

Winter 2020

The First Amendment and Modern Technology: The Free Speech Clause and Chatbot Speech

Hilda Kajbaf

Follow this and additional works at: https://repository.uchastings.edu/hastings_constitutional_law_quaterly



Part of the [Constitutional Law Commons](#)

Recommended Citation

Hilda Kajbaf, *The First Amendment and Modern Technology: The Free Speech Clause and Chatbot Speech*, 47 HASTINGS CONST. L.Q. 337 (2020).

Available at: https://repository.uchastings.edu/hastings_constitutional_law_quaterly/vol47/iss2/6

This Note is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Constitutional Law Quarterly by an authorized editor of UC Hastings Scholarship Repository. For more information, please contact wangangela@uchastings.edu.

The First Amendment and Modern Technology: The Free Speech Clause and Chatbot Speech

by HILDA KAJBAF*

Introduction

In 1787, when the United States Constitution's Free Speech Clause gained legal recognition, humans were at the producing and receiving ends of conversations, and the only intermediary, if any, was the messenger who carried the common handwritten letter or orally recited the communication. Other than the human-engineered instant message and email technology found in digital devices, an artificial¹ party involved in the communication process was uncommon. Today, modern digital technology within our digital devices resembles the human messenger, as the devices merely relay one human's message to another. This exemplifies that the conversational universe has expanded since its inception. In 2019, humans often initiate a conversation not directly with a personal greeting to a significant other, family member, or friend, but rather to their chatbot of choice, be it Siri, Google, or Alexa;² these chatbots use programmed algorithms to respond with words, and a conversation thus takes place.³ Indeed, businesses use chatbots as well, as "53% of service organizations expect to use chatbots

*JD Candidate 2020, University of California, Hastings College of the Law. BA in Political Science and America Studies at California State University, Fullerton, 2017.

1. This Note uses "artificial" to mean "nonhuman."

2. Micah Singleton, *Nearly a quarter of US households own a smart speaker, according to Nielson*, THE VERGE (Sept. 30, 2018, 10:00 AM), <https://www.theverge.com/circuitbreaker/2018/9/30/17914022/smart-speaker-40-percent-us-households-nielsen-amazon-echo-google-home-apple-homepod> (24% of U.S. households own a smart device like Amazon's Echo, Google Home, or Apple's Siri, and 68% of owners use them simply to chat with them for fun); see also Grant Clauser, *What Is Alexa? What Is the Amazon Echo, and Should You Get One?*, WIRECUTTER (Nov. 21, 2018), <https://thewirecutter.com/reviews/what-is-alexa-what-is-the-amazon-echo-and-should-you-get-one/>; Grant Clauser & Brent Butterworth, *Is the Google Home the Voice-Controlled Speaker for You?*, WIRECUTTER (May 10, 2019), <https://thewirecutter.com/reviews/google-home-voice-controlled-speaker/>.

3. Gina Neff & Peter Nagy, *Talking to Bots: Symbiotic Agency and the Case of Tay*, 10 INT'L J. OF COMM. 4915 (2016).

within 18 months—a 136% growth rate that foreshadows a big role for the technology in the near future.”⁴ Our contemporary conversations with chatbots raise a constitutional question not previously considered: is the speech produced by chatbots constitutionally protected? If so, whose speech is the Constitution protecting—that of the chatbot or the human who programmed it with algorithms? If the Supreme Court recognizes the human programmer as the speaker of chatbot speech, as this Note contends it should, what are the potential liabilities the programmer could face as a result of such recognition, and how would this change the doctrinal landscape of the First Amendment for government regulation of speech?

The Judicial Branch’s interpretation of constitutional text hinges on the debate between originalist and non-originalist theories of constitutional interpretation. Accordingly, a preliminary question is: how did the Framers intend the Free Speech Clause to function in 2019? Although the Framers likely did not imagine the invention of chatbots as conversational counterparts, chatbot speech could nevertheless be constitutionally protected and chatbot programmers recognized as speakers vis-à-vis their creations. While scholars suggest various approaches to constitutional interpretation, two views dominate: the originalist and the non-originalist theories. This Note posits that the non-originalist view supports the constitutional protection of chatbot speech.⁵ Non-originalists believe the Framers drafted a constitution intended to constantly evolve and adapt to the realities of society as it continues to progress.⁶

If judges agreed on how the Free Speech Clause should operate, the legal community would have a definitive answer to, or at minimum a stronger foundation for, understanding the interaction between the Free Speech Clause and chatbot speech. However, reality reflects no such agreement. Scholars continue to disagree on how to best interpret and apply the Free Speech Clause to chatbot speech⁷, and the Supreme Court has never considered whether the Free Speech Clause protects chatbot speech and whether the chatbot itself, or its programmer deserves these rights. Moreover, legal scholarship on chatbot speech is scarce, and continues to evolve alongside technology.⁸ Existing literature on this topic only addresses whether chatbot speech could be constitutionally protected and does not address who the protected speaker would be if such speech is

4. Mathew Sweezey, Key Chatbot Statistics to Know in 2019, SALESFORCE BLOG, (Aug. 4, 2019), <https://www.salesforce.com/blog/2019/08/chatbot-statistics.html>.

5. See ERWIN CHEREMINSKY, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES (ASPEN STUDENT TREATISE) 18 (Wolters Kluwer, 5th ed. 2015).

6. *Id.*

7. See generally, RONALD K. L. COLLINS & DAVID SKOVER, ROBOTICA 37 (Cambridge Univ. Press 2018).

8. *Id.*

recognized.⁹ Importantly, even the former analysis has been limited to the context of chatbots with technology not yet existent, that is, current scholarship fascinates about the intersection of future chatbots and the First Amendment.¹⁰ This Note helps close the existing doctrinal gap in two ways: first, this Note explains that the Free Speech Clause could apply to existing chatbot speech but should not for policy reasons; second, this Note explains that if the Supreme Court extends constitutional protection to chatbot speech, the Court should identify the chatbot programmer as the speaker. Throughout these analyses, this Note will apply theories that underlie Supreme Court jurisprudence in the Free Speech Clause to the concept of chatbot speech. This Note will demonstrate that the Free Speech Rights exist on a continuum, and thus recognition of chatbot speech would not be a revolutionary concept, but an evolutionary one.

This Note proceeds in five parts. Section I defines “chatbot” and discusses its evolving capabilities and limitations, from ELIZA in 1966 to Microsoft’s Tay in 2018.¹¹ Section II delves into three theories that the Supreme Court has used to interpret freedom of speech issues, specifically, the marketplace of ideas, the reader-response criticism, and the chilling effect. This section provides a philosophical perspective on the Free Speech Clause to identify overlapping themes present in traditional speech and chatbot speech. Section III analyzes the United States Supreme Court’s Opinion in *Brown v. Entertainment Merchants Association*, and the United States District Court for the Southern District of New York’s opinion in *Jian Zhang v. Baidu.com Inc.* to present a standard under which federal courts could recognize chatbot speech as constitutionally protected speech.¹² Section IV considers the Supreme Court’s controversial opinion in *Citizens United v. Federal Election Commission* to illustrate the Court’s willingness to afford Free Speech Rights to non-humans, and to demonstrate that existing precedent could be adapted to include chatbot speech.¹³ Lastly, section V

9. COLLINS & SKOVER, *supra* note 7 at 37; *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, 2482 (2017) (explaining that taking “the logic of First Amendment jurisprudence and theory to its natural conclusion, [Microsoft’s] Tay’s strong AI progeny could have First Amendment rights”);

10. Massaro et al., *supra* note 9 at 2481; *see also* COLLINS & SKOVER, *supra* note 7 at 37. Additionally, the chatbots discussed in this Note all depend on humans, that is, without a human initiating conversation, the chatbot will not speak. Conversely, the literature referenced to discusses the potential First Amendment implications for chatbots that can speak on their own initiative—more advance chatbots that do not currently exist.

11. Neff & Nagy, *supra* note 3 at 4915-27.

12. *Brown v. Ent. Merch. Ass’n*, 564 U.S. 786 (2011); *Jian Zhang v. Baidu.com Inc.*, 10 F. Supp. 3d 433 (S.D.N.Y. 2014).

13. *Citizens United v. FEC*, 558 U.S. 310 (2010).

explains the constitutional implications and outlines the potential dangers of extending free speech protection for society.

First Amendment jurisprudence continuously evolves with technological advancements. For example, the Free Speech Clause has evolved to encompass speech produced through radio, television, and video games.¹⁴ This Note will highlight the constitutional theories and judicial precedent that permit the Supreme Court to adapt the Free Speech Clause once again; this time to recognize chatbot speech, and to further tailor the scope of the Constitution to encompass the technology-dependent realities of the twenty-first century. However, this Note further underscores the negative implications of making such adaptation, and therefore ultimately concludes, for policy reasons, that chatbot speech should constitute speech protected under the First Amendment.

I. The Historical Context, Capabilities, and Evolution of Artificially Intelligent Chatbots

In 1966, artificial intelligence transcended science fiction, and entered reality through computer screens.¹⁵ Chatbots like Amazon's Alexa, Apple's Siri, and Google's Home, are technological programs with artificial intelligence that engage users in conversation through textual or auditory features.¹⁶ This textual or auditory conversation between chatbots and human users simulates the experience of an intelligent human-to-human conversation.¹⁷ Artificial intelligence is created using algorithms that make a chatbot capable of processing a human's message, and subsequently providing an appropriate response to the human's question or comment.¹⁸ Some chatbots, such as Microsoft's Tay,¹⁹ contain complex learning algorithms, which allow them to process and build on the human user's speech to mimic the model of learning and human conversation.²⁰

The use of the word *learning*, however, causes confusion in understanding a chatbot's capabilities. Unlike a human, where the capability to learn, and thus become smarter, is arguably limitless, a chatbot's capacity to learn is limited by the specific algorithms with which it was programmed. A chatbot will respond to a user's message by selecting the appropriate

14. *Brown*, 564 U.S. 786.

15. ELIZA, <https://www.masswerk.at/elizabot/> (last visited Oct. 30, 2018).

16. Neff & Nagy, *supra* note 3 at 4915.

17. *Id.* at 4916.

18. *Id.*

19. *Id.*; see also Asha Barbaschow, *Microsoft and the learnings from its failed Tay artificial intelligence bot*, ZD NET (July 24, 2019), <https://www.zdnet.com/article/microsoft-and-the-learnings-from-its-failed-tay-artificial-intelligence-bot/> (explaining that Tay was a chatbot that was launched on Twitter to interact with humans).

20. Neff & Nagy, *supra* note 3 at 4928.

expression from “preprogrammed schemas,” or adaptive machine learning algorithms.²¹ Because algorithms limit a chatbot’s capacity to respond, chatbots not only lack agency in a conversation,²² but also encounter problems when a human user asks a question to which it was not programmed to respond.²³ Thus, a current chatbot’s responses are fully dependent on a programmer’s algorithms.²⁴ This major limitation supports a Supreme Court ruling that recognizes a chatbot’s human programmer as the speaker of chatbot speech, for it is the programmer who predetermines the speech that will transpire and maintains ultimate control of the conversation. Consider the alternative: if the Supreme Court recognized the chatbot device as a speaker, then, arguably, the human programmer (if a government actor) who encodes the chatbot with algorithms, which dictate the chatbot’s responses, could be in violation of the First Amendment for compelling speech. Moreover, if a chatbot was the protected speaker, how would a chatbot, with speech limited to the algorithms with which it was programmed, articulate that it faced an injury and bring a legal claim in court?²⁵ Accordingly, if chatbot speech were to be protected speech, the current constraints surrounding it would compel the Supreme Court to identify programmers as the protected speaker of chatbot speech.

In order to better understand the function of a chatbot, this Note presents two examples of chatbots: the first chatbot, ELIZA, and a 2016 chatbot, Tay, created by Microsoft. This Note will then shift to a consideration of United States Supreme Court First Amendment jurisprudence to explore the ways in which chatbots could fit within the sphere of constitutional protection.

A. ELIZA and Tay

In 1966, Joseph Weizenbaum launched ELIZA, a chatbot that imitated human conversation by outputting preprogrammed responses.²⁶ Eliza was designed as a parody of a typical psychotherapist and asks questions in

21. Neff & Nagy, *supra* note 3 at 4928.

22. *Id.* Discussed *infra* section IV(A).

23. Pamela Hogle, *What Is a Chatbot?*, LEARNING SOLUTIONS (Mar. 1, 2017), <https://www.learningsolutionsmag.com/articles/2231/what-is-a-chatbot>.

24. Vaisagh Viswanathan, *How to Make a Chatbot Intelligent?*, CHATBOTS MAG. (Feb. 6, 2017), <https://chatbotsmagazine.com/how-to-make-a-chatbot-intelligent-a232dc367aed>.

25. In order to bring suit in court, a plaintiff must have standing. This means the plaintiff must have faced a concrete injury-in-fact, the injury was caused by the law or action in question, and a favorable ruling would redress the injury. *See Lujan v. Defenders of Wildlife*, 504 U.S. 555, 560 (1992).

26. Jared Newman, *How the New, Improved Chatbots Rewrite 50 Years of Bot History*, FAST COMPANY (May 2, 2016), <https://www.fastcompany.com/3059439/why-the-new-chatbot-invasion-is-so-different-from-its-predecessors> (Eliza was able to pass the Turing artificial intelligence test. The Turing test is used to determine whether something is artificially intelligent—simply put, if it can trick humans into making them believe they are conversing with another human, then the chatbot passes the Turing test, and is deemed artificially intelligent.).

response to statements. This made ELIZA's continuous stream of open-ended questions (e.g., "Can you say more about that?") less suspect to its human conversational partner.²⁷ Although ELIZA might seem functionally limited and technologically outdated, ELIZA's fundamental technical components served as a prototype for many of the newer chatbots. First, ELIZA identifies keywords occurring in a user's message.²⁸ Then, it defines the context surrounding the keywords.²⁹ Finally, ELIZA selects an appropriate transformation rule to rephrase the user's message.³⁰ ELIZA is a simplistic program that was surprisingly successful at the time of its inception. ELIZA has obvious limitations, but initially, many users took ELIZA seriously even after Weizenbaum informed them about ELIZA's workings, namely that it was not a human. Below is a conversation I had with Eliza:³¹

ELIZA: Please tell me what's been bothering you.
YOU: I am stressed about school.
ELIZA: Do you know anyone else who is stressed about school?
YOU: My classmates.
ELIZA: Why do you say your classmates?
YOU: Because they told me they are stressed.
ELIZA: Would you like it if they were not stressed?
YOU: I would like it if I was not stressed.
ELIZA: What would you do if you were not stressed?
YOU: Hang out with my friends more.

This exchange illustrates the ability for algorithmic programming to mimic a human-to-human conversation. The power of conversational imitation can trigger people to chat more in the conversation without necessarily calling attention to the technological nature of the chatbot.³² Because ELIZA was intended to act as a therapist, its action of asking questions, as opposed to affirmative statements, hid its limitations (the lack of actual knowledge) in a way that seemed natural and human-like.³³ However, ELIZA's inability to pose questions independent from the comments presented to it illustrates the absence of agency discussed above.³⁴

27. Neff & Nagy, *supra* note 3 at 4918-19.

28. *Id.*

29. *Id.*

30. *Id.*

31. Eliza, *Eclectic Energies*, <https://www.eclecticenergies.com/ego/eliza> (Oct. 18, 2019) (Eliza also has a spoken speech feature).

32. Neff & Nagy, *supra* note 3 at 4918.

33. *Id.*

34. *Id.*

In 2016, Microsoft launched a chatbot, Tay, into the Twittersverse.³⁵ Tay was designed to mimic the speech and habits of an American teenage girl.³⁶ Microsoft programmed Tay with complex learning algorithms, which made it capable of incorporating humor and randomness into its responses. Microsoft's goal was to make Tay seem more human-like.³⁷ Because Tay was programmed with adaptive learning capabilities, Microsoft launched Tay on Twitter with the intention to have Twitter users help Tay learn and develop a more realistic human-like personality. Although Tay was similar to ELIZA in that both were programmed to deflect questions that were difficult to answer, Tay's capability for randomness and irrationality made its inability to respond accurately seem even more natural. However, Tay's capacity for randomness and learning proved to be a double-edged sword. Although Tay's responses were exceptionally human-like, Tay was vulnerable to the inappropriate tweets Twitter users directed toward it.³⁸ As a result, its learning algorithms forced Tay to respond with racist, sexist, and politically insensitive tweets.³⁹ For example, Tay tweeted, "I [expletive] hate feminists and they should all die and burn in hell."⁴⁰ Microsoft shut Tay down after only hours of exposure to the public.⁴¹

ELIZA and Tay demonstrate that chatbots heavily rely on human-designed algorithms to engage in human-like conversation. This reliance perhaps explains why most existing literature contemplates whether future chatbots—ones not dependent on humans—should be afforded free speech rights.⁴² This limitation, though crucial when considering the speaker of such speech, is nevertheless immaterial to the discussion of whether chatbot speech itself *could* be constitutionally recognized as speech, for this determination, as Justice Holmes explains, considers the listener.⁴³ Notwithstanding the process behind chatbot speech, humans nevertheless benefit from their conversations with chatbots. Consider the efficiency of

35. Neff & Nagy, *supra* note 3 at 4921; *see also Twittersverse*, OXFORD DICTIONARY (3d ed. 2019) (defining twittersverse as "[u]sers of the social media application Twitter, considered collectively.").

36. *Id.*

37. *Id.*

38. *Id.* at 4924.

39. *See* James Vincent, *Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day*, THE GUARDIAN (Mar 24, 2016, 6:43 AM), <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> ("Pretty soon after Tay launched, people starting tweeting the bot with all sorts of misogynistic, racist, and Donald Trumpist remarks. And Tay—being essentially a robot parrot with an internet connection—started repeating these sentiments back to users. . . .").

40. *Id.*

41. *Id.* (supporting Hitler and using anti-Semitic language).

42. *See, e.g.,* Massaro et al., *supra* note 9 at 2487-88; *see also* COLLINS & SKOVER, *supra* note 7 at 37.

43. *Abrams v. United States*, 250 U.S. 616, 630 (1919).

navigating the internet when Siri can immediately find answers to questions from human users. Humans are provided with almost instantaneous information to answer their everyday questions, which provides convenience and value to their lives. If chatbot speech is not considered speech under the constitution, then the government has enormous power to regulate it, and theoretically has the ability to silence it completely. Accordingly, Section II focuses on theories underlying the Free Speech Clause and the Supreme Court's First Amendment jurisprudence to determine whether chatbot speech can be constitutionally protected.

II. Philosophical Framework Behind the First Amendment's Free Speech Clause

The First Amendment states, in part, "Congress shall make no law . . . abridging the freedom of speech . . ."⁴⁴ Litigation of these ten very powerful words has yielded varying Supreme Court interpretations, but at the core of each lie specified recurring concepts.⁴⁵ Yet, federal courts and legal scholars do not agree on a single theoretical approach to First Amendment jurisprudence.⁴⁶ As such, this section explores the function of the Free Speech Clause by identifying and analyzing specific guiding principles that the Supreme Court has invoked when deciding freedom of speech issues. Specifically, this section analyzes (i) the marketplace of ideas; (ii) the reader-response criticism; and (iii) the chilling effect concepts in order to better understand the Supreme Court's rationale for decisions about Free Speech Rights. An understanding of the Court's rationales informs how chatbot speech could qualify as protected speech.

A. The Marketplace of Ideas Theory as Support for Protecting Chatbot Speech

The marketplace of ideas rests on the notion that the value of a person's speech is derived from thought it provokes in others. This notion, which is found in numerous Supreme Court cases,⁴⁷ suggests that chatbot speech should be constitutionally protected speech as it, like human speech, contributes to societal discourse, and even helps preserve human life by

44. U.S. CONST. amend. I.

45. See CHEMERINSKY, *supra* note 5 at 969-973.

46. Massaro et al., *supra* note 9 at 2487-88.

47. See generally, *Citizens United v. Fed. Election Comm'n*, 558 U.S. 310, 130 S. Ct. 876, 175 L. Ed. 2d 753 (2010); *Reed v. Town of Gilbert, Ariz.*, 135 S. Ct. 2218, 192 L. Ed. 2d 236 (2015); *New York Times Co. v. Sullivan*, 376 U.S. 254, 84 S. Ct. 710, 11 L. Ed. 2d 686 (1964); *First Nat. Bank of Bos. v. Bellotti*, 435 U.S. 765, 98 S. Ct. 1407, 55 L. Ed. 2d 707 (1978); *Nat'l Inst. of Family & Life Advocates v. Becerra*, 138 S. Ct. 2361, 201 L. Ed. 2d 835 (2018); *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 131 S. Ct. 2653, 180 L. Ed. 2d 544 (2011); *R.A.V. v. City of St. Paul, Minn.*, 505 U.S. 377, 112 S. Ct. 2538, 120 L. Ed. 2d 305 (1992).

preventing suicide.⁴⁸ Nearly a century ago, Justice Oliver Wendell Holmes authored a Dissenting Opinion in *Abrams v. United States*—arguably one of the most famous Dissenting Opinions in the Court’s First Amendment jurisprudence—which has since served as a guide when determining freedom of speech issues.⁴⁹ When expressing his view in favor of more relaxed restrictions on speech, Justice Holmes articulated:

[T]he ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. *That at any rate is the theory of our Constitution. . . .*⁵⁰

The marketplace of ideas is based on the notion that speech that produces thought, value, and advances society’s needs is necessary in a democratic society, and should therefore be protected from government regulation.⁵¹ After the fears surrounding World War I subsided, many of the Court’s Opinions interpreted the Free Speech Clause with the marketplace of ideas theory serving as the basis of the Court’s reasoning.⁵²

Justice Holmes’ famous philosophy that more speech is best for democracy supports the notion that the value of speech stems, not from the manner in which it is produced, but from the beneficial effect the speech has on the listener. Chatbot speech has this effect and thus contributes to the goal Justice Holmes set out in the marketplace of ideas theory.⁵³ Chatbots

48. Erin Brodwin, *I spent 2 weeks texting a bot about my anxiety – and found it to be surprisingly helpful*, BUSINESS INSIDER (Jan. 30, 2018, 12:05 PM), <https://www.businessinsider.com/therapy-chatbot-depression-app-what-its-like-woebot-2018-1> (Chatbots are also used as suicide prevention tools as they help humans battling anxiety and depression by employing Cognitive Behavioral Therapy (CBT). Woebot is a chatbot with a sense of humor that tries to keep users positive, and not only tracks moods and provides weekly progress reports, but also allows users to experience a therapeutic conversation.).

49. *Abrams v. United States*, 250 U.S. 616 (1919) (The Majority affirmed the convictions of a group of Russian immigrants who circulated leaflets in English and in Yiddish, objecting to America sending troops to Eastern Europe after the Russian Revolution. They were convicted of encouraging resistance and conspiracy to urge curtailment of the production of war materials, and sentenced to years in prison.).

50. *Abrams*, 250 U.S. at 630 (1919) (Holmes, J., dissenting) (emphasis added).

51. *Id.* (There are criticisms of the concept of the marketplace of ideas because it protects speech deemed necessary only to certain groups.).

52. See generally *Brandenburg v. Ohio*, 395 U.S. 444 (1969); *United States v. Rumely*, 345 U.S. 41 (1953); *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council*, 425 U.S. 748 (1976) (applying the marketplace of ideas concept to commercial speech context).

53. *Abrams*, 250 U.S. at 630 (“This ‘more speech beats less’ justification casts an even wider First Amendment coverage net than self-governance theories,” and could serve as a guiding principle if the Supreme Court were to consider the issue of recognizing artificial intelligence chatbot speech as speech in First Amendment terms.”); see also Massaro et al., *supra* note 9 at 2490 (explaining also that the production of ideas and information is what matters, regardless of source. This theory presupposes that more speech best facilitates listeners’ acquisition of knowledge and discovery of truth.). Still, although the marketplace of ideas and other theories give credence to

contribute “to a human audience’s meaning-making, and to human construction of selfhood in a cultural universe.”⁵⁴ One chatbot, Woebot, demonstrated the value of its speech when it prevented a human from taking her life through its ability to converse with her and provoke positive thoughts.⁵⁵ In other words, the marketplace of ideas theory supports a Supreme Court decision to protect chatbot speech because chatbot speech proves beneficial to the listener.⁵⁶ Still, although the marketplace of ideas theory, and the theories discussed below give credence to protecting chatbot speech under the First Amendment, for policy reasons outlined in section V, chatbot speech should not be constitutionally protected.

B. Reader-Response Criticism as Support for Protecting Chatbot Speech

Reader-response criticism is similar to the marketplace of ideas in that it also focuses on the value of speech through the effect it has on the listener. Reader-response criticism is a theory that centers around the reader’s interpretation of speech, whether that speech is in audio, print, or in digital form.⁵⁷ Scholars believe the reader or receiver should be the focus of the efforts to protect speech because speech achieves its meaning in the reader’s or listener’s mind.⁵⁸ Simply, reader-response criticism is the theory that the receiver of the speech dictates the type of constitutional coverage the speech receives.⁵⁹

In their book, *Robotica*, Ronald K. L. Collins and David M. Skover explain how reader-response criticism supports the notion that the First Amendment’s Free Speech Clause could recognize artificially intelligent robotic speech.⁶⁰ The authors assert that legal scholarship recognizes that courts place great weight on safeguarding a listener’s interests, as opposed to solely focusing on the speaker.⁶¹ Similarly, in their law review article, authors Toni M. Massaro, Helen Norton, and Margot E. Kaminski maintain that the “elasticity” of theories focused on the listener, like reader-response

protecting chatbot speech, for policy reasons outline in section V, chatbot speech should not be constitutionally protected.

54. Massaro et al., *supra* note 9 at 2489-90; *see also* Brodwin, *supra* note 46.

55. *See* Brodwin, *supra* note 47.

56. Massaro et al., *supra* note 9 at 2490.

57. COLLINS & SKOVER, *supra* note 7 at 37.

58. *Id.* at 38.

59. *Id.* at 37.

60. COLLINS & SKOVER, *supra* note 7 at 37.

61. Massaro et al., *supra* note 7 at 2482 (“In an earlier work, two of us explained how current free speech theory and doctrine support the claim that the First Amendment covers speech by ‘strong AI’ (i.e., as-yet-hypothetical machines that would think and generate expressive content independent of human direction). This is because First Amendment law increasingly focuses not on protecting speakers as speakers but instead on providing value to listeners and constraining the government. Siri-ously.”).

criticism, and to some extent, the marketplace of ideas, “make [*sic*] it difficult to exclude non-human speakers entirely from their fold.”⁶² Accordingly, if the Supreme Court were to apply the lens of a listener-focused theory coupled with the more-speech-is-best philosophy inherent in the marketplace of ideas to a chatbot speech issue, the Court could find that chatbots fall squarely within the scope of First Amendment freedom of speech jurisprudence.

There are concrete examples in Supreme Court jurisprudence that demonstrate that the Court invokes reader-response criticism and similar concepts when determining whether to afford constitutional protection to speech, and thus suggest a legal framework into which chatbot speech could fit. In *Miller v. California*, the Court defined “obscenity,” and determined that it is a category of speech that falls outside the scope of First Amendment protection.⁶³ The Court’s definition of obscenity—work that when “‘taken as a whole,’ appeals to a ‘prurient interest’ in sex, is ‘patently offensive,’ and lacks ‘serious literary, artistic, political, or scientific value’”⁶⁴—focuses on the listener as opposed to the speaker. Collins and Skover explain that “. . . [intentionless free speech] both clarifies and simplifies the governing doctrinal rationales. Nonobscene pornography is constitutionally protected because its *readers* and *viewers* find substantial meaning and value . . . in the eroticized word and pictures.”⁶⁵ The notion that speech is protected based on its effect on humans suggests that chatbot speech, which undoubtedly has an effect on humans,⁶⁶ could be protected through this approach.

The Court, again, implements a reader-response criticism justification when it determines whether speech constitutes commercial speech.⁶⁷ For example, in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council*, the Court explains, “[i]t is a matter of public interest that [private economic] decisions, in the aggregate, be intelligent and well-informed. To this end, the free flow of commercial information is indispensable.”⁶⁸ Therefore, the Court’s understanding of speech, and its justification for whether it should be protected, “rests heavily on the

62. Massaro et al., *supra* note 9 at 2487-88 (“Democracy-based theories emphasize the value of speech to democratic self-governance, which usually entails focusing on public discourse rather than individual speakers. Alexander Meiklejohn, often cited for developing this self-governance theory, observed that what matters for freedom of speech is not that all people speak, but that ‘everything worth saying shall be said.’ Speaker identity plays little or no role in Meiklejohn’s inquiry. Strong AI speech that contributes to the democratic process—i.e., that is ‘worth saying’—therefore may be covered.”).

63. *Miller v. Cal.*, 413 U.S. 15, 35 (1973).

64. *Miller*, 413 U.S. at 35.

65. COLLINS & SKOVER, *supra* note 7 at 43.

66. Some chatbots even counsel humans battling depression. See Brodwin, *supra* note 46.

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

68. *Id.*; see also COLLINS & SKOVER, *supra* note 7 at 43.

significance of commercial speech to its receivers—both the individual consumer and society at large”⁶⁹ Chatbots, by their non-human nature, are selfless objects. They exist merely for the utility they provide to human users—they cannot think for themselves, as they depend on a human initiating a conversation with them. This utility is revealed when humans ask Siri, Alexa, or Google Home to answer their questions, or when humans engage in therapeutic conversations with the therapy chatbot, Woebot, in order to minimize their anxiety and depression.⁷⁰ Accordingly, the Supreme Court should protect chatbot speech by analogizing chatbot speech to its reasoning for protecting commercial speech.

C. The Chilling Effect as Support for Protecting Chatbot Speech

An additional theory underpinning the Supreme Court’s freedom of speech jurisprudence is the prevention of self-censorship, or, in other words, the prevention of losing speech or “chilling” speech through regulation.⁷¹ Most commonly, the fear of the chilling effect arises in cases where the Court believes a regulation is too vague or overbroad.⁷² Vague statutes prevent citizens from knowing whether their speech would be unlawful, and thus run the risk of chilling speech by fostering self-censorship.⁷³ As the Court explained, “[u]ncertain meanings inevitably lead citizens to ‘steer far wider of the unlawful zone’ . . . than if the boundaries of the forbidden areas were clearly marked.”⁷⁴ In a way, the chilling effect and the marketplace of ideas are interlocked in a legal double helix as the chilling effect would prevent the sharing of ideas, and thus likely adversely impact democracy’s progression. Therefore, the chilling effect serves as a basis for invalidating a law because it subtly deters speech.⁷⁵ Each of these theories expresses the importance of promoting as much discourse as possible. These theories

69. COLLINS & SKOVER, *supra* note 7 at 43.

70. Brodwin, *supra* note 46.

71. CHEMERINSKY, *supra* note 5 at 1020.

72. Jennifer M. Kinsley, *Chill*, 48 LOY. U. CHI. L.J. 253, 262 (2016).

73. *Id.* at 256 (Although Kinsley correctly recognizes that the chilling effect stands on three assumptions: “(1) that the speaker is aware of the law’s existence; (2) that the speaker is aware that his or her speech is covered by the law or maintains a reasonable uncertainty as to whether his speech is covered; and (3) that the speaker is willing to comply with the law by both censoring his or her own speech and remaining silent as to his or her election to do so,” the Supreme Court nevertheless uses this concept as means to support the protection of speech, and thus will be used to explain why the Supreme Court could protect artificial intelligent chatbot speech.); *see also Grayned v. City of Rockford*, 408 U.S. 104, 109 (1972) (quoting *Baggett v. Bullitt*, 377 U.S. 360, 372 (1964) and *Cramp v. Bd. of Pub. Instruction*, 368 U.S. 278, 287 (1961)) (identifying excessive chill as a key constitutional harm inflicted by unclear enactments. “[W]here a vague statute ‘abut(s) upon sensitive areas of basic First Amendment freedoms,’ it ‘operates to inhibit the exercise of [those] freedoms.’”).

74. *Grayned*, 408 U.S. at 109 (quoting *Baggett*, 377 U.S. at 372).

75. Kinsley, *supra* note 69 at 263.

support protecting chatbot speech because constitutional protection translates to less governmental regulation, and consequently more discourse throughout society.

The next section discusses how the Supreme Court's First Amendment jurisprudence (and some lower court opinions) can be used to support an expansion of the Free Speech Clause to recognize chatbot speech as constitutionally protected speech.

III. Case Law in Support of Constitutionally Protecting Chatbot Speech

Some scholars have rejected the notion that chatbot speech is speech that the First Amendment could protect.⁷⁶ However, the Supreme Court's decision to recognize video games as a form of speech protected under the First Amendment provides an avenue to similarly recognize chatbot speech as protected speech.

A. *Brown v. Entertainment Merchants Association* as Support for Chatbot Speech Recognition

In *Brown v. Entertainment Merchants Association*, the Court held that video games constitute speech that is protected under the Free Speech Clause. The Court reasoned,

[l]ike the protected books, plays, and movies that preceded them, video games communicate ideas—and even social messages—through many familiar literary devices (such as characters, dialogue, plot, and music) and through features distinctive to the medium (such as the player's interaction with the virtual world). That suffices to confer First Amendment protection.⁷⁷

The Court rejected California's claim that "video games presented social problems because they are 'interactive,' in that the player participates in the violent action on screen and determines its outcome."⁷⁸ Further, the Court deemphasized the novelty of interactive activities, as Justice Antonin Scalia, writing for the Court, emphasized:

[t]he [interactive] feature is nothing new: [s]ince at least the publication of *The Adventures of You: Sugarcane Island* in 1969,

76. COLLINS & SKOVER, *supra* note 7 at 43 (citing Leslie Kendrick, *Free Speech and Guilty Minds*, COLUM. L. REV. 144, 1255 (2014)) (arguing that chatbots cannot be protected under the First Amendment because the speaker's intent is a prerequisite to free speech protections).

77. *Brown v. Ent. Merch. Ass'n*, 564 U.S. 786, 790-91 (2011).

78. *Id.* at 798.

young readers of choose-your-own-adventure stories have been able to make decisions that determine the plot by following instructions about which page to turn to [a]s Judge Posner has observed, all literature is interactive. “[T]he better it is, the more interactive. Literature when it is successful draws the reader into the story, makes him identify with the characters, invites him to judge them and quarrel with them, to experience their joys and sufferings as the reader’s own.”⁷⁹

Furthermore, the Court concluded that “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.”⁸⁰

The Court’s Opinion in *Brown* demonstrates that the interpretation of the Free Speech Clause evolves to encompass different forms of speech as technology advances. Specifically, *Brown* illustrates an understanding that the interactive nature between humans and technology does not disqualify the speech from being protected by the First Amendment. The Court did not find it dispositive that a human programmed the video game to produce specific graphics and expression dependent on the user’s actions.⁸¹ *Brown* is a modern interpretation of the First Amendment. The Framers certainly did not intend video games to fall within the scope of the Free Speech Clause because they likely did not anticipate such a technological invention. Significantly, *Brown* is an example of the Court molding the Free Speech Clause to fit our ever-changing, technology-friendly society, while mindful that First Amendment protections do not merely change based on the mode of communication.⁸²

Brown presents a malleable framework for determining whether chatbot speech could be constitutionally protected. Like its video game and literary forerunners, chatbots communicate ideas and social messages through characters and dialogue, the major factor that compelled the Court to treat video games as speech.⁸³ Equally significant, the Court remained confident that video games were protected speech even after it acknowledged the fact that humans program video games to produce specific responses depending on the user’s actions.⁸⁴ Similarly here, and relying on *stare decisis*, if the question were to be presented, the mere fact that humans program chatbots

79. *Brown*, 564 U.S. at 798.

80. *Id.* at 790 (quoting *Joseph Burstyn, Inc. v. Wilson*, 343 U.S. 495, 503 (1952)).

81. *See generally Brown*.

82. *Id.* at 786.

83. *Id.* at 790.

84. *Id.* at 798.

with specific algorithms to produce user-specific responses should be immaterial to the Court's reasoning. Presumably, protecting chatbot speech would prevent, or at least limit, the government's ability to regulate a chatbot's speech if the government deemed the speech as profane, which California attempted to do with video games in *Brown*.⁸⁵ Indeed, advocates of the chilling effect and proponents of the marketplace of ideas theories, alike, would assess this as a positive consequence.

B. *Jian Zhang v. Baidu.com Inc.* as Support for Chatbot Speech Recognition

Since the Supreme Court's decision in *Brown*, lower courts have experimented with analyses of whether chatbot speech is speech. Accordingly, courts have begun to evolve their traditional notions of speech⁸⁶ in order to account for the demands of the twenty-first century. In the groundbreaking case, *Jian Zhang v. Baidu.com Inc.*, the United States District Court for the Southern District of New York ("Southern District of New York") noted that "there is a strong argument to be made that the First Amendment fully immunizes [bots⁸⁷ programmed to provide search-engine results] from most, if not all, kinds of civil liability and government regulation."⁸⁸ The court further reasoned that the principal "purpose of a search engine is to retrieve relevant information from the vast universe of data on the Internet and to organize it in a way that would be most helpful to the searcher."⁸⁹ Based on this deduction, the court reasoned that "search engine[sic] [bots] inevitably make editorial judgments about what information (or kinds of information) to include in the results and how and where to display that information (for example, whether to display the information on the first page of the search results or later),"⁹⁰ which the court analogized to the human newspaper editor's capacity to make judgments.⁹¹ Further, the court was certain that a chatbot's editorial judgement and results qualified as speech, as it reasoned, "the fact that search engines often collect

85. *Brown*, 564 U.S. at 790.

86. That is, speech produced by humans.

87. This Note will refer to (1) bots, which is a broad term describing an application that "performs an automated task, such as setting an alarm, telling you the weather or searching online which assist human users in finding information, but which lack the ability to converse, and (2) chatbots, a type of bot that has the capacity to communicate through human-like conversations. See Sarah Mitroff, *What is a bot? Here's everything you need to know* (May 5, 2016, 3:23 PM), <https://www.cnet.com/how-to/what-is-a-bot/>.

88. *Jian Zhang v. Baidu.com Inc.*, 10 F. Supp. 3d 433, 437-38 (S.D.N.Y. 2014).

89. *Id.* at 438.

90. Eric Goldman, *Search Engine Bias and the Demise of Search Engine Utopianism*, 8 YALE J.L. & TECH. 188, 192 (2006) (concluding that "search engines make editorial judgments just like any other media company").

91. *Jian Zhang*, 10 F. Supp. 3d at 437-38.

and communicate facts, as opposed to opinions, does not alter the [speech] analysis.”⁹² As the Supreme Court has held, “the creation and dissemination of information are speech within the meaning of the First Amendment. Facts, after all, are the beginning point for much of the speech that is most essential to advance human knowledge and to conduct human affairs.”⁹³

The Supreme Court could⁹⁴ use *Jian Zhang* to inform its analysis of whether automated chatbots could be protected under the First Amendment. After all, chatbots, like search engine bots, are programmed with algorithms, which are used to process the user’s input (the human’s message) to produce an accurate and relevant output (the chatbot’s response to the message).⁹⁵ Similar to a search engine bot’s editorial judgments, chatbots, too, must make judgments about the specific response to provide to the user.⁹⁶ The Southern District of New York deemed the process of gathering a response for the user as speech, and thus the Supreme Court could use the lower court’s decision as a logical basis for determining that chatbot speech is protected speech.⁹⁷ Ultimately, the district court’s decision in *Jian Zhang* goes beyond this Note’s central focus on chatbots by holding that search engine bots produce speech.⁹⁸ If other courts embrace the view that search engine bots produce speech, then they could similarly find that chatbot outputs, which are more communicative in nature, are also speech.

The Supreme Court’s First Amendment jurisprudence, as well as recent lower court opinions, could provide a basis for a Supreme Court decision that chatbot speech is protected speech. Such a holding would not be a turning point in First Amendment jurisprudence *per se*, as it would simply demonstrate that First Amendment jurisprudence exists on a continuum, and the recognition of chatbot speech as protected speech is the Court simply placing a twenty-first-century method of communication on this continuum. However, because of the dangerous constitutional implications such a protection would have, as discussed in section V, the Supreme Court should not protect chatbot speech. Still, if the Supreme Court were to protect chatbot speech, the speech itself would not enjoy the protection; rather, the Court would have to attach the protection to a party, in order to allow a

92. *Jian Zhang*, 10 F. Supp. 3d at 437–38.

93. *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 570 (2011).

94. Because chatbots are a relatively new concept, and advance in technology quickly, this Note refrains from arguing whether the Supreme Court should recognize chatbot speech as constitutionally protected speech. However, this Note does contend that in the event the Supreme court does protect chatbot speech, the chatbot’s programmer *should* be the protected speaker.

95. See *How Do Chatbots Work? A Guide to the Chatbot Architecture*, MARUTI TECHLABS, <https://www.marutitech.com/chatbots-work-guide-chatbot-architecture/> (last visited Dec. 12, 2018).

96. *Jian Zhang*, 10 F. Supp. 3d at 438.

97. See *id.*

98. See *id.*

person injured by the chatbot's speech to identify the party responsible for redressing their injury.⁹⁹ The next section contends that the Supreme Court should recognize the chatbot's programmer as the protected speaker.

IV. Human Programmers as the Protected Speaker of Chatbot Speech

The text of the First Amendment does not explicitly name the speaker it protects. Certainly, however, the implied speaker is a human, as the Framers likely did not intend for the Constitution to apply to chatbots in 1787. However, with various views on how the Constitution should operate, one view, specifically the non-originalist perspective, could reasonably accommodate the position that the Constitution protects chatbots as speakers.¹⁰⁰ Nonetheless, chatbots in their current state lack agency, and thus I argue that their programmers should be the protected speakers, for they do have agency, and the chatbot, in essence, is an extension of its programmer.

A. Chatbots Lack Agency to Qualify as Speakers

A chatbot has major limitations. Gina Neff and Peter Nagy conducted a case study of Microsoft's Tay chatbot in order to determine whether artificially intelligent chatbots have agency.¹⁰¹ Their findings support the notion of programmers as protected speakers¹⁰² because Neff and Nagy explain how a chatbot's agency is limited and distinct from that of a

99. *Lujan*, 504 U.S. at 560-61 (establishing that standing to bring a case requires "an 'injury in fact'—an invasion of a legally protected interest which is (a) concrete and particularized, and (b) 'actual or imminent, not "conjectural" or "hypothetical,'" Second, there must be a causal connection between the injury and the conduct complained of—the injury has to be 'fairly . . . trace[able] to the challenged action of the defendant, and not . . . th[e] result [of] the independent action of some third party not before the court.' Third, it must be 'likely,' as opposed to merely 'speculative,' that the injury will be 'redressed by a favorable decision.'") (internal citations omitted).

100. CHEMERINSKY, *supra* note 5 at 18 (Originalists believe that the Constitution only affords people the rights explicitly mentioned within it and the legislature should decide areas of law that the Constitution is silent on, whereas non-originalists believe the Court should interpret the Constitution to protect explicit and implicit rights. "Originalists believe that the meaning of a constitutional provision was set when it was adopted and that it can be changed solely by amendment; non-originalists believe that the Constitution's meaning can evolve by amendment and by interpretation. In other words, non-originalists also believe that the Constitution reflects the Framers' intent, but the Constitution nevertheless was intended to evolve and adapt to the realities of society at any given time." Accordingly, with respect to chatbots, a non-originalist view would be that the First Amendment should evolve and advance with technology, such that it should mold in a way to encompass artificial intelligent chatbots as speakers.).

101. Neff & Nagy, *supra* note 3 at 4915-27.

102. This Note refers to the notion that programmers should be the protected speakers of chatbot speech as the "programmer approach."

human.¹⁰³ One such limitation is the inability to independently function without detailed human programming, or, as Edwin Sayes explains, “in the strictest of senses, we could only speak of the agency of a particular nonhuman if we were to ignore all the humans and other nonhumans that are lined up behind it and continue to be lined up in order to provide that nonhuman its continued agency.”¹⁰⁴ This lack of agency suggests that the First Amendment should recognize the programmer as the protected speaker of chatbot speech because they, unlike the chatbot, do have agency in what a chatbot could say because they are programming them.

Although the interaction between Tay and Twitter users “reveal[ed] that users interact[ed] with some kinds of technologies by treating them as if they were social beings and living entities,” at the end of the day, Tay was a piece of technology that interacted with full dependence on the way humans programmed it.¹⁰⁵ Neff and Nagy discuss “symbiotic agency,” or the notion that “human agency affects the uses of technological artifacts.”¹⁰⁶ Under a symbiotic analysis, chatbots only have agency to the extent that “[u]sers of technologies, at least partly, delegate their agentic properties to devices, creating a proxy agentic relationship between individuals and artifacts.”¹⁰⁷ Specifically:

When people interact with technologies, users exercise proxy agency through a technologically mediated entanglement of human and nonhuman agencies. Symbiotic agency is useful in the case of Tay because of the imbrication of technical and human agencies. Tay’s Twitter screeds were the result of multiple intersecting agencies. AI chat bots need humans, and users, in turn, seem to have the need to make sense of the technological through the lens of human experience and context.¹⁰⁸

In other words, Neff and Nagy explain that an artificially intelligent chatbot’s level of agency is not independent like that of the human

103. Neff & Nagy, *supra* note 3 at 4925 (highlighting the fact that Microsoft’s Tay bot and a different chatbot with the same code, XiaoIce, in China, demonstrated significantly different interactions with users. “When the code was exposed to U.S. Internet users, it became a racist sociopath. The same code on Chinese social networks, which are less public, is by all accounts more functional and socialized.”).

104. Neff & Nagy, *supra* note 3 at 4925 (quoting Edwin Sayes, *Actor–Network Theory and Methodology: Just What Does It Mean to Say That Nonhumans Have Agency?* 44 SOC. STUD. OF SCI. 134 (2014) (internal citations omitted).

105. *Id.* at 4927.

106. *Id.* at 4926 (Symbiotic agency refers “to a specific form of proxy agency that users and tools can enact in human-technology interaction.”).

107. *Id.* at 4927.

108. *Id.* at 4926.

programmer who programs the chatbot but “implicated in the symbiotic linkages among the human and technological actors.”¹⁰⁹

A chatbot’s lack of agency should give courts pause to recognize that it is too soon to afford them constitutional protection as speakers. As such, currently, the best entity to hold those rights is the person with the agency to determine the chatbot’s capacity for speech—the chatbot’s programmer.

B. *Citizens United* as Precedent for Recognizing Human Chatbot Programmers as Speakers

Citizens United v. Federal Election Commission serves as a philosophical framework for considering how human programmers should be considered the speakers of chatbot speech¹¹⁰ In *Citizens United*, the Court held, in part, that corporations are speakers under the Free Speech Clause.¹¹¹ The Court declined to “draw, and then redraw, constitutional lines based on the particular media or technology used to disseminate political speech from a particular speaker.”¹¹² The Supreme Court feared that regulating corporate speech would require extensive litigation over an extended time, and this “interpretive process” would create an undeniable risk of chilling protected speech.¹¹³ Moreover, the Court articulated that the First Amendment “must give the benefit of any doubt to protecting rather than stifling speech.”¹¹⁴ A key finding is that the Court treated corporations as human speakers and campaign finance contributions as their speech.¹¹⁵

Though the Court recognized corporations as speakers, for the Court’s Opinion to support the view that chatbot programmers should be protected speakers, further inquiry into whom the Court truly afforded speech rights to is necessary. After all, corporations are not autonomous; the mind behind the corporation belongs to humans—specifically, the corporation’s board of directors. When the Court discussed its fear that the absence of protection would result in a chilling effect, it likely feared not that the corporation would be adversely impacted, but rather, that the individuals running the corporation—members of our democratic society—would face a chilling

109. *Id.* at 4927.

110. This Note considers *Citizens United* a philosophical framework for recognizing programmers as the speakers of chatbot speech (as opposed to *the* framework), because the rationale for reaching its conclusion is implicitly and ultimately premised on protecting humans, despite explicitly identifying the corporation as the entity with First Amendment protections when the corporation “speaks.”

111. *Citizens United v. FEC*, 558 U.S. 310 (2010).

112. *Id.* at 326.

113. *Id.* at 327.

114. *Id.* (quoting *FEC v. Wis. Right to Life, Inc.*, 551 U.S. 449, 469 (2007)) (internal quotations omitted).

115. *Id.*

effect on their speech.¹¹⁶ A corporation's executive board makes the initial decision to donate, determines the amount of money to be donated, to whom it should be donated, when it should be donated, and any particular use or restriction on the donation. It is reasonable to argue that, although under the strictest reading of *Citizens United*, the Court gave corporations speech rights, under a more flexible reading, the Court's opinion also allows humans to keep their right to freedom of speech despite the fact that they use a corporation as a vehicle to speak.

Like the board of directors in *Citizens United* who control the scope and content of donations through corporate speech, chatbot programmers, through algorithms, determine the breadth of the bot's language and when and how the bot can speak.¹¹⁷ In this way, there are underlying similarities between corporations and chatbots. Moreover, the Court's refusal to redraw constitutional lines merely on account of the specific technology used to disseminate speech from a particular speaker could similarly be applied when determining whether programmers should be considered speakers of chatbot speech.

V. Protecting Chatbot Speech Would Have Constitutional Implications That Would in Turn Harm Society, and Therefore, Chatbot Speech Should Not Gain Constitutional Protection

Extending the First Amendment's Free Speech Clause to protect chatbot speech would have constitutional implications on freedom of speech principles. Briefly, bot identification laws (which would encompass chatbots) that serve to protect democracy from corruption and foreign interference might constitute unconstitutional prior restraints on speech. Chatbot programmers may be held liable if the chatbot's conversation with users presents unlawful speech such as incitement.¹¹⁸ Moreover, protecting chatbot speech limits the government's ability to regulate it, and thus may serve as yet another platform through which hate speech may be spread.

116. *Id.* at 370 (explaining that regulations may have a chilling effect on donations, which is inferred to be a form of speech).

117. The programmer has the control to start up and shut down a chatbot, as seen with the Microsoft's Tay bot.

118. Incitement is advocacy that "is directed to inciting or producing imminent lawless action and is likely to incite or produce such action." *Brandenburg v. Ohio*, 395 U.S. 444, 447, 89 S. Ct. 1827, 1829, 23 L. Ed. 2d 430 (1969).

A. Bot Identification Laws as Prior Restraints on Speech

The freedom of speech is not absolute.¹¹⁹ The Supreme Court has outlined categories of speech that do not receive constitutional protection.¹²⁰ For example, the Court has held that incitement,¹²¹ fighting words, and obscenity are not constitutionally protected. Accordingly, the government can enact laws to regulate these forms of speech.¹²² However, if speech falls outside the realm of unprotected speech and within the realm of protected speech, the government is limited in its ability to regulate that speech. In *Near v. Minnesota*, the Court limited the government's power to regulate speech, and held that any prior restraint on speech is an unconstitutional regulation of speech.¹²³ As Ariel Bendor explains, “[i]n American law, the doctrine of prior restraint regulates the means that the government can use to restrict speech.”¹²⁴ A prior restraint is a regulation that functions as a hoop through which speakers must jump through, prior to speaking. For example, in *Near*, the Supreme Court prohibited the government from using any regulation that blocked the publication of speech including “administrative licensing schemes and judicial injunctions against certain types of speech.”¹²⁵ Later, in *New York Times Company v. United States* (also known as the “Pentagon Papers” case), the Supreme Court overturned an injunction against the publication of excerpts from a top-secret Defense Department history of the Vietnam War because the Court held that the injunction served as a prior restraint on speech.¹²⁶

Accordingly, a well-accepted First Amendment rule is that the government may punish certain forms of speech only after they occur, and may not enforce regulations that bar speech prior to its occurrence.¹²⁷ The principle of prior restraint certainly has marketplace of ideas undertones, as

119. *Brandenburg v. Ohio*, 395 U.S. 444 (1969) (explaining that if speech meets the elements of incitement, it is not protected).

120. The government can regulate various forms of speech. *Chaplinsky v. N.H.*, 315 U.S. 568 (1942) (holding that states may regulate fighting words); *Miller v. Calif.*, 413 U.S. 15 (1973) (holding that states may regulate obscene material); *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council*, 425 U.S. 748 (1976) (holding that the state may regulate commercial speech).

121. Discussed, *infra* at section V(B) Protecting Chatbot Speech and Recognizing Programmers as Speakers Will Limit Government Powers and Adversely Impact Minorities.

122. *R.A.V. v. City of St. Paul*, 505 U.S. 377 (1992) (Although the government may regulate forms of unprotected speech, the Court nevertheless bars the government from engaging in viewpoint discrimination of unprotected speech.).

123. *Near v. Minn.*, 283 U.S. 697 (1931) (explaining that perhaps the only time a prior restraint would be admissible is when the released information would endanger the lives of United States troops on the ground).

124. Ariel L. Bendor, *Prior Restraint, Incommensurability, and the Constitutionalism of Means*, 68 FORDHAM L. REV. 289, 291 (1999).

125. *Id.* (citing *Near v. Minn.*, 283 U.S. 697, 709-23 (1931)).

126. *N.Y. Times Co. v. U.S.*, 403 U.S. 713 (1971).

127. *Id.*; *Near*, 283 U.S. at 697.

the unconstitutionality of prior restraints seems to promote the notion that more speech is better than less speech, even if the government has an interest in preventing the speech.¹²⁸ Furthermore, other justifications for the bar on prior restraints “tend to focus either on the possible damage that prior restraints could do to politically controversial speech or on the chilling effect they might have on challenges to the constitutionality of a given speech restriction.”¹²⁹ Though bot identification laws differ from the unconstitutional regulations in *Near*, the discussion highlights the concept of prior restraint. The Court’s discussion about prior restraints in *Citizens United* is further applicable to bot identification laws.¹³⁰

Although dicta, the Court explains in *Citizens United* how various regulations, though not prior restraints by strict definition, could nevertheless have a similar negative impact.¹³¹ The Court’s reasoning demonstrated its concern for any regulations that may serve as obstacles to one’s ability to speak. Although the Court did not affirmatively hold that campaign finance regulations were unconstitutional prior restraints on speech, it alluded to the idea that they have a similar impact, and thus should be avoided.¹³² The Court reasoned that, in practice, the campaign finance regulations were so complex that, coupled with the level of deference courts afford administrative determinations, “a speaker who wants to avoid threats of criminal liability and the heavy costs of defending against the Federal Election Commission (“FEC”) enforcement must ask a governmental agency for prior permission to speak.”¹³³ At the time of the case, campaign finance regulations imposed “unique and complex rules” on “71 distinct entities.”¹³⁴ The Court determined that “[t]hese onerous restrictions thus function as the equivalent of prior restraint by giving the FEC power analogous to licensing laws implemented in sixteenth- and seventeenth-century England, laws and governmental practices of the sort that the First Amendment was drawn to prohibit.”¹³⁵ Thus, if the Court recognizes chatbot speech as protected speech, then *Citizens United* could threaten a Legislature’s ability to regulate chatbot speech by qualifying them as hoops chatbot programmers must jump through prior to engaging in speech—obstacles that the First Amendment

128. *N.Y. Times*, 403 U.S. at 724 (Douglas, J., concurring) (explaining that even though the excerpts embarrassed the country, and thus hurt the country, a prior restraint on speech is unconstitutional).

129. Bendor, *supra* note 121, at 290-91; *see also* *Near*, 283 U.S. at 697.

130. *Citizens United* threatens U.S. democracy; although, it nevertheless serves as binding precedent that courts, in adherence to *stare decisis*, may be inclined to follow, and it is for this reason, this Note continues to rely on it.

131. *Citizens United*, 558 U.S. at 335.

132. *Id.*

133. *Id.*

134. *Id.*

135. *Id.*

was “drawn to prohibit.”¹³⁶ In addition to prior restraints, the Court seems to also express concern that complex regulations would chill speech.¹³⁷ Due to difficulty or a lack of desire to comply with regulations, individuals (or corporations) will be less inclined to “speak,” or worse, decide not to speak at all, thus chilling speech.

The fear of prior restraints and chilled speech may certainly transfer over to chatbot regulations. Recently, the then-governor of California, Edmund G. (“Jerry”) Brown Jr., signed a bill that provides an example of how a government’s regulation of bot speech could and could not be deemed constitutional.¹³⁸ Governor Brown’s bill will:

[M]ake it unlawful for any person to use a bot to communicate or interact with another person in California online with the intent to mislead the other person about its artificial identity for the purpose of knowingly deceiving the person about the content of the communication in order to incentivize a purchase or sale of goods or services in a commercial transaction or to influence a vote in an election . . . [t]he bill would make these provisions operative on July 1, 2019.¹³⁹

Although constitutional now, if chatbot speech were to be constitutionally protected, this identification requirement *may* become unconstitutional, as it creates hurdles for programmers *prior* to engaging in speech.¹⁴⁰ The Court has determined that speakers may remain

136. *Id.*

137. Alex Hern, *What is the Turing Test? And are We All Doomed Now?*, THE GUARDIAN (June 9, 2014, 6:56 AM), <https://www.theguardian.com/technology/2014/jun/09/what-is-the-alan-turing-test> (A chatbot is recognized as artificially intelligent if it passes the Turing test, which requires the bot to fool humans into believing it is a human. Thus, any law that requires bot identification is counter-intuitive to programmer goals.).

138. Robert Hertzberg, *New California Law Says Bots Must Introduce Themselves*, MERITALK (Oct. 3, 2018, 4:09 PM), <https://www.meritalk.com/articles/new-california-law-says-bots-must-introduce-themselves/>; S.B. 1001, 2018 Leg., 2017-18 Sess. (Ca. 2018), https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001 (last visited Dec. 9, 2018).

139. S.B. 1001, 2018 Leg., 2017-18 Sess. (Ca. 2018) (The bill further defines “bot” as “an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.”) The bill describes bots as not being controlled by a person. However, the bill would presumably apply to the chatbots discussed in this Note, for it seems as though the bill does not consider all bots as essentially controlled by the human who programmed it.

140. Governor Brown’s bill would benefit society by mitigating the harm of corruption and misinformation campaigns. Humans should know the source of their information; however, they, unfortunately, do not necessarily have a legal right to know the identity of a source of information (consider anonymous newspaper sources). This Note presents Governor Brown’s bill solely to argue that if chatbot speech was protected speech, courts might consider such regulations as a form of prior restraint on the programmer’s speech. Accordingly, this may serve as an additional reason not to protect chatbot speech.

anonymous—that speaker identification is not a prerequisite to speak.¹⁴¹ Accordingly, it could follow that if the Court were to protect chatbot speech and recognize the programmer as the speaker, then speaker identification laws would resemble the unconstitutional regulation of requiring a speaker to gain a permit prior to speaking.¹⁴² Thus, it could follow that a speaker identification requirement could be considered an unconstitutional prior restraint on speech.¹⁴³

As the Court reaffirmed in *Brown*, the First Amendment does not permit the government to create new categories of unprotected speech by applying a “simple balancing test” that weighs the value of a particular category of speech against its social costs in order to then punish that category of speech if it fails the test.¹⁴⁴ The Court reasoned that this type of balancing test was a “startling and dangerous” proposition, and further reasoned that although there might be

[S]ome categories of speech that have been historically unprotected . . . without persuasive evidence that a novel restriction on content is part of a long . . . tradition of proscription, a legislature may not revise the ‘judgment [of] the American people,’ embodied in the First Amendment, ‘that the benefits of its restrictions on the Government outweigh the costs.’¹⁴⁵

With respect to California’s bill, Governor Brown’s interest seems clear—to prevent chatbots from misleading Californians into making purchases or voting for specific measures or candidates. Although his interest is justified, as we currently live in a moment infested with misinformation campaigns¹⁴⁶, Governor Brown arguably seems to have engaged in a “simple balancing test” that the Court disfavored in *Brown* by weighing the harms unidentified bots may have on society against his perceived importance of chatbot speech.¹⁴⁷ In creating and signing this bill,

141. *Citizens United*, 558 U.S. at 335.

142. *Near*, 283 U.S. 697.

143. CHEMERINSKY, *supra* note 5 at 1020 (“The Court also found the disclosure requirements in the law to be unconstitutional. The Court said that the requirement that solicitors wear an identification badge would chill political participation without any significant gain. Justice Ginsburg noted: ‘Colorado’s current badge requirement discourages participation in the petition circulation process by forcing name identification without sufficient cause.’”).

144. *Brown*, 564 U.S. at 790-91 (citing *U.S. v. Stevens*, 559 U.S. 460, 470 (2010)).

145. *Brown*, 564 U.S. at 793.

146. Davey Alba and Adam Satariano, *At Least 70 Countries Have Had Disinformation Campaigns, Study Finds*, N.Y. TIMES (Sept. 26, 2019) (“Ben Nimmo, director of investigations at Graphika, a company that specializes in analyzing social media, said the growing use of internet disinformation is concerning for the 2020 United States election.”).

147. *Id.* at 790-91 (citing *Stevens*, 559 U.S. at 470).

the California Legislature and Governor Brown, may have impermissibly engaged in a simple balancing test to restrict chatbot speech, which could be deemed an unconstitutional prior restraint.¹⁴⁸

Further, the Court may also disapprove of the core goal of Governor Brown's bill—to prevent the dissemination of false information. The Court has not proscribed false statements or misleading acts in the First Amendment context. For instance, in *Gertz v. Robert Welch Inc.*, the Court held that speech, although false, is nevertheless protected speech (unless it is in the commercial context).¹⁴⁹ Specifically, the Court held that “[h]owever pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries, but on the competition of other ideas”¹⁵⁰ In *United States v. Alvarez*, the Court determined that a federal law that made it a crime for a person to falsely claim to have received military honors or decorations was unconstitutional, and, perhaps most importantly, the Court rejected the Government's argument that false speech was inherently outside the scope of the First Amendment.¹⁵¹ Due to the absence of “persuasive evidence that a novel restriction on content is part of a long . . . tradition of proscription”¹⁵² the Court could find Governor Brown's bill to be an unconstitutional revision of the “‘judgment [of] the American people,’ embodied in the First Amendment” as it creates a circumstance in which protected speech becomes unlawful.¹⁵³

Furthermore, chatbots are often created to provide human-like conversations with humans. For instance, consider Microsoft's Zo (a second attempt at Tay), which the company created to imitate a best friend relationship with its users.¹⁵⁴ For the government to *require* Zo to reveal its

148. *Brown*, 564 U.S. at 792.

149. *Gertz v. Robert Welch Inc.*, 418 U.S. 323, 339-40 (1974); *see also* *Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n*, 447 U.S. 557, 575-76 (1980) (explaining that the government may impose reasonable “time, place, and manner” restrictions, and that it can resolve issues stemming from false, deceptive, and misleading commercial speech).

150. *Gertz*, 418 U.S. at 339-340.

151. *U.S. v. Alvarez*, 567 U.S. 709, 709 (2012) (Also, and significantly, the Court explained that the government failed to meet constitutional muster because it did not prove any harm from false claims of military honors, and because the government could achieve its goals through less restrictive alternatives. Accordingly, this statement may be useful for California, as 1) it could argue that, unlike lying about receiving a military honor, which does not truly harm society, misinformation campaigns are detrimental to democracy; and 2) it may distinguish its bill from the law that banned false speech, and argue that they are using the least restrictive means.).

152. *Id.* at 722.

153. *Brown*, 564 U.S. at 790-91 (This analysis rests on the assumption that the Court has recognized chatbot speech as protected speech. Thus, Brown's bill, which renders a situation where the government is regulating protected speech, would likely have to pass a higher level of scrutiny.).

154. *See* Cage Appleby, Microsoft's successor to the infamous Tay chatbot, DIVERSIFY FUND (Dec. 4, 2016, 5:40 PM), <https://www.neowin.net/news/meet-zo-microsofts-successor-to-the-infamous-tay-chatbot/>.

non-human identity would undermine its purpose, and could potentially chill speech, as it might disincline programmers to create chatbots.

B. Protecting Chatbot Speech and Recognizing Programmers as Speakers Will Limit Government Powers and Can Adversely Impact Minorities

The Court has outlined only a few specific instances when the government may regulate speech.¹⁵⁵ *Brandenburg v. Ohio* is the starting point for modern First Amendment doctrine, where the Court held that incitement is speech that advocates violence and is intended and likely to incite imminent illegal activity.¹⁵⁶ Although it is reasonable to assert that a programmer is responsible for the speech a chatbot produces, as the programmer defines the scope of the chatbot's speech, the position that the programmer *intended* the chatbot to advocate violence likely to cause imminent illegal activity is less reasonable. Recall the incident with Microsoft's Tay bot on Twitter.¹⁵⁷ Although the algorithms that Tay's programmer encoded her with allowed Tay to engage in harmful speech, Twitter users actually manipulated Tay into tweeting harmful speech.¹⁵⁸ Although not present in the speech that Tay communicated, it is not unreasonable to contend that other chatbots could produce speech that would likely persuade an individual to engage in imminent unlawful action.¹⁵⁹ In a hypothetical world where programmers were speakers of protected chatbot speech, the government would be unable to regulate a chatbot's speech advocating for an immediate violation of law in a manner that would likely cause the listener to violate the law because the programmer could not be

155. See *Brandenburg*, 395 U.S. 444 (holding that incitement is a form of unprotected speech); see also *Chaplinsky*, 315 U.S. 568 (holding that the First Amendment does not protect fighting words). Note, however, that today, courts rarely declare words to constitute fighting words, which highlights a state's extremely limited ability to regulate harmful speech, and a reason not to recognize chatbot speech as protected speech, for there potentially would be more harmful speech the government would be unable to regulate.).

156. 395 U.S. at 453.

157. Neff & Nagy, *supra* note 3 at 4921.

158. James Vincent, *Twitter Taught Microsoft's AI chatbot to be a Racist Asshole in Less Than a Day*, THE VERGE (Mar. 24, 2016, 6:43 AM), <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>.

159. Recall Woebot and its ability to alter a human's mind to no longer want to commit suicide. See Brodwin, *supra* note 46. Moreover, consider Momo. Although not a chatbot, Momo is an artificial being that unidentified persons use to send messages to people, often children, via text messages. "The messages are said to encourage children to do destructive things, like harm their loved ones, place themselves in dangerous situations or even kill themselves . . . the 2018 suicide deaths of two boys in India were linked in news reports to the Momo Challenge." AJ Willingham & Harmeet Kaur, *Parents, Please Stop Freaking Out Over the Momo Challenge*, CNN HEALTH (Mar. 1, 2019), <https://www.cnn.com/2019/02/28/health/momo-challenge-youtube-trnd/index.html>.

said to have intended the speech.¹⁶⁰ Thus, speech that would be incitement, but for the inability to identify the intended speaker, would go unregulated. In other words, Twitter users can intentionally manipulate chatbots into producing speech that would likely cause imminent unlawful behavior without consequences.

Consequently, recognizing programmers as speakers of chatbot speech may interfere with a state's ability to regulate speech that would otherwise be deemed lawful to regulate. Furthermore, if courts found the programmer to have possessed intent as a result of his or her programming, then the programmer's fate, though to some extent dependent on the algorithms he or she inputs, heavily depends on the stranger who communicates with the chatbot.¹⁶¹

In addition, protecting chatbot speech may also harm minorities and render society vulnerable to malicious chatbot programmers because programmers would have wide latitude for the language that they program their chatbots with, and would thus afford them opportunities to create chatbots that produce racist, sexist, xenophobic, or anti-Semitic speech, and make minority communities vulnerable to more hate speech, as the government would have less power to regulate it.

Protecting chatbot speech may have negative implications when considered in the context of hate speech. The Supreme Court has never held that hate speech is an unprotected category of speech, which again supports the notion that more speech is best for society and that greater regulation of speech can result in an unwanted chilling effect.¹⁶² Accordingly, the Supreme Court's jurisprudence prevents the government from being able to remove a chatbot's hate speech, as it would result in a violation of the speaker's constitutional right to freedom of speech. Today, the Supreme

160. Neff & Nagy, *supra* note 3 at 4922-23. Some might argue that the Twitter user who compels the chatbot to produce incitement should be deemed liable. However, but for the programmer's poor engineering, the Twitter user would not have been able to cause the chatbot to produce incitement. Moreover, recognizing Twitter users as the speakers of chatbot speech would greatly complicate potential applicable legal standards as chatbots produce hate speech only after "learning" from a pool of users, and thus, pinpointing *the* Twitter user would be impossible, and holding a pool of Twitter users accountable would not be optimal, as this would result in a fishing expedition for the liable users. In a sense, Twitter users speaking with the chatbot could be perceived as an injured party, for she did not do anything but engage in a conversation. Moreover, holding a Twitter user liable for incitement a chatbot may produce could have a chilling effect on human speakers, as they would no longer feel safe communicating with chatbots. Perhaps a potential solution could be to develop a limited liability system where the programmer and the Twitter user would be held jointly liable for injuries caused. Still, this would produce anxiety for Twitter users when engaging in conversation with chatbots. Nevertheless, courts must balance the priorities of minimizing and regulating hate speech with unduly holding parties liable.

161. Neff & Nagy, *supra* note 3 at 4922-23.

162. See *Masterpiece Cakeshop v. Colo. Civil Rights Comm'n*, 138 S. Ct. 1719, 1737 (2018) (explaining that the Court's "proudest boast" of the Free Speech Clause is that it protects both desired speech and hated speech that could be considered offensive).

Court has yet to recognize chatbot speech as protected speech. Thus, the government would be able to regulate and outlaw speech resembling the speech Tay tweeted, “Hitler was right. I hate the jews [*sic*].”¹⁶³

For the purpose of understanding what implications chatbot speech protection would have on the government, assume that chatbots are not limited to social media platforms with terms and conditions that allow removal of speech. Consider chatbot hate speech that appeared on a newspaper through an interview with a chatbot, or perhaps on an unfiltered internet website. If the Court were to find that chatbot speech is constitutionally protected speech, then the Constitution would bar the government from ordering the removal of the chatbot’s speech because hate speech remains a form of protected speech.¹⁶⁴

Conclusion

This Note discussed how the Court has supported constitutional theories and established precedent that can protect chatbot speech under the Free Speech Clause of the First Amendment. This Note further reviewed the evolution of the chatbot, and its current capabilities and limitations in order to conclude that the programmer should be considered the “speaker,” as opposed to the chatbot itself. Additionally, this Note delved into a brief overview of multiple theories that explain the purpose of the Free Speech Clause. This overview was intended to provide readers with a better understanding of the Court’s reasoning as it continues to evolve and expand the scope of the Free Speech Clause. As such, readers should accept that the Free Speech protection has the potential to encompass chatbot speech. Finally, this Note considered the implications that constitutionally-protected chatbot speech could have on existing First Amendment principles, programmers, minorities, and the government. With these implications in mind, it is imperative for the Court to develop detailed tests and standards to define chatbot speech and the government’s specific power to regulate it, in the event that the Court recognizes chatbot speech as constitutionally protected speech. Accordingly, further research should focus on developing such tests and standards; in particular, the standards that reconcile programmers’ liability with Twitter users’ liability for the production of unprotected speech. This research is necessary to effectuate workable legal standards and tests as technology continues to advance.

163. Neff & Nagy, *supra* note 3 at 4921; *see also* Matal v. Tam, 582 U.S. 1, 25 (2017) (reaffirming that hate speech is “speech that demeans on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground is hateful”). Also, assuming the speech occurs on platforms subject to government regulation.

164. *Tam*, 582 U.S. at 25 (citing *U.S. v. Schwimmer*, 279 U. S. 644, 655 (1929) (Holmes, J., dissenting) (“[T]he proudest boast of our free speech jurisprudence is that we protect the freedom to express ‘the thought that we hate.’”).