

110. *Id.* at 1525.

111. *Id.* at 1524.

112. Benjamin, *supra* note 10; Massaro & Norton, *supra* note 10; see also Toni M. Massaro, Helen Norton & Margot E. Kaminski, *Siri-ously 2.0: What Artificial Intelligence Reveals About the First Amendment*, 101 MINN. L. REV. 2481 (2017).

113. Benjamin, *supra* note 10, at 1479.

such a high level of independent volition that their human creators cannot be said to direct their substantive message should their communications fall outside the realm of First Amendment protection, he argues.¹¹⁴

Looking forward to the possibility of strong AI, Massaro and Norton argue that existing First Amendment doctrine lays the groundwork for protection of robotic speech, even when it is far-attenuated from any human creator.¹¹⁵ They emphasize that much of our existing free speech protection is predicated on the importance of “facilitat[ing] listeners’ discovery of truth and distribution of knowledge through a robust exchange of ideas,”¹¹⁶ which does not require that speakers be human. They also point to an important category of nontraditional speakers whose speech rights the Supreme Court has already recognized: corporations.¹¹⁷

Ronald Collins and David Skover argue extensively and persuasively for speech protection for robots on the theory that the First Amendment is and has always been largely predicated on audience interests, which benefit from speech irrespective of the speaker.¹¹⁸ Collins and Skover begin by observing that many new forms of speech, starting with the written word, were initially met with skepticism and even censorship. But as the utility of these new forms of communication became evident, courts began to recognize the need for protections. Thus, for Collins and Skover, the operative constitutional question is whether bots and other forms of automated speech have utility, not only to speakers, but to listeners.¹¹⁹

We agree with the emerging scholarly consensus that the First Amendment likely applies to automated speech. That does not mean, however, that all bot speech receives absolute immunity from all regulation. Rather, the fact that a bot plays a role in communication should have little bearing on the constitutionality of any proposed regulation. Instead, the constitutionality of any law regulating bots would be assessed pursuant to traditional First Amendment principles, as discussed in the next Part.

114. *Id.* at 1481–82.

115. See Massaro & Norton, *supra* note 10, at 1189. This Article generally discusses speech in the context of strong AI, which by nature would be far-attenuated from a human creator.

116. *Id.* at 1178.

117. *Id.* at 1183.

118. See generally RONALD K. L. COLLINS & DAVID M. SKOVER, *ROBOTICA: SPEECH RIGHTS AND ARTIFICIAL INTELLIGENCE* (John Berger ed., Cambridge Univ. Press 2018).

119. *Id.* at 48–64 (chapter entitled “The New Norm of Utility”).

III. THE PROTECTION QUESTION: CAN BOTS BE FORCED TO SELF-DISCLOSE?

This Part goes beyond the threshold question of coverage to examine how the constitutionality of bot speech regulation might differ from traditional speech regulation jurisprudence under the First Amendment. For purposes of discussion, we examine a generic law that would require bots to identify themselves as nonhuman in all contexts. The ingredients of the generic bill are:

Defines a “bot” as an automated account that interacts socially online.¹²⁰

Requires such a bot to clearly indicate that it is automated.

Numerous commentators have proposed such a requirement.¹²¹ Recently, the California Senate overwhelming voted to adopt a bill that applied to “make it unlawful for any person to use a bot, as defined, to communicate or interact with natural persons in California online, with the intention of misleading.”¹²² The California Assembly later modified the bill such that the law itself narrows its scope of application to only commercial bots and bots seeking to influence an election.¹²³ Meanwhile, the U.S. Senate has begun to consider a blanket bot disclosure requirement.¹²⁴ Given the diversity of potential laws, we will address the basic requirement that bots self-identify but also discuss some of the nuances of the California bill.

The proposals to regulate bot speech that motivate this paper do not amount to censorship per se. If they “abridge” speech, they do so by requiring a new 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. However, the very ambiguity around who is speaking may form an integral part of the message.¹²⁵ Moreover, it may prove impossible

120. This is to distinguish bots from, for example, web-crawlers (also known as spiders) or automated trading algorithms.

121. See, e.g., Etzioni, *supra* note 7; Mark Cuban (@mcuban), TWITTER (Jan. 28, 2018, 10:49 AM), <https://twitter.com/mcuban/status/957686987229618176?lang=en> [<https://perma.cc/EZZ9-M569>] (“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 individual account.”).

122. S.B. 1001 § 17941(a), Cal. Leg., 2017–2018 Sess. (Cal. 2018).

123. S.B. 1001 § 17941(a), Cal. Leg., 2017–2018 Sess. (Cal. 2018) (as amended by Cal. State Assemb., June 21, 2018).

124. See Bot Disclosure and Accountability Act of 2018, S. 3127, 115th Cong. (2018).

125. See *supra* Subpart I.C.

to enforce a bot-disclaimer without identifying an otherwise anonymous speaker or providing the scaffolding for censorship.

The First Amendment is, of course, a bulwark against government censorship. Outside specifically delineated circumstances, the government is not free to curtail protected speech, particularly out of disagreement with its message.¹²⁶ Thus, were a state to ban all automated speech, all automated speech on a particular topic, or even all automated speech in a particular category (such as all commercial speech), then courts would likely see this as an abridgement of speech in contravention of the First Amendment. Bot disclosure laws, however, are a far cry from censorship.¹²⁷ After all, such laws do not on their face limit the volume or content of bot speech. Rather, they require only a label informing the audience about its origins.

Still, a deeper assessment reveals a variety of entanglements with free speech doctrine. First, forcing a bot operator to reveal that his or her creation is a bot may amount to compelled speech.¹²⁸ In certain circumstances, such as the commercial or electoral context, the government may show it is justified in compelling a person to disclose some fact about herself or her products. But in others, no such justification may be available. The second is that enforcing bot disclosure laws will be difficult without compromising the right to speak anonymously. Bot disclosure proposals to date do not call for unmasking, but nor do they provide a process by which a speaker can verify she is human without also confirming her identity. Finally, without expressly censoring anyone, bot disclosure laws may nevertheless permit or encourage censorship by private parties or other jurisdictions that lack robust protections for free speech. We address each of these issues in turn.

A. Justifying Disclosure

Commentators and legislators propose variations on the same basic rule: to require automated accounts that interact socially online to identify themselves as nonhuman. The gravamen of this requirement is that a speaker—the person or organization behind the bot—must make a claim about the world they would not necessarily make on their own. Accordingly, bot disclosure laws are best

126. See, e.g., *Snyder v. Phelps*, 562 U.S. 443, 458 (2011); *Texas v. Johnson*, 491 U.S. 397, 414 (1989).

127. Censorship is a somewhat nebulous concept. See Laura E. Little, *Laughing at Censorship*, 28 *YALE J.L. & HUMAN.* 161, 162–63; Robert C. Post, *Project Report: Censorship and Silencing*, 51 *BULL. AM. ACAD. ARTS & SCI.* 32, 34 (1998). We use censorship to refer specifically to removal, deletion, or silencing of material by government actors.

128. See *infra* notes 124–129 and accompanying text.

understood as instances of compelled speech. One might also think of both disclosure laws as a time, place, or manner regulation, because they predicate the ability to participate on a special condition.¹²⁹ Under either set of doctrines, the government must have legitimate ends and pursue those ends through sufficiently careful means.¹³⁰

1. Compelled Speech

Compelled speech doctrine is something of a hodgepodge of cases unified by a single theme. Broadly, it is the notion that the government may not force a speaker to say something she does not wish to say absent an appropriate justification.¹³¹ One subset of compelled speech jurisprudence is the “right of reply” line of cases. These consider the question of whether a speech-facilitating entity (such as a newspaper or radio station) can be required to give political candidates air time or space in print to respond to criticism that the entity publishes.¹³² A second subset of compelled speech cases consider when people may be required to effectively endorse an idea they disfavor, such as driving a car required to bear New Hampshire’s “Live Free or Die” license plate slogan or being required to stand for the Pledge of Allegiance.¹³³ A third relates to campaigning, where the courts have upheld reasonable accountability measures to preserve the sanctity of elections that require attribution for advertisements.¹³⁴

Generally, the government compels speech most often and with the greatest success in the realm of commercial products and services. Commercial actors may be compelled to disclose certain information about their products. For example, agency regulations may require commercial products to conspicuously bear “the name and place of business of manufacturer, packer, or distributor.”¹³⁵

129. See *infra* Subpart III.A.2.

130. See *infra* Subpart III.A.3.

131. See generally Jonathan H. Adler, *Compelled Commercial Speech and the Consumer “Right to Know”*, 58 ARIZ. L. REV. 421 (2016); Jennifer M. Keighley, *Can You Handle the Truth? Compelled Speech and the First Amendment*, 15 U. PA. J. CONST. L. 539 (2012); Robert Post, *Compelled Commercial Speech*, 117 W. VA. L. REV. 867 (2015).

132. See, e.g., *Miami Herald Publ’g Co. v. Tornillo*, 418 U.S. 241, 243 (1974); *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 372 (1969).

133. See, e.g., *Janus v. AFSCME, Council 31*, 138 S. Ct. 2448 (2018); *Wooley v. Maynard*, 430 U.S. 705, 706 (1977); *W. Va. State Bd. of Educ. v. Barnette*, 319 U.S. 624, 627–29 (1943).

134. See, e.g., *McIntyre v. Ohio Elections Comm’n*, 514 U.S. 334, 348 (1995).

135. 21 C.F.R. § 101.5 (2018); 21 C.F.R. § 201.1 (2018). It’s current CFR: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=101.5> and <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=201.1>.

as well as other details such as nutrition information.¹³⁶ 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 discovery 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 protecting commercial speech, and requiring disclosure of truthful information promotes that goal.¹³⁷

The governing standard of review for commercial disclosures comes from *Zauderer v. Office of Disciplinary Counsel*,¹³⁸ in which the Supreme Court of Ohio held that the state could require disclosures that are “reasonably related” to preventing consumer deception.¹³⁹

Even *Zauderer*’s relaxed standard of review does not permit unlimited disclosure requirements, however, particularly when the mandatory disclosure forces the speaker to express a view adverse to her own position. The Sixth Circuit and the D.C. Circuit disagreed on whether requiring cigarette companies to include graphic visual warnings on cigarette packaging constituted a permissible “mere information” disclosure¹⁴⁰ or impermissibly “were aimed at changing behavior and hence pressed the cigarette industry into regulatory service.”¹⁴¹ The D.C. Circuit found that such requirements effectively enabled the FDA to force cigarette companies to spread “an ideological message, a point of view on how people should live their lives: that the risks from smoking outweigh the pleasure that smokers derive from it, and that smokers make bad personal decisions, and should stop smoking.”¹⁴² The graphic images thus constituted “not warnings, but

136. See generally FDA, NUTRITIONAL LABELING AND EDUCATION ACT (NLEA) REQUIREMENTS (1994), <https://www.fda.gov/iceci/inspections/inspectionguides/ucm074948.htm> [<https://perma.cc/3SFL-6ARR>].

137. *N.Y. State Rest. Ass’n v. N.Y. City Bd. of Health*, 556 F.3d 114, 132 (2d Cir. 2009) (quoting *Nat’l Elec. Mfrs. Ass’n v. Sorrell*, 272 F.3d 104, 113–14 (2d Cir. 2001)).

138. 471 U.S. 626 (1985).

139. *Id.* at 651.

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

141. Ryan Calo, *Code, Nudge, or Notice?*, 99 IOWA L. REV. 773, 777 (2014); see also *R.J. Reynolds Tobacco Co. v. FDA*, 696 F.3d 1205, 1221–22 (D.C. Cir. 2012), *overruled on limited grounds by Am. Meat Inst. v. Dep’t of Agric.*, 760 F.3d 18 (D.C. Cir. 2014) (overruling with regard to whether *Zauderer* standard applies only to correcting deception of consumers).

142. *R.J. Reynolds*, 696 F.3d at 1211.

admonitions: “[D]on’t buy or use this product.”¹⁴³ The Supreme Court recently considered the limits of commercial disclosure requirements, ruling that providers of pregnancy-related services could not be required to share information about abortion with patients.¹⁴⁴

Similarly, the Supreme Court overturned a California ruling requiring a private utility company to cede space in its own newsletter (mailed along with monthly billing statements) to organizations advocating for decreased electricity consumption in *Pacific Gas & Electric Co. v. Public Utilities Commission of California*.¹⁴⁵ The Court noted that such a requirement would force the utility company to effectively endorse speech antithetical to its own interests, contravening the First Amendment principle that “the choice to speak includes within it the choice of what not to say.”¹⁴⁶ While case law around these kinds of disclosures is sparse, existing law may generally suggest that the government cannot force commercial speakers to endorse ideas contrary to their own interests under the guise of providing consumers with mere information.

2. Time, Manner, or Place

Time, place, and manner restrictions are relatively common regulations of speech, especially in the context of public forums.¹⁴⁷ They include rules limiting, for example, the time of day when protesters may march, the volume at which a concert may be played, and the area in which a crowd may gather.¹⁴⁸ They are generally upheld by courts absent evidence that the restriction is pretextual or heavy-handed. And while the majority of Supreme Court jurisprudence regarding the constitutionality of time, place, and manner restrictions involves public forums, the Court has not expressly limited this line of analysis to restrictions on speech in public forums.¹⁴⁹

143. *Id.*

144. *Nat’l Inst. of Family & Life Advocates v. Becerra*, 138 S. Ct. 2361 (2018).

145. 475 U.S. 1, 8, 20–21 (1986).

146. *Id.* at 16.

147. *See, e.g., Cox v. New Hampshire*, 312 U.S. 569, 575–76 (1941).

148. *See, e.g., Ward v. Rock Against Racism*, 491 U.S. 781 (1989); *Clark v. Cmty. for Creative Non-Violence*, 468 U.S. 288 (1984); *Cox v. Louisiana*, 379 U.S. 536, 558 (1965).

149. In *Ward v. Rock Against Racism*, 491 U.S. at 791, the Court wrote that “*even* in a public forum the government may impose reasonable restrictions on the time, place, or manner of protected speech” (emphasis added), thus suggesting that a time, place, or manner regulation could also exist outside the public forum context. This is further supported by the Court’s analysis in *Pacific Gas & Electric Co.* In addition to analyzing California’s requirement that the private utility company cede space in its billing envelopes to other organizations as impermissible compelled speech (discussed *supra* Subpart III.A.1), the Court also considered the State’s argument that the requirement was a time, place, or manner regulation. The Court

Time, place, and manner regulations are almost always grouped together in First Amendment decisions but represent 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.

What it means to regulate manner, by contrast, is far less clear. The Supreme 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*.¹⁵⁰ 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.”¹⁵¹ In *Clark v. Community for Creative Non-Violence*, the Court held that barring protestors from sleeping on federal land in Washington, 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.”¹⁵² Ultimately, the concept of *manner* is less concrete and tidy than time or place, giving courts ample leeway to determine what constitutes a permissible manner regulation.

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. However, the government’s ability to restrict these aspects of speech is also limited. The government may regulate the time, place, or manner of speech only as long as the regulation (1) is content-neutral; (2) is narrowly

noted that the envelopes constituted a “private forum,” yet proceeded with a full time, place, and manner analysis. *Pac. Gas & Elec. Co.*, 475 U.S. at 9, 20–21 (ultimately striking down the regulation as impermissibly content-based). However, even if time, place, and manner analysis was limited to public forums, there are many online contexts in which the government does create public forums online where bots may speak. For example, many suspect bots generated the thousands of identical “citizen” comments to the Federal Communications Commission when the Commission proposed removing its net neutrality rules. See Paul Hitlin & Skye Toor, *Public Comments to the Federal Communications Commission About Net Neutrality Contain Many Inaccuracies and Duplicates*, PEW RES. CTR. (Nov. 29, 2017), <http://www.pewinternet.org/2017/11/29/public-comments-to-the-federal-communications-commission-about-net-neutrality-contain-many-inaccuracies-and-duplicates> [https://perma.cc/LL7N-YD9H]. The state may also host online discussion on websites such as www.whitehouse.gov.

150. 491 U.S. at 803.

151. *Id.* at 802.

152. *Clark*, 468 U.S. at 297.

tailored to serve a significant governmental interest; and (3) leaves open “ample alternative channels” to communicate the information.¹⁵³

The first of these prongs, content neutrality, is 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.”¹⁵⁴ With regard to time, place, and manner, the content neutrality requirement demands that the regulation must be “justified without reference to the content of the regulated speech.”¹⁵⁵ As clarified by the Supreme Court in *Reed v. Town of Gilbert*, content-based discrimination is easily conflated with viewpoint-based discrimination.¹⁵⁶ Any regulation aimed at a particular subject matter is not content-neutral, “even if it does not discriminate among viewpoints within that subject matter.”¹⁵⁷ By way of example, the Court noted that “a law banning the use of sound trucks for political speech—and only political speech—would be a content-based regulation, even if it imposed no limits on the political viewpoints that could be expressed.”¹⁵⁸ Thus, a valid time, place, or manner regulation must be specific only to particular noncontent elements of speech, such as volume, location, or format. It may not apply specifically to certain subject matters.

3. Means and Ends

Whether the government is seeking to compel speech or seeking to regulate the manner of speech, the key inquiry is the same: First Amendment doctrine requires narrow tailoring to a significant government interest. Achieving this narrow tailoring is neither straightforward nor simple. A significant government interest could be anything from reducing crime¹⁵⁹ to national security¹⁶⁰ to protecting citizens from unwelcome and excessive noise, even in traditional public forums such as parks.¹⁶¹ Government interventions also tend to be highly context-specific. While there may be plausible justifications for regulating a particular type of bot, it is difficult to imagine a justification that makes sense across multiple contexts.

153. *Id.* at 293 (citations omitted).

154. *Police Dep’t v. Mosley*, 408 U.S. 92, 95 (1972).

155. *Clark*, 468 U.S. at 293.

156. 135 S. Ct. 2218, 2230 (2015).

157. *Id.*

158. *Id.*

159. *City of Los Angeles v. Alameda Books, Inc.*, 535 U.S. 425, 430 (2002).

160. *Wayte v. United States*, 470 U.S. 598, 611 (1985).

161. *Ward v. Rock Against Racism*, 491 U.S. 781, 796 (1989).

It is far easier to conceive of narrowly tailored context-specific regulations, however. For example, commercial bots could be regulated under similar principles as other for-profit social media accounts. 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.¹⁶² 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.”¹⁶³ 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. This is bolstered by the fact that commercial speech generally receives a lower standard of protection than political, artistic, or other kinds of speech. Under the intermediate scrutiny *Central Hudson* test, a commercial speech regulation will survive so long as it directly advances a substantial government interest and is not more extensive than necessary.¹⁶⁴

In light of the widespread concern about foreign interference in the 2016 presidential election through social media bots, an automation disclosure requirement could be justified by a significant government interest with regard to political bots *in particular settings*. Even so, the line between “political” bots and private individuals expressing political views is a hazy one.¹⁶⁵ So too is the line between “election” related speech and general political speech—as the Supreme Court recently reaffirmed when it struck down Minnesota’s ban on “political” apparel in polling places.¹⁶⁶ Thus, while preserving free and fair elections can serve as 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.

162. *The FTC’s Endorsement Guidelines: What People Are Asking*, FTC: TIPS & ADVICE (Sept. 7, 2017), <https://www.ftc.gov/tips-advice/business-center/guidance/ftcs-endorsement-guides-what-people-are-asking> [<https://perma.cc/6MK3-3XF6>].

163. *Id.*

164. *Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm’n*, 447 U.S. 557, 561–66 (1980).

165. *See generally* *Fed. Elections Comm’n v. Wis. Right to Life, Inc.*, 551 U.S. 449 (2007) (deciding whether issue advertisements constituted express advocacy or its functional equivalent).

166. *Minn. Voters All. v. Mansky*, 138 S. Ct. 1876 (2018). In striking down the ban on “political” apparel in polling places, the Court emphasized the inherent vagueness in the notion of political apparel: “[P]erfect clarity and precise guidance have never been required even of regulations that restrict expressive activity.” *Id.* at 1891 (quoting *Ward*, 491 U.S. at 794). “But the State’s difficulties with its restriction go beyond close calls on borderline or fanciful cases. And that is a serious matter when the whole point of the exercise is to prohibit the expression of political views.” *Id.*

Even if the same rationales behind FTC and FEC regulations could justify regulation of commercial and political bots, respectively, the question remains whether any rationale could justify disclosure by all bots in all settings. At a minimum, any omnibus attempt to require all bots to identify themselves in all contexts would need to be a sort of Frankenstein's monster of government interests, where the government enumerates one or more significant rationale for each context it governs.¹⁶⁷ Moreover, when new uses for bots emerge, the legislature would need to revisit regulations and provide additional, responsive rationales. Ultimately, any omnibus bot disclosure bill seems doomed to be overinclusive, and thus likely to censor valuable speech. Moreover, the unique value of that speech may rely in part on the ambiguity of whether the speaker is automated or human.

Arts and entertainment furnish good examples. Science fiction is replete with figures whose basic humanity is in question, or where the dramatic tension depends on an ambiguity as to the nature and origins of a particular character.¹⁶⁸ Society's obsession with the Turing test—which asks whether a robot can fool a human into thinking it is not a robot—shows no signs of waning.¹⁶⁹ Online artists and storytellers increasingly incorporate bots and use them as a unique medium of expression.¹⁷⁰ These artists may express themselves through the intentional haziness of social media accounts that make us ask, “Is it a bot or not?” Forcing an artist to say whether a person is behind, for example, @MagicRealismBot, which may or may not be an automated source for magical realism storylines,¹⁷¹ interferes with his or her ability to tell a story in a particular way.

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

167. We owe this insight to David Skover.

168. See, e.g., NEAL STEPHENSON, *THE DIAMOND AGE: OR, A YOUNG LADY'S ILLUSTRATED PRIMER* (1995).

169. See, e.g., Gary Marcus, *What Comes After the Turing Test?*, *NEW YORKER: ANNALS OF TECH.* (June 9, 2014), <https://www.newyorker.com/tech/annals-of-technology/what-comes-after-the-turing-test> [<https://perma.cc/T7Y-AK6E>]; Lance Ulanoff, *Did Google Duplex Just Pass the Turing Test?*, *MEDIUM* (May 8, 2018), <https://medium.com/@LanceUlanoff/did-google-duplex-just-pass-the-turing-test-ffcfe6868b02> [<https://perma.cc/XV4A-V6PD>]; Yongdong Wang, *Your Next New Best Friend Might Be a Robot*, *NAUTILUS* (Feb. 4, 2016), <http://nautil.us/issue/33/attraction/your-next-new-best-friend-might-be-a-robot> [<https://perma.cc/6YEF-8WZN>].

170. Neyfakh, *supra* note 71.

171. See Ethan Chiel, *Magic Realism Bot Can Meet Some of Your Very Specific Literary Needs*, *SPLINTER* (Nov. 24, 2015), <https://splinternews.com/magic-realism-bot-can-meet-some-of-your-very-specific-l-1793853236> [<https://perma.cc/24LN-62X9>].

channels of communication available to convey her message.¹⁷² 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.¹⁷³ 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.

In addition to the prospect of bot disclosure laws being overinclusive—in the sense of sweeping in more speech than is needed to effectuate the government's goals—there is the prospect that bot disclosure laws will be underinclusive. Take, for example, the pending California legislation. The legislative history for SB 1001 specifically mentions Russian interference with the 2016 United States election in part through the use of automated accounts.¹⁷⁴ The bill itself mentions electoral interference.¹⁷⁵ Presumably, the idea is that California citizens will be better protected against manipulation if they understand the source of the message to be a bot.

However, as political bot researchers such as Philip Howard and Samuel Woolley show, many of the harms of bots occur at scale.¹⁷⁶ These include creating the perception that a politician is more popular than he is by padding his following; causing a particular idea or theory to “trend,” that is, to be highlighted as popular by a platform; and “hijacking” a hashtag that could further legitimate discussion or community building by flooding it with nonsense or vitriol. While a spot check may reveal that certain followers, amplifiers, or distractors are not real people, ultimately the harm is experienced in the aggregate. Unmasked bots can still perpetuate these scale-base harms, calling into question whether, for example, California's restriction of speech is accomplishing the government's ends. A large enough disconnect between the harm and the solution jeopardizes the means-ends requirement that any valid manner regulation must satisfy.

172. *Clark v. Cmty. for Creative Non-Violence*, 468 U.S. 288, 293 (1984).

173. Billboards, television advertisements, and other forms of online advertising, just to name a few.

174. S.B. 1001 § 17941(a), Cal. Leg., 2017–2018 Sess. (Cal. 2018) (as amended by Cal. Senate, Mar. 14, 2018).

175. S.B. 1001 § 17941(a), Cal. Leg., 2017–2018 Sess. (Cal. 2018) (prohibiting the use of unidentified bots to “influence a vote in an election”).

176. Philip N. Howard, Samuel Woolley & Ryan Calo, *Algorithms, Bots, and Political Communication in the U.S. 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration*, 15 J. INFO. TECH. & POL. 81, 83–91 (2008).

The most obvious response to the means-ends fit problem is to proceed in a piecemeal fashion, regulating only certain types of bots—commercial or electioneering bots, for example—and justifying each intervention separately. As we discuss in the final Part of this Article, we believe this to be the wiser course for regulators: A government body interested in regulating bot speech should articulate the specific harm or harms in a particular context that justify imposing limits there.¹⁷⁷ It is important to note, however, that such an approach forecloses justifying a given law as a mere time, manner, and place regulation. This is so because time, manner, and place regulations must be content-neutral, that is, they may not privilege certain categories of speech over others.¹⁷⁸ Rather, the proper free speech analysis for a law, such as California's, that singles out categories of speech such as commercial or electoral is that of coerced speech.

B. Unmasking Through Enforcement

Imagine for the purpose of argument that the government furnished a variety of legitimate justifications for bot disclosure in narrow contexts where the state has the authority to regulate. For example, the state might successfully invoke consumer protection interests to enact a law that requires self-identification when using an automated agent for marketing, or might invoke the sanctity of the election process in prohibiting automated attempts to influence an election without disclosure.¹⁷⁹ Such laws would likely be facially constitutional so long as they did not discriminate on the basis of viewpoint.

However, absent careful thought regarding the enforcement mechanism, actual application of such statutes could nevertheless veer into unconstitutional territory. To date, no bot disclosure proposal has described the specifics of *how* such a requirement would be enforced. Among our chief concerns is the prospect that enforcement of a generic bot disclosure law would interfere with the right to speak anonymously. Unmasking the status of a bot as a bot differs from penalizing an anonymous speaker. But unmasking anonymous speakers may be inevitable absent the establishment of a mechanism by which a human can prove her status as a natural person without divulging her identity.

Arguably distinct from other forms of mandatory disclosure, which typically require information about the subject of speech rather than the speaker, the right to anonymity is well-established in American jurisprudence.

177. See *infra* Part III.

178. See *supra* Subpart III.A.2.

179. Again, drawing the line between election-related and other political speech is notoriously difficult. We are assuming the problem is tractable for purposes of this Subpart.

Discussions of the right inevitably begin with the Federalist papers.¹⁸⁰ However, the Supreme Court did not directly speak on the issue of anonymity protection until the mid-1900s. In *NAACP v. Alabama ex rel. Patterson*,¹⁸¹ 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 the First Amendment right to freedom of association rather than freedom of speech.¹⁸² Noting that disclosure of NAACP membership lists had exposed members to threats, public hostility, and economic harm in the past, the Court reasoned that compelling disclosure would adversely affect the group's ability to advocate effectively.¹⁸³ The Court found freedom of association to be inextricably linked to the right to "privacy in one's associations."¹⁸⁴

The right to anonymity was first formally linked to free speech protection two years later in *Talley v. California*,¹⁸⁵ 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 the 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.¹⁸⁶ 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."¹⁸⁷

The Court reaffirmed a staunch protection of the right to speak anonymously thirty-five years later in *McIntyre v. Ohio Elections Commission*,¹⁸⁸ 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

180. *McIntyre v. Ohio Elections Comm'n*, 514 U.S. 334, 343 n.6 (1995).

181. 357 U.S. 449 (1958).

182. *See id.* at 460–62.

183. *See id.* at 462–63.

184. *Id.* at 462.

185. 362 U.S. 60 (1960).

186. *Id.* at 64.

187. *Id.*

188. 514 U.S. 334 (1995).

in the advocacy of political causes”¹⁸⁹ that serves as “a shield from the tyranny of the majority.”¹⁹⁰

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 generally:¹⁹¹ “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.”¹⁹²

Litigation furnishes another relevant context. Although the existence of the right to bring suit under a pseudonym is well-established, the Supreme Court has not yet delineated the contours of when pseudonymous litigation is permissible.¹⁹³ A plaintiff must first obtain permission from the court in order to proceed without revealing her real name,¹⁹⁴ and although the Supreme Court has permitted the practice in numerous cases¹⁹⁵ (perhaps most famously in *Roe v. Wade*),¹⁹⁶ it has never expressly addressed when a plaintiff may litigate pseudonymously. The Circuit Courts permit the practice according to varying standards, many of which include consideration of factors such as whether the plaintiff would be forced to reveal sensitive personal information if not permitted

189. *Id.* at 343.

190. *Id.* at 357.

191. Margot Kaminski, *Real Masks and Real Name Policies: Applying Anti-Mask Case Law to Anonymous Online Speech*, 23 *FORDHAM INTELL. PROP. MEDIA & ENT. L.J.* 815, 834–35 (2013). Anti-mask laws prohibit the wearing of masks in public. *Id.* at 848.

192. *Id.* Kaminski has written extensively on the right to anonymity in the context of anti-mask laws, arguing 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.” *Id.* at 850. 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. *Id.* at 850–74. Some decisions find that they established an independent right to anonymity under the First Amendment, while others characterize the right as a component of the right to free expression. *Id.* Other courts declined to apply the First Amendment altogether in the context of anti-mask laws, most notably the Second Circuit in its *Kerik* decision. *Id.* at 865 (discussing *Church of the Am. Knights of the Ku Klux Klan v. Kerik*, 356 F.3d 197 (2d Cir. 2004)).

193. Jayne S. Ressler, *#WorstPlaintiffEver: Popular Public Shaming and Pseudonymous Plaintiffs*, 84 *TENN. L. REV.* 779, 810–11 (2017).

194. A. Michael Froomkin, *Anonymity and the Law in the United States*, in *LESSONS FROM THE IDENTITY TRAIL: ANONYMITY, PRIVACY, AND IDENTITY IN A NETWORKED SOCIETY* 441, 458 (Ian Kerr et al. eds., Oxford Univ. Press 2009).

195. *Id.*

196. 410 U.S. 113 (1973).

to proceed pseudonymously,¹⁹⁷ whether the plaintiff seeks to challenge government activity,¹⁹⁸ the risk of prejudice or unfairness to the defendant,¹⁹⁹ the potential for retaliatory harm to the plaintiff,²⁰⁰ and whether the plaintiff is a child.²⁰¹

The protection of anonymity established in these and other 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.²⁰² The government bears the burden of proving that a challenged law is narrowly tailored in furtherance of a compelling state interest.²⁰³ Accordingly, there are two particular areas where courts have recognized valid government interests that justify unmasking: electoral speech and at certain points in litigation.

Electoral speech. In *McIntyre*, the Court noted that a narrow identification requirement might be justified on the basis of certain government interests.²⁰⁴ 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.²⁰⁵ 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.²⁰⁶ 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.²⁰⁷ That same year, the Court upheld the Bipartisan Campaign Reform Act’s disclaimer and disclosure requirements for campaign advertisements in *Citizens United v.*

197. See, e.g., *James v. Jacobson*, 6 F.3d 233, 238–39 (4th Cir. 1993).

198. See, e.g., *id.*; *Doe v. Porter*, 370 F.3d 558, 560–61 (6th Cir. 2004).

199. See, e.g., *Doe v. City of Chicago*, 360 F.3d 667, 669 (7th Cir. 2004).

200. See, e.g., *Does I Thru XXIII v. Advanced Textile Corp.*, 214 F.3d 1058, 1068 (9th Cir. 2000).

201. See, e.g., *James*, 6 F.3d at 238–39; *Porter*, 370 F.3d at 560–61.

202. *McIntyre v. Ohio Elections Comm’n*, 514 U.S. 334, 334–35 (1995).

203. See *id.*; *First Nat’l Bank of Bos. v. Bellotti*, 435 U.S. 765, 786 (1978).

204. See *McIntyre*, 514 U.S. at 353.

205. See *John Doe No. 1 v. Reed*, 561 U.S. 186 (2010).

206. *Id.* at 195.

207. See *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. See *Kaminski*, *supra* note 191, at 839.

FEC.²⁰⁸ In doing so, the Court emphasized the public’s “informational interest” and the importance of “making informed choices in the political marketplace.”²⁰⁹

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 result in *McIntyre*,²¹⁰ and Justice Thomas, who believed that neither law was justified under strict scrutiny.²¹¹

Litigation. The standards for unmasking anonymous defendants vary both by jurisdiction and by the nature of the claims at issue. When a plaintiff sues for defamation, the most widely employed standards for unmasking an unknown defendant (usually an online speaker) are variations on the New Jersey *Dendrite*²¹² standard: First, plaintiffs must notify anonymous speakers in order to provide them with a reasonable opportunity to contest a potential unmasking; second, they must identify precisely which statements are allegedly defamatory; third, they must produce prima facie evidence supporting every element of their claim; and finally, the court must weigh the risk of unmasking the defendant against the harm to the plaintiff on a case-by-case basis.²¹³ Less stringent anonymity protections apply to third parties in litigation such as witnesses and subpoena recipients. For example, in *Branzburg v. Hayes*,²¹⁴ the Supreme Court held that requiring journalists to reveal their sources when subpoenaed by a Grand Jury did not violate the First Amendment.²¹⁵

208. 558 U.S. 310 (2010).

209. *Id.* at 369, 367.

210. *Reed*, 561 U.S. at 219–20 (Scalia, J., concurring).

211. *Id.* at 239 (Thomas, J., dissenting).

212. *Dendrite Int’l, Inc. v. Doe*, 775 A.2d 756 (N.J. Super. Ct. App. Div. 2001).

213. See Jason A. Martin & Anthony L. Fargo, *Anonymity as a Legal Right: Where and Why It Matters*, 16 N.C. J.L. & TECH. 311, 342 (2015).

214. 408 U.S. 665 (1972).

215. The Ninth Circuit’s recent ruling in *United States v. Glassdoor*, 875 F.3d 1179, 1191 (9th Cir. 2017), extended this reasoning to hold 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, see, e.g., Lisa A. Hayes, *Anonymous Speech Online Dealt a Blow in U.S. v. Glassdoor Opinion*, CTR. FOR DEMOCRACY & TECH. (Nov. 8, 2017) <https://cdt.org/blog/anonymous-speech-online-dealt-a-blow-in-us-v-glassdoor-opinion> [<https://perma.cc/TW32-7LU4>], 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.

In short, the First Amendment protects the right to speak and even litigate anonymously. The exceptions are narrow and vary by context, such that officials and courts may require self-disclosure if there is a pressing enough need. The trouble with generic bot disclosure laws—and each version of actual bills we have seen introduced—is that they fail to provide for a process for accused humans to prove they are not actually bots. This issue will only become more acute as bots become more adept at mimicking people.

Consider a hypothetical Twitter account that frequently tweets eccentric critiques of local politicians. One such politician, up for reelection, objects to a pattern he sees whereby the account replies to (“@”) comments of his followers with variations on the same criticism of his record. This politician suspects that the account is actually a bot and refers the matter to a local prosecutor, citing a bot disclosure requirement around election-related speech. If the prosecutor agrees, how will she go about enforcing the statute?

Calls for bot disclosure bills are silent on this point. But it is of critical importance to preserving the right to speak anonymously, for the only way to investigate, let alone prosecute, the politically critical account is to pursue the person or people behind it. If the account is a bot and its operator stands in violation of the statute, then the platform (here, Twitter) will have to disclose the operator’s identity for purposes of facing charges. But if there is a real person behind the account, that person will also have to come forward and prove they are, in fact, human. As there is no mechanism in place for verifying that a person is a person without revealing *which* person, we must assume virtually every instance of enforcement will involve unmasking.

There are a number of means to domesticate the unmasking problem. For example, Twitter and other platforms could devise a system by which to certify the human status of users without using their real names. Or Twitter could automatically detect and label bots (thereby removing user discretion). As we discuss in the final Subpart of this Part, Twitter might simply ban automated accounts to avoid the hassle. A trusted third party could accomplish the same by creating a process whereby a person could attest that they are the speaker behind an accused account.²¹⁶

The problem with these and similar approaches, apart from the fact that they do not exist as of this writing, is that they are not officially sanctioned. Neither the prosecutor nor the accuser is obligated to take the platform or a third party’s word for the humanness of an accused account, let alone that of the accused herself. The prosecutor could, in theory, continue to pursue the claim up

216. We owe this insight to commentary by Michael Froomkin.

to the point of unmasking the person behind it. It is hard to see how this unmasking would be justified under the narrow exceptions to the right to anonymous speech. But it is equally difficult to see how bot disclosure laws would ever be enforced if the person behind the account could not be identified.

Today this problem might be mitigated by the ease with which the human behind an account can prove they are human merely through their ability to engage in dynamic conversation. Although they date back decades, chat bots are still in their infancy; telling a person from a bot is not hard due to bots' limited discursive capabilities. The problem becomes more and more acute, however, as bots become increasingly capable of dissembling convincingly. Already "robocallers," or phone-based bots, deny that they are automated.²¹⁷ Future bots could have specific protocols for facing down accusers, including reaching out for help from a person who is on call for this purpose.²¹⁸

C. Automating Censorship

A final consideration sounds less in First Amendment doctrine than in general principles of free speech in cyberspace. A general requirement that bots disclose themselves in a sense delineates "bot speech" as a separate category. Sometimes we do draw such lines around speech. For example, commercial speech—defined as speech that "propose[s] a commercial transaction"—exists apart from other forms of expression and may be amenable to greater control.²¹⁹ As we discussed in Part II, certain communication falls outside of speech protections altogether.²²⁰ But there seems to be no basis by which to lump all bot speech together as a category. To do so, especially at this early stage of its development, would prejudice bot speech and perhaps enable or even encourage the censorship of that speech.

Successfully requiring bots to self-identify could lead to attempts to limit bot speakers in new ways. As discussed above, no official path exists by which to contest accusations of automation without exposing the real person behind the speech at issue. Platforms may respond to this messiness by prohibiting bot speech altogether so as to avoid getting between the user and the accuser. This

217. Dvorsky, *supra* note 30.

218. Indeed, while we do not address the question in this Article, the prospect of human-bot hybrids seems to present special difficulties in drafting bot disclosure rules.

219. *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc.*, 425 U.S. 748, 762 (1976).

220. *See supra* Part II.

would dramatically reduce the volume and variety of bot speech on the Internet, all without official censorship.²²¹

Even absent choices by platforms, if all bot speakers must identify themselves as fitting in a unitary category—nonhuman speech—then all speech online becomes searchable and sortable. The very sortability of bot speech could lead to further demands, if not by government, then by individuals and groups, that platforms create a means by which to identify and block all speech in the bot category. This action would not be *state* action, of course. But because the disclosure requirement that makes such blocking possible would be a government mandate, subsequent platform blocking would be enabled by state action. Because these platforms operate internationally, in foreign jurisdictions wherein the First Amendment does not apply, a U.S.-based bot disclosure requirement could empower non-U.S. officials to block bot speech entirely as a category—for example, if bots are being used to criticize the government—without having to do the work of identifying bots.

In cyberlaw terms, broadly applied bot disclosure laws will alter the architecture, and with it the regulability, of an emerging form of digital speech.²²² While this form of regulation may not register to courts as violating the First Amendment, bot disclosure laws nevertheless implicate the forms and possibilities of human communication and hence, free expression itself.

IV. REGULATING BOT SPEECH: A CASE FOR CAUTION

To summarize the argument so far: Bots have many forms and purposes. In some cases, bots have contributed to a variety of information-based harms. These harms are visible enough to occasion calls for and attempts at regulation, specifically in the form of a requirement that bots disclose themselves as nonhuman. Although such a requirement seems unproblematic from a First Amendment perspective, a deeper analysis reveals a variety of free speech concerns. These include the difficulty of narrowly tailoring regulation to the government interest at stake, as well as the prospect that questions around anonymous speech and other related issues will arise in the enforcement of bot disclosure laws.

We do not deny that bots generate significant harms in the current environment. Nor do we argue against the regulation of bots, including

221. See Assaf Hamdani, *Who's Liable for Cyberwrongs?*, 87 CORNELL L. REV. 901, 916–18 (2002) (discussing how strict regulation of online content in general encourages excessive self-censorship by internet services providers and third-party platforms).

222. See LAWRENCE LESSIG, *CODE: VERSION 2.0*, at 32 (2006).

potentially through mandatory disclosure. Rather, we believe that current approaches—which contemplate blanket requirements of bot self-disclosure without reference to context—tend to offend principles of free speech. And we worry more broadly about the free speech consequences of attempting to channel a new medium of speech at its inception. Accordingly, in this final Part, we urge caution and self-restraint in regulating this widely varied and still emerging form of communication.

Eight years ago, in *Brown v. Entertainment Merchants Association*, the Supreme Court confronted the question of whether the government could treat violent, interactive videogames as unprotected speech.²²³ The Court, without denying (or affirming) the prospect that violent videogames could be harmful, answered this question with a resounding no. “[W]hatever the challenges of applying the Constitution to ever-advancing technology,” the Court observed, the requirements of the First Amendment “‘do not vary’ when a new and different medium for communication appears.”²²⁴ Indeed, the government is not free to add new categories of unprotected speech merely because the legislature “concludes certain speech is too harmful to be tolerated.”²²⁵

Like videogames, bots can be a vehicle for speech that society finds problematic—speech that, for instance, foments strife, deeply offends, or attempts to manipulate. But this capacity for harm does not confer a license upon the state to shunt bots into a category of speech deserving of lesser protection. And the harms of bots should be considered not only alongside the potential benefits, but with the understanding that bots represent—like videogames or even the Internet—an emerging form of speech. The scale issues that make bots problematic also permit an engagement at scale. The very ambiguity between human and machine that makes bots feel dangerous is also a source of novel forms of expression, research, and critique.

Ultimately, the communicative potential of bots urges, if not inaction, then at least caution. We therefore conclude this Article with a series of recommendations to policymakers around how best to approach bot speech in these early days.

First, to the extent feasible, governments should begin by updating and leveraging existing law to address harms caused by bots. If the concern is, for instance, bots being used to harass individuals or groups, to exploit the elderly,

223. *Brown v. Entm’t Merchs. Ass’n*, 564 U.S. 786 (2011); see also *United States v. Stevens*, 559 U.S. 460, 472 (2010) (“Our decisions . . . cannot be taken as establishing a freewheeling authority to declare new categories of speech outside the scope of the First Amendment.”).

224. *Brown*, 564 U.S. at 790 (quoting *Joseph Burstyn, Inc. v. Wilson*, 343 U.S. 495, 503 (1952)).

225. *Id.* at 791.

children, or other vulnerable individuals, or to interfere with an election, then each of these problems might be addressable without enacting new, untested laws with the capacity to infringe on speech.

Second, and relatedly, governments should regulate bot speech, if at all, through individual restrictions aimed at (1) particular categories of bots, (2) within specific contexts, and (3) supported by the specific harms the government hopes to mitigate. Thus, for example, if the concern is commercial bots leaving scathing reviews on Yelp about competitors, or political bots created by Political Action Committees coordinating with a candidate or his campaign, governments should address these concerns specifically rather than require all bots to behave a particular way. As discussed above, however, any such intervention should be assessed under a coerced speech standard, because mere regulations of time, manner, and place must be content-neutral.

Legislatures might consider starting with regulation of commercial bots, given the lower standard of scrutiny generally applied to commercial speech and the well-established importance of consumer protection and other related interests. Governments should in all instances interrogate whether their proposed solution sweeps in harmless speech and, conversely, whether it actually addresses the harmful activity at issue. And they should acknowledge that, for at least some categories of bot speech, the requirement to self-identify itself operates as a restriction on expression.

Third, governments should anticipate and address inevitable issues around enforcement. With respect to a generic bot disclosure law, there will be many instances in which an official or citizen suspects noncompliance. It is the government's obligation to create a viable, constitutional path by which individuals and groups can prove they are compliant without having to give up speech rights such as the right to speak anonymously. Such a path could include a means by which the platform or another third party can verify the human nature of a given account, or provide penalties for attempting to silence an individual by falsely reporting her to be a bot. Despite the many precedents devoted to proceeding anonymously in public or in court, no bot disclosure scheme proposed to date gives any guidance on this question.²²⁶

And fourth, governments should acknowledge the downstream effects of officially differentiating between bot speech and other forms of online communication. In theory, bot disclosure laws merely offer signals to individuals

226. As discussed above, California's law prohibits misleading consumers or citizens as to the human nature of the bot but does not specify any procedure for verifying that a bot is human without unmasking the person behind it. No such procedure appears in the federal bill around political bots nor in any call for bot regulation we have seen in the press.

as they navigate a complex information ecosystem. But in practice, those signals may come to serve as the scaffolding for private or, outside of jurisdictions with a robust free speech tradition, public censorship of bots as a category of speech. Perhaps limitations of bot speech will arise from market forces as platforms such as Twitter and Facebook respond to user demand for greater civility and transparency. The question is under what conditions governments can, or at any rate should, alter the character of speech to make it more susceptible to various forms of suppression.

We believe our analysis of the speech concerns around bot disclosure yields certain lessons. Other analyses and experiences may yield many more. None of these principles, alone or in combination, guarantees the constitutionality of bot disclosure or other laws that affect bot speech. But together they point the way toward more thoughtful interventions geared to address the automated speech phenomenon without arbitrarily limiting a new and still unfolding communication medium.

CONCLUSION

In their aforementioned book on rights in robot speech, Collins and Skover observe a certain pattern around new technologies of communication.²²⁷ Beginning as far back as the written word, novel forms of information and media have been met with skepticism. This skepticism has in turn led to censorship: of the written word, of the printed page, of the Internet. Over time, however, Collins and Skover observe that useful technologies of communication have found a way to flourish.²²⁸ Free speech has expanded to meet these technologies largely on their own terms.

Time will tell whether the many and varied bots of today and tomorrow meet this threshold of utility. They have already displayed the capacity for significant mischief, and some measure of wonder. This Article has shown that a popular response to the harms of bots may look innocuous on the surface but, upon deeper analysis, implicates core free speech concerns. Bots represent a new form of communication—whether in their capacity to surprise, their ability to produce speech at scale, or the way in which some bots test our intuitions about the boundary between person and machine. This novelty is frightening and even harmful. Any response must nevertheless be measured and respect age-old principles of free expression.

227. COLLINS & SKOVER, *supra* note 118.

228. *Id.* at 31.