

# LET’S GET REAL: WEAK ARTIFICIAL INTELLIGENCE HAS FREE SPEECH RIGHTS

James B. Garvey\*

*The right to free speech is a strongly protected constitutional right under the First Amendment to the U.S. Constitution. In 2010, the U.S. Supreme Court significantly expanded free speech protections for corporations in Citizens United v. FEC. This case prompted the question: could other nonhuman actors also be eligible for free speech protection under the First Amendment? This inquiry is no longer a mere intellectual exercise: sophisticated artificial intelligence (AI) may soon be capable of producing speech. As such, there are novel and complex questions surrounding the application of the First Amendment to AI. Some commentators argue that AI should be granted free speech rights because AI speech may soon be sufficiently comparable to human speech. Others disagree and argue that First Amendment rights should not be extended to AI because there are traits in human speech that AI speech could not replicate.*

*This Note explores the application of First Amendment jurisprudence to AI. Introducing relevant philosophical literature, this Note examines theories of human intelligence and decision-making in order to better understand the process that humans use to produce speech, and whether AI produces speech in a similar manner. In light of the legal and philosophical literature, as well as the Supreme Court’s current First Amendment jurisprudence, this Note proposes that some types of AI are eligible for free speech protection under the First Amendment.*

INTRODUCTION.....	954
I. AI, THE FIRST AMENDMENT, AND HUMAN INTELLIGENCE.....	957
A. <i>What Is AI?</i> .....	957
B. <i>The First Amendment and Freedom of Speech</i> .....	961
1. Content-Based and Content-Neutral Speech Restrictions.....	962

---

\* J.D. Candidate, 2023, Fordham University School of Law; B.A., 2019, Binghamton University. First, many thanks to Professor Abner S. Greene for all his guidance. Thank you to my fantastic editor, Jamie Reinah, and the other editors of the *Fordham Law Review* for their time and feedback. I also would like to thank my family for their constant encouragement, love, and support.

2. Corporate Speech .....	966
3. Theoretical Justifications for Free Speech .....	968
C. <i>Theories of Cognition</i> .....	970
1. Philosophical Theories of Human Intelligence .....	970
2. The Role of Emotion in Human Decision-Making ....	972
II. ISSUES IN THE APPLICATION OF FIRST AMENDMENT RIGHTS TO AI.....	975
A. <i>Doctrinal and Theoretical Issues for Nonhuman     Speech</i> .....	975
B. <i>Philosophy of Cognition and AI</i> .....	979
III. WEAK AI SHOULD RECEIVE FREE SPEECH RIGHTS UNDER THE FIRST AMENDMENT.....	982
A. <i>Weak AI Is Similar to Human and Corporate     Speakers</i> .....	982
B. <i>Weak AI Has Relevantly Similar Traits to Some     Protected Human Speech</i> .....	986
C. <i>“Dogood” the AI</i> .....	987
CONCLUSION.....	991

#### INTRODUCTION

Artificial intelligence (AI)<sup>1</sup> is increasingly entrenched in our daily lives and government. AI technologies serve in militaries,<sup>2</sup> categorize emergency calls,<sup>3</sup> grade exams,<sup>4</sup> and predict when crimes might happen and recommend appropriate responses.<sup>5</sup> Perhaps, AI could even write briefs and be a judge.<sup>6</sup> With significant involvement in civil society and increasing sophistication, AI will likely be used to make “speech” in the near future.

---

1. This Note uses the term “AI” to refer to both artificial intelligence as a descriptor and artificial intelligence as a type of program or technology.

2. See Gerrit De Vynck, *The U.S. Says Humans Will Always Be in Control of AI Weapons. But the Age of Autonomous War Is Already Here.*, WASH. POST (July 7, 2021, 10:00 AM), <https://www.washingtonpost.com/technology/2021/07/07/ai-weapons-us-military/> [https://perma.cc/L28B-KRR3].

3. See Gregory S. Dawson, Kevin C. Desouza & Rashmi Krishnamurthy, *Learning from Public Sector Experimentation with Artificial Intelligence*, BROOKINGS (June 23, 2017), <https://www.brookings.edu/blog/techtank/2017/06/23/learning-from-public-sector-experimentation-with-artificial-intelligence/> [https://perma.cc/5ENW-8W6V].

4. See Emma Martinho-Truswell, *How AI Could Help the Public Sector*, HARV. BUS. REV. (Jan. 29, 2018), <https://hbr.org/2018/01/how-ai-could-help-the-public-sector> [https://perma.cc/A5F5-WJKP].

5. See Will Douglas Heaven, *Predictive Policing Algorithms Are Racist. They Need to Be Dismantled.*, MASS. INST. TECH. REV. (July 17, 2020), <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/> [https://perma.cc/BJH7-GN48].

6. See Eugene Volokh, *Chief Justice Robots*, 68 DUKE L.J. 1135, 1156 (2019).

AI is generally described as a program that performs tasks with a “‘human-like’ intelligence or better.”<sup>7</sup> Current AI technology is primarily “goal-oriented,” meaning that a human programmer decides what goal the AI will pursue, but the AI, largely independently, decides how to achieve that goal.<sup>8</sup> This means that, at least sometimes, the AI will act in unintended ways not totally controlled or anticipated by the programmer.<sup>9</sup> While a skittish self-driving car slamming the brakes for a plastic bag is not particularly concerning, another AI, perhaps a military program, could cause serious unintended consequences as the AI diligently seeks to achieve its programmed goal.<sup>10</sup>

There are two types of AI: “weak” or “artificial narrow intelligence” (ANI)<sup>11</sup> and “strong” or “artificial general intelligence” (AGI).<sup>12</sup> ANI is described as a “one-trick pony” and performs a particular task in a particular way, like a self-driving car or search engine.<sup>13</sup> ANI is the most common type of AI; most of the recent technological progress in AI has been made with respect to ANI.<sup>14</sup> AGI is a type of AI that can do anything a human can do but is currently more of a concept or theory with less technological advancement.<sup>15</sup>

Recent developments in AI technology and utilization prompt questions concerning whether AI is eligible to receive free speech protections under the First Amendment to the U.S. Constitution.<sup>16</sup> Unlike a traditional computer program, which produces speech that can be directly traced back to its creator,<sup>17</sup> an AI’s speech would be significantly more independent because AI do not rely on a precise set of instructions or code when producing speech.<sup>18</sup> While some scholars, including Professors Toni M. Massaro and Helen Norton, have argued that it is plausible for strong AI to be agents with free speech protections,<sup>19</sup> strong AI is unlikely to be developed in the

---

7. NOAH WAISBERG & ALEXANDER HUDEK, AI FOR LAWYERS: HOW ARTIFICIAL INTELLIGENCE IS ADDING VALUE, AMPLIFYING EXPERTISE, AND TRANSFORMING CAREERS 5 (2021).

8. See Natalie Wolchover, *Artificial Intelligence Will Do What We Ask. That’s a Problem.*, QUANTA (Jan. 30, 2020), <https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/> [<https://perma.cc/MU8Q-QC63>].

9. See *id.* (“When programmers try to list all goals and preferences that a robotic car should simultaneously juggle, the list inevitably ends up incomplete.”). An alternate theory on how to structure AI, promoted by Professor Stuart J. Russell, is that AI should aim to identify and promote human preferences rather than specific goals. See *id.*

10. See *id.*

11. This Note uses “weak AI” and “ANI” interchangeably.

12. This Note uses “strong AI” and “AGI” interchangeably.

13. Andrew Ng, *AI for Everyone: Introduction*, COURSERA (Feb. 2019), <https://www.coursera.org/learn/ai-for-everyone/lecture/SRwLN/week-1-introduction> [<https://perma.cc/YN52-GXEN>].

14. See *id.*

15. *Id.*

16. U.S. CONST. amend. I.

17. See *infra* Part I.A.

18. See *infra* Part I.A.

19. See Toni M. Massaro & Helen Norton, *SIRI-Ously?: Free Speech Rights and Artificial Intelligence*, 110 NW. L. REV. 1169, 1172 (2016).

foreseeable future.<sup>20</sup> Thus, the question of whether weak AI, the most common type of AI, could be eligible for First Amendment protection is an important one.

In its First Amendment jurisprudence, the U.S. Supreme Court has repeatedly held that when the government regulates the content of an individual's speech, those regulations are subject to the highest level of constitutional scrutiny.<sup>21</sup> However, the Supreme Court has also held that when the government regulates speech *irrespective of its content*, those regulations are reviewed under various formulations of intermediate scrutiny.<sup>22</sup> Relying on these principles, the Court recently extended free speech protections to nonhuman speakers.<sup>23</sup> In *Citizens United v. FEC*,<sup>24</sup> a case involving political campaign financing, the Court held that a corporation is sufficiently similar to a traditional human speaker and is thus eligible for free speech protection under the First Amendment.<sup>25</sup> Broadly speaking, the Court in *Citizens United* extended the constitutional analysis it applies to traditional human speakers to nonhuman speakers, which is particularly important when considering whether AI should be granted free speech rights.<sup>26</sup>

This Note discusses the application of First Amendment free speech doctrine to AI.<sup>27</sup> Since AI is not a typical “speaker”—that is, not a natural person or legal entity—this Note addresses whether AI produces speech in a way that is relevantly similar to protected human or corporate speech. To answer this question, this Note compares current Supreme Court doctrine regarding corporate speech and current literature regarding protections for AI with the modern philosophical and psychological understanding of “intelligence.” This Note explores human emotional intuitions and their relationship to moral decisions.<sup>28</sup> This Note then considers the implications

---

20. See Ng, *supra* note 13 (“[T]here’s a lot of progress in [weak] AI, which is true. But that has caused people to falsely think that there might be a lot of progress in [strong AI] as well which is leading to some irrational fears about evil clever robots coming over to take over humanity anytime now. I think [strong AI] is an exciting goal for researchers to work on, but . . . it may be decades or hundreds of years or even thousands of years away.”).

21. The Supreme Court has characterized these types of regulations as “content based.” See *Reed v. Town of Gilbert*, 576 U.S. 155, 163 (2015) (“Government regulation of speech is content based if a law applies to particular speech because of the topic discussed or the idea or message expressed.”). Content-based regulations are subject to strict scrutiny. See *id.*; see also *infra* Part I.A.

22. The Supreme Court has characterized these types of regulations as “content neutral.” See generally *Ward v. Rock Against Racism*, 491 U.S. 781 (1989); *United States v. O’Brien*, 391 U.S. 367, 376–77 (1968). For a more detailed discussion on various types of content-neutral speech regulations, see *infra* Part I.B.1.

23. See generally *Citizens United v. FEC*, 558 U.S. 310, 372 (2010).

24. 558 U.S. 310 (2010).

25. See *id.* at 341.

26. See *infra* Part I.B.2.

27. Because strong AI has not yet been fully developed, this Note will focus exclusively on whether weak AI is eligible for free speech protection under the First Amendment.

28. See Jonathan Haidt, *The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 PSYCH. REV. 814, 814 (2001).

of the relationship between intuitions and decisions on free speech doctrine as it relates to AI that possesses comparable intelligence to humans.<sup>29</sup>

Part I provides relevant background on AI, the First Amendment, and philosophical foundations on human cognition relevant to the production of speech. Part II explains why current First Amendment theory will likely be applied to weak AI and analyzes the current philosophical debate regarding AI's consciousness. Part III proposes that, in some circumstances, weak AI will possess relevantly comparable intelligence to humans and produce similar speech and, therefore, will be eligible for free speech protection under the First Amendment.

### I. AI, THE FIRST AMENDMENT, AND HUMAN INTELLIGENCE

AI technology is quickly developing, prompting questions of whether AI could be eligible for free speech protections under the First Amendment. This part provides an overview of the development of AI technology, free speech doctrine, and philosophical discussions regarding human intelligence. Part I.A provides an overview of basic AI technology and implications for the development of AI in the near future. Part I.B discusses the Supreme Court's current First Amendment jurisprudence, focusing specifically on the Court's recent cases involving other nonhuman speech actors, corporations. Part I.B also explains the current philosophical and legal theories justifying freedom of speech. Lastly, Part I.C identifies the modern debate among scholars over human consciousness and intelligence.

#### A. *What Is AI?*

Weak AI, the most common type of AI, may be described as a program that performs a particular, narrow type of task, such as winning a chess match, with “‘human-like’ intelligence or better.”<sup>30</sup> Weak AI designed to play chess is wholly incapable of driving a car or analyzing trends in home sales—it only plays chess.<sup>31</sup> Most of the recent technological progress in AI has been made with respect to weak AI.<sup>32</sup>

Traditional computer programs are designed in a rules-oriented fashion.<sup>33</sup> More specifically, traditional computer programs are created using a series of conditional statements that decide what, if any, action is taken.<sup>34</sup> With

---

29. See Volokh, *supra* note 6, at 1137–38.

30. WAISBERG & HUDEK, *supra* note 7, at 5. This Note focuses on weak AI and thus uses “AI” and “weak AI” interchangeably.

31. See, e.g., Ng, *supra* note 13.

32. See *id.*

33. See *id.*

34. See *id.* For instance, if a computer user clicks the spam button in their email inbox, the email program is designed to remove the email from the user's inbox, place it in the spam folder, and mark it for deletion in thirty days. Although this process feels seamless in the email context, it is very difficult to implement in large-scale projects. See Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 90–95 (2014).

traditional computer programs, human programmers must handwrite every conditional statement, a time-consuming and therefore costly process.<sup>35</sup>

AI can use different technological processes to improve efficiency, including machine learning systems.<sup>36</sup> Machine learning is a process whereby a program, independent from human control, reviews data and adjusts its procedures to improve efficiency and accuracy.<sup>37</sup> Instead of programming a specific set of rules, the creator “turns the key,” and the program adapts on its own.<sup>38</sup> This allows designers to increase the scale of projects far beyond what humans are capable of. As a result, AI can include more variables and data points than a human ever could.<sup>39</sup> Additionally, machine learning can help identify human error, allowing programmers—or AI—to spot the issue and correct it.<sup>40</sup> Since much of the work is done autonomously, machine learning programs are far more cost-effective as compared to traditional programming.<sup>41</sup>

A common version of a machine learning system is “pattern-based.”<sup>42</sup> An example of a pattern-based machine learning system is an automatic spam filter.<sup>43</sup> When a human user marks an email as spam, an AI reviews the email for discernable patterns.<sup>44</sup> Emails with phrases such as “reverse-mortgage” or “extended warranty” tend to be spam, so the AI can identify emails as spam based on these common phrases. The AI may also recognize spam emails based on other patterns, such as senders from particular locations.<sup>45</sup> This process alone is useful, but it becomes exponentially more effective once the AI starts to identify more complex patterns and makes connections between previously recognized patterns.<sup>46</sup> For instance, the AI operates more efficiently when it can identify an email as spam based on both the phrase “extended warranty” and its sender’s particular location.<sup>47</sup>

An AI’s process of pattern recognition, pattern review, and rule formation can occur without human input or oversight.<sup>48</sup> Machines—and AI—“learn”

---

35. See Surden, *supra* note 34, at 93–94.

36. See *id.* at 94.

37. See *id.* This Note assumes that weak AI will utilize machine learning.

38. See Ng, *supra* note 13.

39. See *id.*

40. See *id.*

41. See *id.*

42. See Surden, *supra* note 34, at 90.

43. See *id.*

44. See *id.*

45. See *id.* at 92.

46. See *id.* at 91.

47. See *id.*

48. See Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3044–45 (2014). This process is termed “unsupervised learning.” See *Unsupervised Learning*, IBM, <https://www.ibm.com/cloud/learn/unsupervised-learning> [https://perma.cc/8CMJ-9CZ9] (last visited Nov. 7, 2022). Specifically, the AI is given a set of unlabeled data and discerns the patterns itself. See *id.* This process is different from the spam email filter example, where the data sets are labeled, and the AI can easily identify the patterns. See Julianna Delua, *Supervised vs. Unsupervised Learning: What’s the Difference?*, IBM

or “adapt” by utilizing programming tools written by the creators to achieve preestablished goals. This process is best described by Professor Tom M. Mitchell: “A computer program is said to *learn* from experience *E* with respect to some class of tasks *T* and performance measure *P*, if its performance at tasks in *T*, as measured by *P*, improves with experience *E*.”<sup>49</sup> Returning to the spam email example discussed above, the task is to label certain emails as spam, the performance measure is the accuracy of the labeling, and the experience is the process of labeling each individual email. So the spam email filter program “learns” by recognizing more patterns and becoming more accurate at labeling.

AI using machine learning typically acts contrary to designer expectations because of problems in the design of the performance measure or due to a lack of data.<sup>50</sup> For example, in the design of facial recognition software, AI were consistently worse at recognizing darker-skinned females compared to any other group.<sup>51</sup> After reviewing the systems, the designers sought to correct the imbalance and improved accuracy<sup>52</sup> by modifying the performance measure, specifically changing the “testing cohorts” and by increasing data collection to include specific demographic data.<sup>53</sup> The spam email filter program and facial recognition program examples both demonstrate the way in which extensive and varied data is essential for AI to act congruently with human expectations.

As compared to both traditional programming and human operations, AI can be designed to be different in degree. Traditional programming cannot be used in very complex predictive systems, such as cancer-screening programs.<sup>54</sup> It would simply be too complex for humans to program a series of conditional statements accurate enough to examine and identify radiological reports. AI can do that and can be as good as doctors at reading the reports.<sup>55</sup> In this example, the difference between an AI and a human doctor is one of degree—the AI can be quicker but, for all intents and purposes, is performing a process similar to the human doctor’s. However,

---

(Mar. 12, 2021), <https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning> [<https://perma.cc/D9WL-CCP9>].

49. TOM M. MITCHELL, *MACHINE LEARNING 2* (1997).

50. Alex Najibi, *Racial Discrimination in Face Recognition Technology*, SITN (Oct. 24, 2020), <https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/> [<https://perma.cc/QU96-R44X>].

51. *See id.*

52. *See* INIOLUWA DEBORAH RAJI & JOY BUOLAMWINI, *ACTIONABLE AUDITING: INVESTIGATING THE IMPACT OF PUBLICLY NAMING BIASED PERFORMANCE RESULTS OF COMMERCIAL AI PRODUCTS* (2019), [https://www.aies-conference.com/2019/wp-content/uploads/2019/01/AIES-19\\_paper\\_223.pdf](https://www.aies-conference.com/2019/wp-content/uploads/2019/01/AIES-19_paper_223.pdf) [<https://perma.cc/78QY-XCS8>].

53. *See* Najibi, *supra* note 50.

54. *See Artificial Intelligence: Opportunities in Cancer Research*, NAT’L INSTS. HEALTH (Aug. 31, 2020), <https://www.cancer.gov/research/areas/diagnosis/artificial-intelligence> [<https://perma.cc/YSN5-VZ7N>].

55. *See* Fergus Walsh, *AI ‘Outperforms’ Doctors Diagnosing Breast Cancer*, BBC (Jan. 2, 2020), <https://www.bbc.com/news/health-50857759> [<https://perma.cc/TQ9M-LZC8>].

unlike the human doctor, an AI can account for more factors than a human realistically could.<sup>56</sup>

The implementation of machine learning systems to large-scale projects like cancer screening or facial recognition requires significant amounts of data and processing power to be efficient. Luckily, for AI at least, the necessary technology is here and rapidly improving.<sup>57</sup> Data is the “new oil,” and “big data”<sup>58</sup> is driving rapid advances in AI efficacy.<sup>59</sup> A machine-learning AI system may use big data as a massive reservoir of information to develop more nuanced and effective methods to achieve its purposes.<sup>60</sup> Additionally, improvements in processing power mean that machine learning systems will be better equipped to manage large quantities of data and perform the analysis necessary to function.<sup>61</sup> The combination of big data and advances in processing speed means that machine learning systems, and the AI using them, operate incredibly fast and effectively and are anticipated to continue to improve in the future.<sup>62</sup>

Although AI is continuing to develop, recent technological innovations have shown how AI, using machine learning systems, can operate more efficiently than traditional computer programming, without constant human oversight or input, while often acting in unintended ways.

---

56. See *Artificial Intelligence: Opportunities in Cancer Research*, *supra* note 54. For example, while a human doctor may review a single patient’s radiological report and check for dozens of markers of cancer, the AI can check thousands of factors, discern patterns connecting the factors, and generally reference more data than a human could. See *id.*

57. See Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1154–56 (2018); see also Pearce, *supra* note 48, at 3044–45.

58. Big data “can be defined as data sets whose size or type is beyond the ability of traditional relational databases to capture, manage and process the data with low latency,” and include characteristics such as “high volume, high velocity and high variety.” See *Big Data Analytics*, IBM, <https://www.ibm.com/analytics/hadoop/big-data-analytics> [<https://perma.cc/C9SU-L8GR>] (last visited Nov. 7, 2022). In other words, big data is, for a human user, incomprehensibly large amounts of data. See *id.* “Sources of data are becoming more complex than those for traditional data” because of AI and social media. *Id.*

59. See *Regulating the Internet Giants: The World’s Most Valuable Resource Is No Longer Oil, But Data*, *ECONOMIST* (May 6, 2017), <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> [<https://perma.cc/D359-T7H2>] (“Smartphones and the internet have made data abundant, ubiquitous and far more valuable. Whether you are going for a run, watching TV or even just sitting in traffic, virtually every activity creates a digital trace—more raw material for the data distilleries. As devices from watches to cars connect to the internet, the volume is increasing: some estimate that a self-driving car will generate 100 gigabytes per second. Meanwhile, artificial-intelligence (AI) techniques such as machine learning extract more value from data. Algorithms can predict when a customer is ready to buy, a jet-engine needs servicing or a person is at risk of a disease.”).

60. See Balkin, *supra* note 57, at 1154–56.

61. See Pearce, *supra* note 48, at 3044–45.

62. See *id.*; see also Balkin, *supra* note 57, at 1154–56.

*B. The First Amendment and Freedom of Speech*

The First Amendment unequivocally states that “Congress shall make no law . . . abridging the freedom of speech.”<sup>63</sup> The Supreme Court has interpreted the First Amendment to guarantee broad protections for various types of speech, including offensive words,<sup>64</sup> nonverbal speech acts,<sup>65</sup> and political expenditures.<sup>66</sup> Notably, there can be no general prohibition on hate speech of a political nature.<sup>67</sup> The Court has categorized restrictions on speech into two groups: content based and content neutral.<sup>68</sup> Content-based restrictions are subjected to strict scrutiny.<sup>69</sup> Within content-neutral restrictions, the Court has carved out two subcategories: “time, place, and manner restrictions” and “*O’Brien* restrictions.”<sup>70</sup> Time, place, and manner restrictions are subject to intermediate scrutiny,<sup>71</sup> while restrictions under *United States v. O’Brien*<sup>72</sup> are subject to their own version of intermediate scrutiny.<sup>73</sup> When analyzing government restrictions on speech, the Supreme Court also frequently considers a range of philosophical and legal theories that seek to explain why free speech protections matter in the first place. These theories are particularly helpful to consider when analyzing whether AI is eligible for free speech protections under the First Amendment. Part I.B.1 describes the primary doctrinal analysis used by the Supreme Court when reviewing the constitutionality of restrictions on speech. Part I.B.2 highlights the Court’s recent jurisprudence concerning corporate speech. Part I.B.3 reviews the main philosophical and legal theories used to justify free speech protections.

---

63. U.S. CONST. amend. I.

64. *See, e.g.*, *Cohen v. California*, 403 U.S. 15, 17 (1971) (concluding that the petitioner’s jacket stating “Fuck the Draft” was protected speech).

65. *See, e.g.*, *Tinker v. Des Moines Indep. Cmty. Sch. Dist.*, 393 U.S. 503, 514 (1969) (holding that students were permitted to wear black armbands in an anti-war protest).

66. *See, e.g.*, *Buckley v. Valeo*, 424 U.S. 1, 51 (1976) (*per curiam*) (finding certain limits on expenditures for political contributions unconstitutional).

67. *See Virginia v. Black*, 538 U.S. 343, 366–67 (2003) (holding that a criminal statute considering cross-burning, even as a political message, as providing “prima facie evidence of intent” was unconstitutional).

68. *See Reed v. Town of Gilbert*, 576 U.S. 155, 159 (2015); *see also Ward v. Rock Against Racism*, 491 U.S. 781, 789 (1989).

69. *See Reed*, 576 U.S. at 159 (“[The regulations] are content-based regulations of speech that cannot survive strict scrutiny.”).

70. *See Ward*, 491 U.S. at 789; *United States v. O’Brien*, 391 U.S. 367, 382 (1968).

71. *See Ward*, 491 U.S. at 791 (“[T]he government may impose reasonable restrictions on the time, place, or manner of protected speech, provided the restrictions ‘are justified without reference to the content of the regulated speech, that they are narrowly tailored to serve a significant governmental interest, and that they leave open ample alternative channels for communication of the information.’” (quoting *Clark v. Cmty. for Creative NonViolence*, 468 U.S. 288, 293 (1984))).

72. 391 U.S. 367 (1968).

73. *See id.* at 382.

### 1. Content-Based and Content-Neutral Speech Restrictions

The Supreme Court has established several tests for determining whether government restrictions on speech are constitutional.<sup>74</sup> If government regulations categorize speech based on its content, the Court analyzes these regulations under strict scrutiny, which is extremely difficult to satisfy.<sup>75</sup> If the regulations are content neutral, such as those that regulate the time, place, or manner (TPM) of speech (“TPM regulations”),<sup>76</sup> or those focused on conduct with expressive qualities (“*O’Brien* regulations”), the Court uses two forms of intermediate scrutiny.<sup>77</sup> Under current jurisprudence, the Court will analyze the free speech claims under one of these three types of scrutiny for an AI to receive free speech protection. This section first reviews content-based regulations, then TPM regulations, and concludes with *O’Brien* regulations, while explaining the different levels of scrutiny that the Court applies to each type of restriction on speech.

Content-based restrictions of speech are subject to the most exacting test—strict scrutiny.<sup>78</sup> In this context, the Court has identified two aspects of content: subject matter<sup>79</sup> and viewpoint or ideology.<sup>80</sup> For a regulation to be upheld, it must promote a “compelling interest” using the “least restrictive means” or “narrowly drawn regulations.”<sup>81</sup> Strict scrutiny is difficult to overcome; “[o]nly one speech regulation has survived strict scrutiny in the Supreme Court.”<sup>82</sup>

For example, in *Reed v. Town of Gilbert*,<sup>83</sup> the Court applied strict scrutiny and struck down the town of Gilbert’s signage regulations.<sup>84</sup> The town had various codes regarding how outdoor signs may be posted, with different rules for various categories.<sup>85</sup> The Court treated the differentiation between

74. See Massaro & Norton, *supra* note 19, at 1186–87.

75. See Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445, 1451 (2013).

76. The classic example of a TPM regulation is a town’s restriction on the time that loud music can be played in a residential neighborhood. See, e.g., *Kovacs v. Cooper*, 336 U.S. 77, 78 (1949).

77. See Benjamin, *supra* note 75, at 1451.

78. See, e.g., *Sable Commc’ns of Cal., Inc. v. FCC*, 492 U.S. 115, 126 (1989) (“The Government may, however, regulate the content of constitutionally protected speech in order to promote a compelling interest if it chooses the least restrictive means to further the articulated interest.”).

79. See *id.* at 118 (invalidating a state statute that criminalized adult access to indecent phone messages).

80. See, e.g., *Rosenberger v. Rectors & Visitors of the Univ. of Va.*, 515 U.S. 819, 822–23 (1995) (invalidating a university plan that provided funding to secular newspaper publications but not to religious publications).

81. *Sable Commc’ns*, 492 U.S. at 126.

82. See Benjamin, *supra* note 75, at 1451 n.19 (referring to *Holder v. Humanitarian Law Project*, 561 U.S. 1, 7–8 (2010)); see also *Burson v. Freeman*, 504 U.S. 191, 193, 195 (1992) (plurality opinion) (applying strict scrutiny and upholding a Tennessee statute restricting electioneering near polling sites).

83. 576 U.S. 155 (2015).

84. See *id.* at 159.

85. See *id.* at 159–60. Specifically, the town regulated the size and placement of ideological signs, political signs, and signs advertising temporary events. *Id.*

the categories as a facially content-based regulation and applied strict scrutiny.<sup>86</sup> Justice Thomas, writing for the Court, noted that “speaker-based” laws are reviewed under strict scrutiny when the classification implicates a content preference.<sup>87</sup> The Court struck down the signage regulations, concluding that the law was not narrowly tailored because the distinctions were “hopelessly underinclusive.”<sup>88</sup>

Still, the Court has permitted content-based restrictions in certain circumstances. For example, a minor’s speech in school may be censored when it interferes with the school’s interests.<sup>89</sup> Further, speech that incites imminent harm,<sup>90</sup> obscenities,<sup>91</sup> and “fighting words”<sup>92</sup> are not protected. False statements are generally protected,<sup>93</sup> but there are limitations such as in cases of defamation.<sup>94</sup>

Unlike content-based regulations, TPM regulations target the time, place, or manner of speech but are neutral as to the subject matter or viewpoint.<sup>95</sup> For example, in *Ward v. Rock Against Racism*,<sup>96</sup> the Supreme Court upheld regulations regarding the use of a bandshell in Central Park in New York City.<sup>97</sup> The Court recognized that TPM restrictions are permissible when they “are justified without reference to the content of the regulated speech, . . . are narrowly tailored to serve a significant governmental interest,

---

86. *See id.* at 164–65.

87. *See id.* at 170 (“[W]e have insisted that ‘laws favoring some speakers over others demand strict scrutiny when the legislature’s speaker preference reflects a content preference.’ Thus, a content-based law that restricted the political speech of all corporations would not become content neutral just because it singled out corporations as a class of speakers.” (quoting *Turner Broad. v. FCC*, 512 U.S. 622, 658 (1994))).

88. *Id.* at 171 (finding that temporary directional signs are no more aesthetically displeasing or likely to cause threats to traffic safety than ideological or political ones).

89. *See, e.g.*, *Morse v. Frederick*, 551 U.S. 393, 397 (2007) (upholding school’s censorship of students’ speech, “BONG HiTS 4 JESUS,” displayed on a poster at a school-sponsored event).

90. *See, e.g.*, *Brandenburg v. Ohio*, 395 U.S. 444, 447 (1969) (“Constitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.”).

91. *See, e.g.*, *Miller v. California*, 413 U.S. 15, 24 (1973) (establishing that obscenities are “works which depict or describe sexual conduct” and may be prohibited under applicable state law).

92. *See, e.g.*, *Chaplinsky v. New Hampshire*, 315 U.S. 568, 572 (1942) (concluding that fighting words are “those [words] which by their very utterance inflict injury or tend to incite an immediate breach of the peace”).

93. *See, e.g.*, *United States v. Alvarez*, 567 U.S. 709, 715 (2012) (holding that the Stolen Valor Act of 2005, which criminalized misrepresentations about military decorations or medals, was unconstitutional).

94. *See, e.g.*, *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 352 (1974) (finding that states may define libel and defamation, so long as liability is not no-fault). *See generally* *N.Y. Times Co. v. Sullivan*, 376 U.S. 254 (1964) (establishing that a defamation suit by a public official regarding their official conduct requires proof of actual malice by the alleged defamer).

95. *See* *Ward v. Rock Against Racism*, 491 U.S. 781, 791 (1989).

96. 491 U.S. 781 (1989).

97. *See id.* at 784.

and . . . leave open ample alternative channels for communication of the information.”<sup>98</sup>

When determining whether a regulation is content neutral, “[t]he principal inquiry . . . is whether the government has adopted a regulation of speech because of disagreement with the message it conveys.”<sup>99</sup> Applying this test, the Court first considered whether New York City regulated the content of speech. New York City’s justification for its regulations was to retain the character of nearby passive recreation areas and not intrude into residential areas.<sup>100</sup> The Court found the justification to be neutral as to the content of the performances.<sup>101</sup>

Next, the Court addressed whether the regulations were narrowly tailored to serve a significant interest.<sup>102</sup> The Court recognized New York City’s “substantial interest in protecting its citizens from unwelcome noise.”<sup>103</sup> Regarding the next prong, the Court stated that a regulation is narrowly tailored “so long as the . . . regulation promotes a substantial government interest that would be achieved less effectively absent the regulation.”<sup>104</sup> In other words, the regulation must not be “substantially broader than necessary,” but the existence of a “less-speech-restrictive alternative” does not make regulation impermissible.<sup>105</sup> The Court found that the city’s regulations were direct and effective (i.e., not substantially broader than necessary), and that the interest in limiting volume would “have been served less well” without the regulations.<sup>106</sup>

The Court has reaffirmed the *Ward* method of analyzing TPM regulations.<sup>107</sup> In response to protesters at abortion clinics, Colorado established an eight-foot buffer zone between individuals entering health-care facilities and protesters who interacted with those entering the facility.<sup>108</sup> The Court upheld the law, characterizing the regulation as “a

---

98. *Id.* at 791 (quoting *Clark v. Cmty. for Creative NonViolence*, 468 U.S. 288, 293 (1984)).

99. *Id.*

100. *See id.* at 792.

101. *See id.* The Court also found a second justification—ensuring adequate sound quality—to be content neutral. *Id.* at 792–93.

102. *See id.* at 796.

103. *Id.*

104. *Id.* at 799.

105. *Id.* at 800. In other cases involving door-to-door solicitation or canvassing, the Court has considered whether a less restrictive alternative exists when analyzing whether the government’s regulation satisfied the narrowly tailored prong. *See, e.g., Watchtower Bible & Tract Soc’y of N.Y., Inc. v. Village of Stratton*, 536 U.S. 150, 156 (2002) (striking down a regulation banning canvassing because the posting of signs saying “No Solicitation” and the ability of the homeowner to turn the solicitor away at the door was a less restrictive means); *Martin v. City of Struthers*, 319 U.S. 141, 146–49 (1943) (invalidating a regulation banning canvassing as an unconstitutional limit on free speech because it is the individual homeowner’s choice, not the government’s, whether to turn the canvasser away).

106. *Ward*, 491 U.S. at 800. The Court also considered the availability of alternative channels for communication but found that the regulations had “no effect on the quantity or content of that expression beyond regulating the extent of amplification.” *See id.* at 802.

107. *See, e.g., Hill v. Colorado*, 530 U.S. 703, 710–11 (2000).

108. *See id.* at 707–08.

minor place restriction.”<sup>109</sup> Importantly, the Court noted that an otherwise constitutional statute does not become unconstitutional if it is only applied to a specific location.<sup>110</sup>

In addition to TPM regulations, the Court has identified a second category of content-neutral regulations: those that address expressive conduct.<sup>111</sup> If a regulation, facially and through the interests that it seeks to promote, targets expressive conduct, the Court applies the *O’Brien* test.<sup>112</sup> For conduct to have sufficient expressive qualities to warrant First Amendment protection, the speaker must have intended to convey a message that was reasonably likely to be understood by those who viewed the conduct.<sup>113</sup>

*O’Brien* concerned a provision of the Military Selective Service Act of 1967<sup>114</sup> that criminalized the destruction of draft cards.<sup>115</sup> On March 31, 1966, David O’Brien and three others burned their draft cards in front of a crowd on the steps of a Boston courthouse.<sup>116</sup> The *O’Brien* Court held that a regulation survives First Amendment scrutiny even if it imposes “incidental limitations” on speech when the regulation (1) is “within the constitutional power of the Government,” (2) “furthers an important or substantial governmental interest” that (3) is “unrelated to the suppression of free expression” and (4) prohibits no more speech than is essential.<sup>117</sup> The Court held that the draft card regulation satisfied the test and concluded that the law was content neutral because the focus was on the “noncommunicative” aspect of draft card destruction—i.e., preventing disruption of the administration of the draft.<sup>118</sup>

Because content-based restrictions on speech are subject to strict scrutiny, courts frequently invalidate such regulations. Therefore, for a government restriction on speech to survive, in most cases, it must be content neutral.<sup>119</sup> The tests for content-based and content-neutral speech regulations are important to consider when thinking about whether AI possesses free speech rights and about government’s decision whether to regulate this speech, as the level of scrutiny applied by courts may be dispositive.

---

109. *Id.* at 723.

110. *Id.* at 724.

111. *See, e.g.*, *United States v. O’Brien*, 391 U.S. 367, 382 (1968).

112. *See Texas v. Johnson*, 491 U.S. 397, 407 (1989). If the regulation focuses on the content of the expression, then it is content based. *See id.* If the regulation focuses on conduct without any expressive qualities, then the regulation is subject to rational basis review. *Id.*; *see also Massaro & Norton, supra* note 19, at 1187.

113. *See Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Bos., Inc.*, 515 U.S. 557, 569 (1995); *see also Spence v. Washington*, 418 U.S. 405, 410–11 (1974) (per curiam).

114. Pub. L. No. 90-40, 81 Stat. 100 (codified as amended in scattered sections of the U.S.C.).

115. *O’Brien*, 391 U.S. at 370.

116. *Id.* at 369.

117. *Id.* at 377.

118. *Id.* at 382.

119. *See Burson v. Freeman*, 504 U.S. 191, 193, 195 (1992) (plurality opinion); *see also Benjamin, supra* note 75, at 1451 n.19 (citing *Holder v. Humanitarian L. Project*, 561 U.S. 1, 7–8 (2010)).

## 2. Corporate Speech

In addition to characterizing regulated speech as content-based or content neutral, the Court has also addressed whether the First Amendment's speech protections extend to nonhuman actors, such as corporate entities. The Court's decision in *Citizens United v. FEC*, which involved the speech rights of a corporation, had a significant effect on free speech doctrine.<sup>120</sup> Most commentators agree that *Citizens United* gave corporate persons the same constitutional speech rights that natural persons have.<sup>121</sup> Thus, *Citizens United* will inevitably be relevant to any discussion of nonhuman speech, including the speech rights of AI.

*Citizens United* involved a documentary film produced by the nonprofit organization Citizens United, which criticized then presidential candidate Hillary Clinton.<sup>122</sup> The organization challenged a federal campaign-finance statute that regulated aspects of campaigning, such as by preventing "corporations and unions from using their general treasury funds to make independent expenditures for speech defined as an 'electioneering communication' or for speech expressly advocating the election or defeat of a candidate."<sup>123</sup> Citizens United argued that "[the statute's] ban on corporate-funded independent expenditures" and the statute's "disclaimer and disclosure requirements" were unconstitutional.<sup>124</sup> Justice Anthony Kennedy, writing for the majority, framed the issue as determining the constitutionality of federal laws regulating political speech based on the speaker's corporate identity.<sup>125</sup>

The Court first established that, because the regulations targeted political speech, they were subject to strict scrutiny, meaning the government had to prove that the regulation "furthers a compelling interest and is narrowly tailored to achieve that interest."<sup>126</sup> The Court noted that there are "certain governmental functions that cannot operate without some restrictions."<sup>127</sup> The Court distinguished political speech, stating that "it is inherent in the

---

120. See, e.g., 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, 2496 (2017); Brandon L. Garrett, *The Constitutional Standing of Corporations*, 163 U. PA. L. REV. 95, 153 (2014); Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1502 (2013).

121. See, e.g., Garrett, *supra* note 120, at 153; Massaro et al., *supra* note 120, at 2496; Wu, *supra* note 120, at 1502.

122. See *Citizens United v. FEC*, 558 U.S. 310, 319–20 (2010).

123. *Id.* at 310 (quoting 2 U.S.C. § 441(b)).

124. *Id.* at 321.

125. See *id.* at 318–19.

126. See *id.* at 340 (quoting *FEC v. Wis. Right to Life, Inc.*, 551 U.S. 449, 464 (2007)).

127. *Id.* at 341 (citing *Bethel Sch. Dist. No. 403 v. Fraser*, 478 U.S. 675, 683 (1986)); see also *Jones v. N.C. Prisoners' Lab. Union, Inc.*, 433 U.S. 119, 129 (1977) (furthering "the legitimate penological objectives of the corrections system"); *Parker v. Levy*, 417 U.S. 733, 759 (1974) (ensuring "the capacity of the Government to discharge its [military] responsibilities"); *Civ. Serv. Comm'n v. Letter Carriers*, 413 U.S. 548, 557 (1973) ("[F]ederal service should depend upon meritorious performance rather than political service.").

nature of the political process that voters must be free to obtain information from diverse sources in order to determine how to cast their votes.”<sup>128</sup>

The Federal Election Commission (FEC) argued that the government could regulate political speech under the “anti-distortion” argument.<sup>129</sup> In *Austin v. Michigan Chamber of Commerce*,<sup>130</sup> the Court held that large corporations would overwhelm public discourse and inaccurately represent to the public (i.e., distort) the actual amount of support for an idea.<sup>131</sup> In *Citizens United*, the Court rejected the anti-distortion argument.<sup>132</sup> The Court stated that adopting the argument would unconstitutionally limit the “voices” of corporations to advise “voters on which persons or entities are hostile to their interests.”<sup>133</sup>

Next, the Court turned to a second justification previously used to uphold campaign finance regulations: preventing corruption or the appearance of corruption.<sup>134</sup> The Court narrowly interpreted its cases to limit the interest to “*quid pro quo* corruption” only.<sup>135</sup> The Court stated that “[t]he appearance of [corporate] influence or access, furthermore, will not cause the electorate to lose faith in our democracy.”<sup>136</sup>

Ultimately, the Court struck down most of the campaign finance regulations as impermissible restrictions on political speech.<sup>137</sup> The Court observed that “[r]apid changes in technology—and the creative dynamic inherent in the concept of free expression—counsel against upholding a law that restricts political speech in certain media or by certain speakers.”<sup>138</sup>

Justice Antonin Scalia penned a concurrence to reinforce the theory underlying the majority’s decision.<sup>139</sup> He stated: “The [First] Amendment is written in terms of ‘speech,’ not speakers. Its text offers no foothold for excluding any category of speaker, from single individuals to partnerships of individuals, to unincorporated associations of individuals, to incorporated associations of individuals.”<sup>140</sup>

---

128. *Citizens United*, 558 U.S. at 341.

129. *See id.* at 348.

130. 494 U.S. 652 (1990).

131. *See id.* at 660 (finding a compelling governmental interest in preventing “the corrosive and distorting effects of immense aggregations of wealth that are accumulated with the help of the corporate form and that have little or no correlation to the public’s support for the corporation’s political ideas”).

132. *Citizens United*, 558 U.S. at 354.

133. *Id.*

134. *See id.* at 356.

135. *Id.* at 359.

136. *Id.* at 360. The Court also briefly dismissed a third justification: that the regulation protects “dissenting shareholders from being compelled to fund corporate political speech.” *Id.* at 361. The Court simply stated that “[t]he First Amendment does not allow that power.” *Id.*

137. *See id.* at 364. In striking down a majority of the campaign finance regulations, the Court rejected its ruling in *Austin*. *See id.* at 365.

138. *See id.* at 364. The Court still upheld the statute’s disclosure requirements. *See id.* at 366–67.

139. *See id.* at 392–93 (Scalia, J., concurring).

140. *Id.*

*Citizens United* established that regulation of a particular category of speakers generally warrants the highest level of scrutiny.<sup>141</sup> Additionally, when nonhuman speech is relevantly similar to human speech, there are few barriers to nonhuman speakers receiving protections comparable to those given to human speakers.<sup>142</sup> Because AI systems are nonhuman speakers, it is important to consider the Court's decision in *Citizens United* when analyzing whether AI has free speech rights.<sup>143</sup>

### 3. Theoretical Justifications for Free Speech

This section reviews the most prominent legal and philosophical theories justifying the right to free speech.<sup>144</sup> There are four main theories: the “negative” theory, the “marketplace of ideas” theory, the “democracy” theory, and the “autonomy” theory.<sup>145</sup> Since the Supreme Court often considers these theories when deciding free speech cases,<sup>146</sup> they are also relevant for determining whether AI qualify for free speech protection.

The first theory is called the “negative” theory. Its basic premise is that free speech operates as a check on government power, and it does not focus on any specific value of the activity.<sup>147</sup> In other words, speech is simply a means to an end, with the end being limiting unwarranted government intrusion on speech.<sup>148</sup> Negative theory is the most instrumental of the four—it implies that there is no inherent value in speech, since both the speaker and listener are somewhat irrelevant to its justification.<sup>149</sup> This theory has two primary concerns: one regarding government's lack of institutional competence in regulating speech and the other regarding “the

---

141. See, e.g., *Reed v. Town of Gilbert*, 576 U.S. 155, 170–71 (2015) (citing *Citizens United v. FEC*, 558 U.S. 310, 340–41 (2010)).

142. Commentators generally agree that the Court's holding in *Citizens United* has greatly expanded free speech protections. See, e.g., Massaro et al., *supra* note 120, at 2496; Garrett, *supra* note 120, at 153; Wu, *supra* note 120, at 1502.

143. The Court has also acknowledged that certain corporations may also receive speech protections under the Religious Freedom Restoration Act of 1993. See generally *Burwell v. Hobby Lobby Stores, Inc.*, 573 U.S. 682 (2014).

144. See FREDERICK SCHAUER, *FREE SPEECH: A PHILOSOPHICAL ENQUIRY* ix (1982) (“Unless we can get clear about the philosophical underpinnings of a political principle, we can hardly navigate successfully through the waters of specific application of that principle.”).

145. Each theory has various sub-theories and nuanced enumerations. See *id.* However, this Note provides a broad overview of each theory to highlight the underpinnings of free speech doctrine in the United States and does not explore the various enumerations of each theory.

146. See Benjamin, *supra* note 75, at 1455.

147. See Massaro et al., *supra* note 120, at 2492; see also Benjamin, *supra* note 75, at 1454; SCHAUER, *supra* note 144, at 80.

148. See Massaro et al., *supra* note 120, at 2492; see also SCHAUER, *supra* note 144, at 47, 81.

149. See SCHAUER, *supra* note 144, at 47. In this context, “instrumental” means that speech does not have any value in and of itself. See *id.* In other words, speech is a mere means to a valued end without relevant independent moral worth. See *id.* This contrasts with a view holding that speech has inherent, and therefore not merely instrumental, value for either the listener or speaker, which is why the distinction is mentioned in text.

government's censorial motives (i.e., its malevolence, its self-interest, or at the very least its paternalism)."<sup>150</sup>

Next is the "marketplace of ideas" or "search for truth" theory. This theory is most commonly invoked by the Supreme Court.<sup>151</sup> The marketplace theory is instrumental—it proposes that speech serves a functional role as a means of communicating ideas.<sup>152</sup> The theory assumes that free communication of ideas will lead to increased truth or knowledge.<sup>153</sup> This increased truth and knowledge is useful because it has "value to listeners' enlightenment."<sup>154</sup> Analogous to the theory of free market economics, the marketplace theory argues that, with minimal regulation, the best and most accurate information will, through speech, be discovered.<sup>155</sup> Therefore, listeners will have the best access to truth and knowledge when there are strong free speech protections and less government regulation of speech.<sup>156</sup>

Another instrumental theory of free speech, the "democracy" theory, is rooted in promoting democracy and self-governance.<sup>157</sup> This theory focuses on free speech's value to members of the sovereign electorate.<sup>158</sup> Generally, the democracy theory argues that free speech protections allow the sovereign electorate to exercise its power in the democratic process.<sup>159</sup> The electorate may use free speech protections to argue issues and share information<sup>160</sup> in public discourse or to criticize government officials and hold them accountable.<sup>161</sup> As with the preceding theories, under the democracy theory, speech is a tool used to achieve an end—here, self-governance of a sovereign electorate.<sup>162</sup>

The final theory justifying free speech is focused on autonomy and self-realization.<sup>163</sup> This theory is not instrumental—it assumes that speech has value on its own, not as a tool to achieve another goal.<sup>164</sup> A popular enumeration of this theory is that speech—and, more broadly, expression—is a necessary trait for personhood (having consciousness, personal identity, plans, goals, reason, etc.), and personhood is necessary for fundamental

---

150. Massaro et al., *supra* note 120, at 2492; *see also* SCHAUER, *supra* note 144, at 81.

151. *See* Benjamin, *supra* note 75, at 1455; *see also* Massaro & Norton, *supra* note 19, at 1178; SCHAUER, *supra* note 144, at 15 ("[T]he predominant and most persevering has been the argument that free speech is particularly valuable because it leads to the discovery of truth.").

152. *See* SCHAUER, *supra* note 144, at 47.

153. *See id.*

154. *See* Massaro & Norton, *supra* note 19, at 1178. Therefore, increased truth or knowledge is the valued end, and speech, as a form of communication, is merely the means to achieving that end. *See id.*

155. *See id.*; *see also* SCHAUER, *supra* note 144, at 15–16, 19–20.

156. *See* Massaro & Norton, *supra* note 19, at 1178; *see also* SCHAUER, *supra* note 144, at 38.

157. *See* Massaro & Norton, *supra* note 19, at 1176.

158. *See* SCHAUER, *supra* note 144, at 47.

159. *See id.* at 36.

160. *See* Massaro & Norton, *supra* note 19, at 1176; SCHAUER, *supra* note 144, at 38.

161. *See* Massaro & Norton, *supra* note 19, at 1167; SCHAUER, *supra* note 144, at 36.

162. *See* SCHAUER, *supra* note 144, at 36, 47.

163. *See* Massaro & Norton, *supra* note 19, at 1178.

164. *See* SCHAUER, *supra* note 144, at 48.

rights.<sup>165</sup> Others view expression as an integral part of self-realization.<sup>166</sup> Therefore, the autonomy theory argues that strong free speech protections preserve both a fundamental right and a necessary condition for full moral personhood.<sup>167</sup>

Each theory has its own literature and debates.<sup>168</sup> While there is no overwhelmingly favored theory in jurisprudence and academia, the marketplace theory is the best-known theory, and courts, especially the Supreme Court, frequently address it when analyzing free speech claims under the First Amendment.<sup>169</sup>

### C. Theories of Cognition

As previously discussed, the Supreme Court often uses legal theories such as the marketplace of ideas when analyzing government regulations on protected speech.<sup>170</sup> For nonhuman actors to receive free speech rights, the necessary characteristics of protected human speech will need to be identified. Therefore, understanding the philosophical theories regarding intelligence—the process of human cognition—is necessary before discussing the potential application of First Amendment protection to AI. Part I.C.1 outlines the debate over the definition and meaning of human “intelligence.” Part I.C.2 reviews recent psychological literature regarding the way in which humans make ethical and moral judgments when producing speech, and relates this literature to the larger philosophical debate.

#### 1. Philosophical Theories of Human Intelligence

The philosophical debate surrounding the definition of human intelligence has two primary views: the “output only” position<sup>171</sup> and the “missing something” position.<sup>172</sup> Both Professor Eugene Volokh and Alan Turing have advanced the output-only position.<sup>173</sup> They argue that to determine whether a program is intelligent, the internal mechanics of the thinking process are irrelevant—only the results matter.<sup>174</sup> Put differently, if the

---

165. See Massaro & Norton, *supra* note 19, at 1179 (citing JOEL FEINBERG, FREEDOM AND FULFILLMENT: PHILOSOPHICAL ESSAYS 52 (1992)).

166. See SCHAUER, *supra* note 144, at 48.

167. See *id.*; see also Massaro & Norton, *supra* note 19, at 1179 (citing JOEL FEINBERG, FREEDOM AND FULFILLMENT: PHILOSOPHICAL ESSAYS 52 (1992)). See generally C. EDWIN BAKER, HUMAN LIBERTY AND FREEDOM OF SPEECH 3 (1989).

168. See, e.g., SCHAUER, *supra* note 144, at 16.

169. See Benjamin, *supra* note 75, at 1455; Massaro & Norton, *supra* note 19, at 1178; see also SCHAUER, *supra* note 144, at 15.

170. See *supra* Part I.B.3.

171. See Volokh, *supra* note 6, at 1137–38.

172. See Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231, 1262 (1992).

173. See A.M. Turing, *Computing Machinery and Intelligence*, 49 MIND 433, 433–34 (1950); Volokh, *supra* note 6, at 1138 (“Whatever goes on under the hood, thinking is as thinking does.”).

174. See Turing, *supra* note 173, at 433; Volokh, *supra* note 6, at 1138.

results appear intelligent, then the program is intelligent.<sup>175</sup> Turing created a test that he termed “The Imitation Game,” to determine intelligence. Turing’s test involves three “players”<sup>176</sup>: one human, one machine, and one tester.<sup>177</sup> Each player is in a separate room and cannot see the others.<sup>178</sup> The tester writes a question to the others, and the others write responses back; the tester does not know which participant is the machine.<sup>179</sup> The goal of the game is for the tester to accurately guess which is the machine and which is the human based solely on their responses to the question.<sup>180</sup> Turing posited that if a machine can reliably trick the tester into guessing wrong, then the machine is intelligent, irrespective of the process used to reach the result.<sup>181</sup> Under the output-only theory, there is a compelling argument that AI may possess aspects of human intelligence.<sup>182</sup>

Other scholars, such as John Searle, advance the missing-something position of human intelligence. This position argues that humans have unique traits that other beings or things, such as animals or computer programs, lack.<sup>183</sup> For AI, the position is that there are certain aspects of human cognition that AI lack, and therefore AI cannot have true intelligence.<sup>184</sup> Searle, in particular, is famous for his “Chinese Room” thought experiment.<sup>185</sup> The experiment describes a room with two letter slots, one for input and one for output.<sup>186</sup> An individual outside the room writes a letter containing a story in English and a question about the story and then pushes the letter into the room.<sup>187</sup> Inside the room, an English-speaking human reads the letter, understands it, and writes a response.<sup>188</sup> The human then pushes the response out of the output letter slot.<sup>189</sup> The human’s output states the correct response to the question about the story.<sup>190</sup>

Next, the individual outside the room inputs the same story and question, but this time, the story is written in Chinese characters.<sup>191</sup> The human inside

---

175. Volokh, *supra* note 6, at 1137–38.

176. *See* Turing, *supra* note 173, at 433–34.

177. *See id.*

178. *See id.*

179. *See id.*

180. *See id.*

181. *See id.*

182. *See infra* Part II.B.

183. *See* Solum, *supra* note 172, at 1262.

184. *See* John R. Searle, *Minds, Brains, and Programs*, 3 BEHAV. & BRAIN SCIS. 417, 417 (1980).

185. *See id.* at 417–18.

186. *See id.*

187. *See id.* at 418. Searle used the following example story and question: “A man went into a restaurant and ordered a hamburger. When the hamburger arrived it was burned to a crisp, and the man stormed out of the restaurant angrily, without paying for the hamburger or leaving a tip. . . . Did the man eat the hamburger?” *See id.* Searle assumed that someone reading the narrative would understand it and consciously formulate a response to the question. *See id.*

188. *See id.*

189. *See id.*

190. *See id.* at 417 (“No, he did not [eat the hamburger].”).

191. *See id.* at 417–18.

the room does not know how to read the characters.<sup>192</sup> Instead, the human pulls out a large guidebook that has, in English, information on how to correspond English words into every possible Chinese character or string of characters.<sup>193</sup> The human, using the guidebook, writes a correct response in Chinese characters and pushes it out of the room.<sup>194</sup> The human did not translate the story into English or understand<sup>195</sup> the story at all but instead used preset instructions to transcribe a response.<sup>196</sup>

Searle's thought experiment illustrates that an AI may appear to be intelligent—by passing the Turing test and producing comparable output—but, “under the hood,” lacks important parts of human cognition<sup>197</sup>—specifically, a deep understanding of the subject necessary for humanlike intelligence.<sup>198</sup> Searle argues that programs such as AI merely respond to stimuli.<sup>199</sup> In other words, AI does not have consciousness, intentionality, or reasoning.<sup>200</sup> When determining whether AI possesses intelligence comparable to human speakers, courts are likely to consider the output-only and missing-something theories of intelligence because they are the most relevant to analyzing AI.

## 2. The Role of Emotion in Human Decision-Making

In addition to theories of intelligence, psychological research on how the human brain functions is relevant to the philosophical debate on AI. The process that humans use to make moral and ethical judgments is of particular importance when analyzing whether AI is eligible for First Amendment protection. Although psychological research<sup>201</sup> is more relevant for analyzing the missing-something theory,<sup>202</sup> it is nonetheless helpful to consider it as part of the larger debate surrounding AI and free speech rights.

---

192. *See id.* at 417.

193. *See id.* at 418.

194. *See id.*

195. “Understand” in this context means consciously and intentionally identifying the subject matter and reasoning a response. *See id.* When the story was written in English, Searle presumed the human understood the subject matter and crafted a response based on that understanding. *See id.*

196. *See id.*

197. *See id.*

198. *See id.*

199. *See id.*

200. *See id.*

201. This Note provides a brief overview of psychological theories in connection with the debate surrounding AI free speech rights but does not provide extensive background on the development of these psychological theories. For additional information regarding these psychological theories, see generally HANDBOOK OF INTUITION RESEARCH (Marta Sinclair ed., 2d ed. 2013).

202. As previously discussed, under the missing-something theory, nonhumans may appear to be intelligent but lack certain cognitive traits necessary for humanlike intelligence. *See Searle, supra* note 184, at 417–18. As applied to AI, John Searle and other scholars argue that AI lacks certain cognitive functions. *See id.* To accurately identify what traits an AI is missing, proponents of the missing-something theory must first understand the traits that are relevant to human decision-making. *See Solum, supra* note 172, at 1262. The psychological research

One predominant school of psychological thought is the “rationalist model.”<sup>203</sup> This model argues that when making a moral judgment, an individual acts in a conscious, intentional, and rational manner.<sup>204</sup> Essentially, an individual is “a judge, weighing issues of harm, rights, justice, and fairness, before passing judgment.”<sup>205</sup> Centuries of philosophers have used this model, including Immanuel Kant.<sup>206</sup> Although rationalist theory was never universally adopted, it significantly affected psychology in the latter half of the twentieth century and remains very influential.<sup>207</sup>

Proponents of the “ethical intuition” model disagree with the rationalists. Ethical intuition is not a novel concept; it was recognized by the founding fathers.<sup>208</sup> Professor Jonathan Haidt, a “social intuitionist,” argues that when making a moral judgment, humans make snap emotional decisions (intuitions) that are justified (to one’s self and expressed to others) by a post hoc rational cognitive process.<sup>209</sup> Essentially, while rationalists argue that humans are judges weighing facts and deciding cases, social intuitionists argue that humans are attorneys arguing for their side.<sup>210</sup> Professor Haidt does not discount reason but disagrees that it is the primary causal factor in human decision-making.<sup>211</sup> His social intuitionist theory combines reason, emotion, and intuition, and emphasizes social influence as an explanation for moral judgments.<sup>212</sup>

Professor Haidt proposes a model with six “links,” or stages of and influences on, cognition for moral judgments.<sup>213</sup> First, the individual makes an “intuitive judgment.”<sup>214</sup> This judgment is sudden; there is no conscious decision-making.<sup>215</sup> Professor Haidt argues that intuitive judgment is predominantly based on social forces.<sup>216</sup> At this point, an individual has already decided the correct response to the issue presented.<sup>217</sup> The second stage is “post hoc reasoning,” during which the individual consciously<sup>218</sup> decides what their explanation—to themselves and to others—for the

---

described in this section provides context for understanding the role of emotion in human decision-making.

203. See Haidt, *supra* note 28, at 815.

204. See *id.*

205. See *id.*

206. See *id.* at 816. However, the model was not universally adopted. Opponents of rationalist theory, including David Hume, argued that reason alone could not be the sole cognitive function but must be partnered with moral sentiments. See *id.*

207. See *id.*

208. See THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776) (“We hold these truths to be self-evident . . .”).

209. See Haidt, *supra* note 28, at 815.

210. See *id.*

211. See *id.* at 816.

212. See *id.* at 828.

213. See *id.* at 818.

214. See *id.*

215. See *id.* at 817.

216. See *id.*

217. See *id.*

218. Professor Haidt defines conscious decision-making as “intentional, effortful, controllable,” a process during which “the reasoner is aware that it is going on.” See *id.* at 818.

judgment will be.<sup>219</sup> This stage—the first conscious one—may create the illusion that the judgment was created by internal reasoning.<sup>220</sup> The third stage is “reasoned persuasion,” during which the individual explains their justification for the judgment to others.<sup>221</sup> The fourth stage is “social persuasion,” during which “the mere fact that friends, allies, and acquaintances have made a moral judgment exerts a direct influence on others, even if no reasoned persuasion is used.”<sup>222</sup> Professor Haidt argues that these four stages—intuitive judgment, post hoc reasoning, reasoned persuasion, and social persuasion—operate in a loop: the latter two influence the former two to produce a judgment.<sup>223</sup>

Professor Haidt argues that a fifth link, “reasoned judgment,” may also influence judgments, but he believes it to be rare.<sup>224</sup> This link requires an individual “by sheer force of logic” to override an intuition.<sup>225</sup> Professor Haidt similarly argues that a sixth link, “private reflection,” may change intuitions, but is also rare.<sup>226</sup> In other words, for most people, moral judgments are produced by unconsciously and intuitively “reading the room” for social clues as to the correct decision, followed by a conscious rationalization.<sup>227</sup> One need not fully subscribe to Professor Haidt’s model to acknowledge that at least some human speech and action is the product of unconscious emotions.<sup>228</sup>

Although few legal scholars have explored the relationship between human intelligence and AI, the relationship is an important part of the discussion on whether AI is eligible for free speech protections. When facing novel free speech issues, the Supreme Court has closely relied on philosophical and legal theories of free speech, as evidenced by its decision in *Citizens United*.<sup>229</sup> Thus, analyzing the question of whether to expand free speech protections to AI requires an examination of the philosophical theories regarding AI and cognition.

---

219. *See id.* at 819.

220. *See id.* at 822.

221. *See id.*

222. *See id.*

223. *See id.*

224. *See id.*

225. *See id.*

226. *See id.*

227. *See id.* (“When asked why he enjoyed a party, a person turns first to his cultural knowledge about why people enjoy parties, chooses a reason, and then searches for evidence that the reason was applicable.”).

228. For example, this idea is also exemplified by research concerning implicit bias—the unconscious tendency for individuals to hold viewpoints that influence their actions without awareness of these views or their influence. *See generally* B. Keith Payne & C. Daryl Cameron, *Divided Minds, Divided Morals: How Implicit Social Cognition Underpins and Undermines Our Sense of Social Justice*, in *HANDBOOK OF IMPLICIT SOCIAL COGNITION: MEASUREMENT, THEORY, AND APPLICATIONS* 445 (Bertram Gawronski & B. Keith Payne eds., 2010).

229. *See supra* Part I.B.2.

## II. ISSUES IN THE APPLICATION OF FIRST AMENDMENT RIGHTS TO AI

While the Court has not entertained a case involving an AI making a serious claim for free speech protection under the First Amendment, recent technological advancements in weak AI signal that these claims may be arising in the near future. Because there is no First Amendment jurisprudence specifically involving AI, most of the relevant legal and philosophical debate regarding free speech rights for AI has occurred in academic publications. Nonetheless, each discipline contains a robust discussion of various issues relating to the application of free speech to weak AI.

This part explores the potential challenges that may arise in connection with extending free speech protection to AI. Part II.A discusses various doctrinal and theoretical issues raised by an AI's claim for free speech protection under the First Amendment. Part II.B further elaborates on the output-only and missing-something theories of human intelligence and explains their application to AI and speech rights.

### A. Doctrinal and Theoretical Issues for Nonhuman Speech

The Supreme Court's First Amendment jurisprudence seems to generally support extending free speech protections to AI. Most commentators interpret *Citizens United* to extend First Amendment protections broadly to nonhumans, and thus arguably to an AI, a nonhuman speaker.<sup>230</sup> As previously discussed, *Citizens United* involved corporate speakers challenging federal campaign-finance restrictions based on a free speech claim.<sup>231</sup> The Court held that, essentially, corporate speakers should be held to the same, or close to the same, standards as natural persons.<sup>232</sup> The Court's decision was significant because it expanded the ability for corporations to participate in and influence elections.<sup>233</sup>

Since the Court in *Citizens United* expanded free speech rights to nonhumans, the decision provides an important framework for expanding free speech rights to AI. Language in the majority opinion and in Justice Scalia's concurrence indicates that the Court would be reticent to immediately discount other nonhuman speakers as not deserving of protection.<sup>234</sup> Notably, Justice Scalia took a strong stance, stating that the identity of the speaker should be irrelevant for protection because "[t]he [First] Amendment is written in terms of 'speech,' not speakers."<sup>235</sup> Although the majority does not explicitly take such a strong stance, the identity or qualities of the speaker were nonetheless largely irrelevant as to whether the corporations were entitled to free speech protection.<sup>236</sup>

---

230. See, e.g., Benjamin, *supra* note 75, at 1447; Massaro & Norton, *supra* note 19, at 1185.

231. See *Citizens United v. FEC*, 558 U.S. 310, 310, 321 (2010).

232. See *id.* at 364; *supra* Part I.B.2.

233. See *Citizens United*, 558 U.S. at 364.

234. See *id.* at 392–93 (Scalia, J., concurring).

235. See *id.* at 392.

236. See *id.* at 341 (majority opinion).

Further, the Court established that it would carefully consider restrictions on speech in connection with new technology: “Rapid changes in technology—and the creative dynamic inherent in the concept of free expression—counsel against upholding a law that restricts political speech in certain media or by certain speakers.”<sup>237</sup> These principles suggest that the Court would be open to the possibility of expanding free speech protections to other nonhuman actors like AI. Additionally, in other contexts, the Court has been reluctant to uphold regulations on nontraditional content and forms of expression.<sup>238</sup> When assessing a novel method of expression that is comparable to other methods, the Court tends to use the standard analytical framework that is applied to all free speech regulations; in other words, the Court applies a principle of equivalence.<sup>239</sup>

Although *Citizens United* generally supports the expansion of free speech rights to AI, there is a potential problem with the Court focusing on corporations as *associations of humans*, as opposed to standalone legal entities. In his concurrence, Justice Scalia listed speakers that should receive equivalent free speech protections: from “single individuals to partnerships of individuals, to unincorporated associations of individuals, to incorporated associations of individuals.”<sup>240</sup> As applied to AI, the Court may determine that an AI is eligible for free speech protection only if it contains some form of human input or presence.<sup>241</sup> In other words, the Court may apply its holding in *Citizens United* to AI but limit its application. This approach would provide a bridge between fully rejecting and fully protecting an AI’s speech.<sup>242</sup>

While the Court’s decision in *Citizens United* greatly expanded speech protections for nonhuman actors, some legal scholars doubt that the Supreme Court will extend its decision to AI. Professor Tim Wu argues that many AI will not receive speech protections because of a “de facto functionality

---

237. *See id.* at 364.

238. *See, e.g., Sorrell v. IMS Health Inc.*, 564 U.S. 552, 557 (2011) (extending free speech protection to the sale of prescription records for drug marketing); *Tinker v. Des Moines Indep. Cmty. Sch. Dist.*, 393 U.S. 503, 514 (1969) (extending free speech protection to students symbolically wearing black armbands in an anti-war protest); *see also Massaro & Norton, supra* note 19, at 1186.

239. *See, e.g., Sorrell*, 564 U.S. at 570 (comparing prescriber-identifying information sold for drug marketing to information provided by other sources such as a beer bottle label). Still, in the case of AI, there is a possibility that the Court may uphold more extensive regulations for AI speech due to a categorical difference between AI speech and other methods. *See supra* Part I.B.1 (explaining the Supreme Court’s treatment of different types of speech, including hate speech and obscenities). As discussed earlier, AI speech may be different from human or other types of speech due to issues of scalability and complexity. *See supra* Part I.A.

240. *See Citizens United*, 558 U.S. at 392–93 (Scalia, J., concurring).

241. *See Benjamin, supra* note 75, at 1479; *see also Wu, supra* note 120, at 1497 (arguing that the more human involvement in an AI program, the more likely courts will be willing to extend First Amendment protections to the AI).

242. If the Supreme Court concludes that an AI is eligible for free speech protection only if it contains human input, the next logical question concerns the amount of human input necessary to warrant free speech protection. Although AI operate independently, both the AI’s programmed goals and the design of the system involve human input. *See supra* Part I.A.

doctrine” present in many of the Court’s First Amendment cases.<sup>243</sup> Professor Wu posits that current First Amendment doctrine already excludes certain speech from protection in two ways.<sup>244</sup> The first is focused on who is involved in the speech, in that protection is not given to “carrier/conduits—actors who handle, transform, or process information, but whose relationship with speech or information is ultimately functional.”<sup>245</sup> The second focuses on “tools—works whose use of information is purely functional, such as navigational charts, court filings, or contracts.”<sup>246</sup> Professor Wu states that courts may fear a misuse of the strong protections afforded to speech by opportunistic attorneys or by clients with motivations contrary to the justifications underlying free speech protections.<sup>247</sup>

As applied to AI, Professor Wu contends that, despite the Court’s holding in *Citizens United*, the de facto functionality doctrine will exclude at least some AI from protection.<sup>248</sup> Professor Wu argues that AI will “handle or transform speech, but [not be] a speaker”<sup>249</sup> and will “perform[] some task other than the communication of ideas.”<sup>250</sup> For example, Professor Wu argues that a Google search is not protected because the search engine itself is functional—it is a tool for indexing results.<sup>251</sup> Additionally, Professor Wu argues that Google is a “carrier/conduit” of information—Google lacks legal responsibility for the content and does not curate content as compared to a newspaper.<sup>252</sup> On the other hand, Professor Wu notes that when a program effectively inherits traits or opinions from its creator, it is more likely to be viewed as more than a tool, and therefore its speech will be protected under the First Amendment.<sup>253</sup> The de facto functionality doctrine may bar free speech protections for some AI, but does not eliminate them for an AI that can emulate the relevant aspects of human cognition.<sup>254</sup> The Court’s

---

243. See Wu, *supra* note 120, at 1497.

244. See *id.*

245. See *id.* (“Definitive examples are Federal Express or the telephone company, common carriers to whom the law does not grant speech rights. Those who merely carry information from place to place (courier services) generally don’t enjoy First Amendment protection, while those who select a distinct repertoire, like a newspaper or cable operator, do. Similarly, those who provide the facilities for job interviews are not recognized as speakers, nor are the manufacturers of technologies that record or transform information from one form into another—like a typewriter, photocopier, or loudspeaker.”).

246. See *id.*

247. See *id.* at 1524. He also notes that “a censorial motive would probably trump the functionality doctrine.” *Id.* Essentially, if the Court determines that the regulation was promulgated with a censorial motive, it will extend First Amendment protection even if, absent the censorial motive, there would be no protection. *Id.*

248. See *id.* at 1520.

249. See *id.*

250. See *id.* at 1521.

251. See *id.* at 1530. *But see* EUGENE VOLOKH & DONALD M. FALK, FIRST AMENDMENT PROTECTION FOR SEARCH ENGINE SEARCH RESULTS 4 (2012) (arguing that the search results are curated, and the engine is not a mere tool).

252. See Wu, *supra* note 120, at 1528–29. Professor Wu notes that Google itself has made efforts to avoid legal responsibility for the websites listed on the search engine, arguing that it is a platform for information and not a speaker. *Id.*

253. See *id.* at 1533.

254. See *infra* Part III.

corporate speech doctrine would seem to be supportive of AI protection, but a requirement of human input or the functionality doctrine might limit protection.

Although there is some tension as to whether First Amendment jurisprudence supports AI receiving free speech protections, the philosophical and legal theories behind free speech favor expanding speech rights to AI.<sup>255</sup> As discussed above, there are four main theoretical justifications for free speech: negative theory, democratic theory, marketplace of ideas theory, and autonomy theory.<sup>256</sup> The more instrumental free speech theories—negative, democracy, and marketplace of ideas—provide strong support for AI having free speech rights.<sup>257</sup>

For the negative theory, an AI may be more effective at reaching the desired end: a check on government power.<sup>258</sup> For example, a “transparency AI” may be designed to continuously scour statutes and produce public updates on rulemaking that may have censorial motives, providing another watchdog against ulterior government motives.<sup>259</sup> Likewise, supporters of the marketplace of ideas theory likely will consider free speech protection for an AI to be a significant benefit—the more viewpoints, the better. If the AI can aid in the search for truth, it deserves First Amendment protection.<sup>260</sup>

For the democracy theory, an AI could help contribute to public discourse and provide benefits to listeners.<sup>261</sup> However, there is a concern that an AI might damage democratic institutions by controlling participation or overrepresenting a particular viewpoint because of technological advantages of the AI.<sup>262</sup> This criticism is conceptually related to the anti-distortion argument that the Supreme Court rejected in *Citizens United*.<sup>263</sup> Still, as far as the philosophical discussion goes, this critique poses a challenge to an AI receiving free speech protection.<sup>264</sup>

In contrast to the other First Amendment theories, the autonomy theory poses serious challenges to an AI receiving free speech protections.<sup>265</sup> Specifically, under the autonomy theory, one could argue that AI lacks the consciousness necessary to have moral personhood, and therefore, fundamental rights such as freedom of speech.<sup>266</sup> However, the fundamental

---

255. See SCHAUER, *supra* note 144, at ix (“Unless we can get clear about the philosophical underpinnings of a political principle, we can hardly navigate successfully through the waters of specific application of that principle.”).

256. See *supra* Part I.B.3 (explaining the theoretical justifications for free speech protections in the United States).

257. See Massaro & Norton, *supra* note 19, at 1176–78.

258. See Massaro et al., *supra* note 120, at 2494.

259. See *id.*

260. See Massaro & Norton, *supra* note 19, at 1178.

261. See *id.* at 1176.

262. See *id.* at 1177.

263. See *supra* Part I.B.3.

264. See Massaro & Norton, *supra* note 19, at 1177.

265. See *id.* at 1178.

266. See SCHAUER, *supra* note 144, at 47, 50, 53; see also Solum, *supra* note 172, at 1262. This argument is a form of the missing-something theory of human intelligence discussed earlier. See *supra* Part I.C.1.

rights version of the autonomy theory is not the only one.<sup>267</sup> If the articulation instead focuses on the self-realization aspect of freedom of speech or expression, then an AI may have a stronger claim for protections.<sup>268</sup> An AI is necessarily an extension of the creator's expression, and limiting the protections for an AI limits the creator's ability to fully self-realize.<sup>269</sup>

Both a doctrinal analysis of First Amendment free speech jurisprudence and a review of free speech theories identify several hurdles that an AI must pass before it could receive protection. Even if those questions are satisfied, the essential issue remains: does an AI have traits that are sufficiently similar to currently protected speakers that would justify extending free speech protections?

### B. *Philosophy of Cognition and AI*

When thinking about whether AI has characteristics sufficiently similar to speakers currently protected by the First Amendment, the theories behind human intelligence and cognition are important to consider. The primary debate over cognition is between the two philosophical viewpoints previously discussed: the output-only and the missing-something theories.<sup>270</sup> When AI is being considered for free speech protection, the choice between the two positions could prove to be dispositive. There are open questions about what courts and scholars should consider when determining whether AI is intelligent and sufficiently similar to currently protected speakers.<sup>271</sup> Each position would likely reach different conclusions.<sup>272</sup> This part explores the tension between the philosophical theories of human intelligence when applied to the debate surrounding AI's eligibility for free speech protections.

The output-only position generally will favor an AI receiving free speech protections. As previously discussed, the output-only position argues that so long as the results of the process satisfy the criteria, then the actual process is irrelevant to the conclusion.<sup>273</sup> The theory purposefully ignores the identity or qualities of the speaker and the mechanisms of the process.<sup>274</sup>

---

267. See *supra* Part I.B.1.

268. See SCHAUER, *supra* note 144, at 47, 50, 53.

269. See Benjamin, *supra* note 75, at 1474.

270. See *supra* Part I.C.

271. See Solum, *supra* note 172, at 1262.

272. To date, legal scholarship on AI has not explicitly addressed the application of philosophical theories of human intelligence to AI. Some prominent scholars, including Professor Lawrence B. Solum, predicted that these theories would become relevant in future discussions surrounding AI free speech rights. *Id.* Specifically, Professor Solum argued that if an AI became sophisticated enough to warrant serious discussion of free speech rights, and legal personhood more broadly, there would be a debate between the missing-something theory and other philosophical theories on human intelligence. *Id.* This Note builds on Professor Solum's predictions and applies philosophical theories of human cognition to the legal questions surrounding potential First Amendment rights for AI.

273. See *supra* Part I.C.1.

274. See Turing, *supra* note 173, at 433; see also Volokh, *supra* note 6, at 1138.

Perhaps the criteria for a sufficient output may be that the speech produced is indistinguishable from human speech.<sup>275</sup> For an AI to act as a judge or attorney, Professor Volokh suggests a “[m]odified John Henry Test.”<sup>276</sup> This test would require an AI to perform as well as an average practitioner.<sup>277</sup> Regardless, assuming that the AI could pass the test, then the output-only position would tend to support free speech protections for a nonhuman. In the free speech context, perhaps the criteria might be something close to the Turing test; AI would thus be eligible for protection when most people could not distinguish the speech produced by the AI from speech produced by a human.<sup>278</sup> On the other hand, a more stringent set of criteria might be established; only AI speech that meets certain metrics of sophistication or accuracy would be eligible for protection. Unless the criteria are extremely demanding, it is very likely that an AI will be able to meet the criteria. Therefore, the output-only position will tend to support free speech protection for an AI.

The missing-something position, most popular in the literature,<sup>279</sup> would seem to reject protection for AI as a threshold matter.<sup>280</sup> Generally, the missing-something position argues that an AI is “missing something”—intelligence, and more specifically, intentionality, rationality, or consciousness—which humans have and are necessary traits to receive free speech protection.<sup>281</sup> For example, Professor Harry Surden states that “AI systems are often able to produce useful, intelligent results without intelligence.”<sup>282</sup> Professor Surden’s position implies that some characteristic of intelligence is a requirement for constitutional protection.<sup>283</sup> Professor Surden’s position is similar to the basic conclusion of Searle’s Chinese Room thought experiment discussed earlier.<sup>284</sup> In the Chinese Room thought experiment, a person inside the room can appear to be using intelligence to craft responses to questions from outside the room, even though no thinking is actually occurring inside.<sup>285</sup> Searle argued that by only considering the appearance of the output of a system, certain necessary traits may be absent.<sup>286</sup>

---

275. See, e.g., Turing, *supra* note 173, at 433.

276. See Volokh, *supra* note 6, at 1138.

277. *Id.*

278. See Turing, *supra* note 173, at 433–34.

279. See Solum, *supra* note 172, at 1262.

280. See *supra* Part I.C.1.

281. See Solum, *supra* note 172, at 1260–62, 1283; see also Massaro & Norton, *supra* note 19, at 1182. Professor Solum notes that the supporters of this position still need to demonstrate why the “something” that an AI is missing matters for free speech rights. See Solum, *supra* note 172, at 1262. This Note assumes that intentionality, rationality, and consciousness are necessary characteristics to warrant free speech protection.

282. See Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U. L. REV. 1305, 1308 (2019).

283. See *id.*

284. See Searle, *supra* note 184, at 417.

285. See *supra* Part I.C.1.

286. See *supra* Part I.C.1.

Applying the missing-something theory to the AI free speech question, certain important and necessary human cognitive processes may always be absent from AI, effectively placing it outside the scope of protected speakers.<sup>287</sup> At this point, the discussion warrants an inquiry into what the relevant traits of human cognition are<sup>288</sup>—specifically, the process that humans use to create speech.<sup>289</sup> This inquiry will require courts to review what traits modern psychological research has identified in the human cognitive process.

Scholars such as Professors Wu and Stuart Minor Benjamin have identified several different traits central to the human cognitive process.<sup>290</sup> Professor Wu’s functionality doctrine argues that most AI will be “missing something.” He identifies “a lack of identification with the information [the AI] handles, along with a lack of specific knowledge” as two traits necessary for protection but absent in AI.<sup>291</sup> Additionally, Professor Wu argues that an AI will operate as a communicative tool rather than as an expression of speech, which means that an AI is missing relevant characteristics to warrant protections.<sup>292</sup> Professor Wu’s argument is also similar to the critique in Searle’s Chinese Room thought experiment. Professor Wu argues that an AI would not identify data in the same way that a human would.<sup>293</sup> Professor Benjamin does not take such a hardline approach but argues that a lack of human decision-making in the AI system could make it ineligible for First Amendment protection.<sup>294</sup> Under the missing-something approach, AI speech must have characteristics similar to human speech—a trait that most scholars doubt an AI does or could ever have.<sup>295</sup>

For an AI to successfully receive protection under the First Amendment, there are multiple hurdles to overcome. Although *Citizens United* paved the way for further expansions of nonhuman speech protections, there is some tension between current First Amendment doctrine and the underlying theories supporting free speech as applied to AI. Further, there are unanswered questions surrounding the applicability of philosophical theories of intelligence to AI. These challenges must be resolved if AI are to have a viable claim for free speech rights.

---

287. *See id.*

288. *See* Solum, *supra* note 172, at 1262; *see also supra* Part I.C.2 (discussing the psychological research surrounding human decision-making).

289. *See* Solum, *supra* note 172, at 1262.

290. Notably, few, if any, of these scholars cite to psychological research regarding human cognition. This Note contends that psychological research regarding human decision-making is a central component to the debate surrounding AI free speech rights. *See supra* Part I.C.2.

291. *See* Wu, *supra* note 120, at 1520.

292. *Id.* Specifically, Professor Wu argues that AI fails the functionality requirement. *See id.*

293. *See, e.g.,* Searle, *supra* note 184, at 417.

294. *See* Benjamin, *supra* note 75, at 1479–82.

295. *See, e.g.,* Solum, *supra* note 172, at 1162, 1283.

### III. WEAK AI SHOULD RECEIVE FREE SPEECH RIGHTS UNDER THE FIRST AMENDMENT

Given AI's rapid development and its ability to produce some speech similar to human speech, there are novel and complex questions regarding whether weak AI—the most common type of AI—is eligible for free speech protections under the First Amendment. While there are multiple hurdles for AI to clear before possibly obtaining free speech rights,<sup>296</sup> this part argues that First Amendment free speech jurisprudence, free speech theories, and philosophical theories of cognition support extending free speech protection to weak AI. Part III.A demonstrates how the Supreme Court's decision in *Citizens United v. FEC*, which extended speech rights to nonhuman actors, presents a compelling framework for granting AI free speech rights. Part III.B establishes how both the output-only and missing-something theories of cognition further support protection for AI speech when the actual process of human cognition is considered. Part III.C proposes a hypothetical AI program to illustrate how the Supreme Court may approach analyzing an AI's claim for free speech protection in light of the current legal and philosophical doctrines.

#### A. Weak AI Is Similar to Human and Corporate Speakers

The Supreme Court's current free speech jurisprudence supports extending free speech protections to AI. As previously discussed, in *Citizens United*, the Court used a principle of speaker equivalence to hold that corporate speakers and natural persons should be given comparable free speech protections.<sup>297</sup> While the principle of speaker equivalence may not extend so far as to require the exact same protection for every type of speaker,<sup>298</sup> it strongly supports the application of a standard analytical framework to novel speakers.<sup>299</sup> Applying this principle to AI, when an AI makes a claim for free speech protection in response to a government regulation, the Court will likely utilize the same free speech framework that it uses for claims by human and corporate speakers. Under this analysis, the Court will first classify the regulation as content based or content neutral, and then review the regulation under the appropriate level of scrutiny.<sup>300</sup>

In addition to applying a principle of speaker equivalence, the Court in *Citizens United* indicated that it would be reluctant to immediately discount new types of speakers as not deserving of First Amendment protection.<sup>301</sup> This approach aligns with the general principles established by the Court's

---

296. See *supra* Part II.

297. See *Citizens United v. FEC*, 558 U.S. 310, 366 (2010).

298. See *id.*

299. See, e.g., *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 570 (2011) (comparing prescriber-identifying information sold for drug marketing to other sources of information like a beer bottle label); see also *Citizens United*, 558 U.S. at 366.

300. See *supra* Parts I.B.1–2; see also *infra* Part III.C (applying the Supreme Court's First Amendment jurisprudence to a hypothetical AI program regulated by a hypothetical congressional enactment).

301. See *Citizens United*, 558 U.S. at 392–93 (Scalia, J., concurring).

decision: although corporations look different from human speakers, they have voices that have a right to be heard without severe regulation, just as human speakers do.<sup>302</sup> AI also have voices,<sup>303</sup> similar to a corporation, and therefore deserve the same consideration regarding government limitations on their speech. Justice Scalia's concurrence, which stated that the identity of the speaker is irrelevant to First Amendment protection, further supports applying the Court's traditional free speech analysis to AI.<sup>304</sup>

Further, many commentators agree that *Citizens United* supports broadly extending equivalent First Amendment protections to AI.<sup>305</sup> First, the Court acknowledged that it would be particularly cautious, and perhaps even err on the side of overprotection, when a claim for free speech protection involves novel technology.<sup>306</sup> Additionally, the Court noted that the very nature of expression warrants a less strict approach to regulation of novel speakers.<sup>307</sup> Second, the Court stated that speaker-based regulations are subject to strict scrutiny review.<sup>308</sup> While the Court stepped back from an absolute rule for speaker-based regulations in *Reed*, the Court closely reviews speaker-based regulations for any potential content-based motivations.<sup>309</sup> The Court is suspicious of speaker-based regulations written to appear neutral but actually designed to target content.<sup>310</sup> For an AI, this means that even a facially neutral regulation would be cause for the Court to closely scrutinize the law's legislative history. As discussed, the Court nearly always strikes down speech regulations under strict scrutiny.<sup>311</sup>

Although *Citizens United* provides a framework for granting AI speech rights, aspects of the case may also be problematic for AI. Justice Scalia at one point refers to corporations as being associations of humans—as opposed to fully independent speakers.<sup>312</sup> If the Court applied this principle to AI, it may determine that AI is not eligible for free speech protections without having human input or presence.<sup>313</sup> However, even if the Court decides that some amount of human presence is necessary, AI still has a strong claim for free speech protection. As previously discussed, there may be some human

---

302. *See id.* at 354 (majority opinion) (noting that corporations have a right to advise “voters on which persons or entities are hostile to their interests”).

303. *See supra* Part I.A.

304. *See Citizens United*, 558 U.S. at 392–93 (Scalia, J., concurring) (arguing that the First Amendment “is written in terms of ‘speech,’ not speakers”).

305. *See, e.g.,* Benjamin, *supra* note 75, at 1447; Massaro & Norton, *supra* note 19, at 1185.

306. The Court noted that “[r]apid changes in technology—and the creative dynamic inherent in the concept of free expression—counsel against upholding a law that restricts political speech in certain media or by certain speakers.” *Citizens United*, 558 U.S. at 364.

307. *Id.*

308. *Id.* at 341.

309. *See Reed v. Town of Gilbert*, 576 U.S. 155, 170 (2015).

310. *See id.*

311. *See Burson v. Freeman*, 504 U.S. 191, 193, 195 (1992) (plurality opinion) (applying strict scrutiny and upholding a Tennessee statute restricting electioneering near polling sites); Benjamin, *supra* note 75, at 1451 n.19 (noting that a majority of the Court has only upheld a speech regulation under strict scrutiny in one case).

312. *See Citizens United*, 558 U.S. at 392–93 (Scalia, J., concurring).

313. *See Benjamin*, *supra* note 75, at 1479.

input within AI programs.<sup>314</sup> For instance, human actors typically play a role in establishing the goals of the AI system; they also can adjust the system's data sets while the AI is operating.<sup>315</sup> In that example, the human presence is similar to the way in which a corporation speaks—while humans decide the goals, the nonhuman entity, corporation or AI, effectuates them.<sup>316</sup> Therefore, human involvement will not be dispositive for an AI seeking protection since there are several areas in which humans are involved within AI programming.

In addition, the de facto functionality doctrine may also block AI from receiving free speech protections.<sup>317</sup> As discussed above, Professor Wu argues that, under the de facto functionality doctrine, many AI lack certain qualities of human speakers.<sup>318</sup> Namely, they are either mere communicative tools, like a map, or they operate as conduits for speech, but do not have enough interaction with the subject material, like a Google search.<sup>319</sup> Because AI lacks qualities of human speakers, Professor Wu contends that it is ineligible for free speech protection.<sup>320</sup>

However, Professor Wu's position does not account for recent technological advancements in AI or the process of human cognition. First, AI will often sufficiently communicate an idea to warrant free speech protection.<sup>321</sup> Advances in AI technology mean that a number of AI will likely meet the Court's standard for speech—expressing an idea reasonably likely to be understood by an observer.<sup>322</sup> Second, accounting for the process of human cognition, certain human speech traditionally protected by the First Amendment would fail the de facto functionality doctrine. In at least some cases, humans currently act as mere conduits for information.<sup>323</sup> In other words, they do not curate or interact with the information to the level that the de facto functionality doctrine seems to require, but the speech is nonetheless protected. Therefore, the line between protected “pure” speech and

---

314. *See supra* Part I.A.

315. *See Ng, supra* note 13.

316. This example ties into a potential issue regarding who would have standing to bring a lawsuit to vindicate the AI's rights. To have standing, a party must have suffered an injury. *See Garrett, supra* note 120, at 153. The Supreme Court requires that a party's injury be “concrete, particularized, and actual or imminent; fairly traceable to the challenged action; and redressable by a favorable ruling.” *Monsanto Co. v. Geertson Seed Farms*, 561 U.S. 139, 149 (2010). There are several ways that the creator of the AI may have standing, either as an organization, association, or a third party. *See, e.g., Burwell v. Hobby Lobby Stores, Inc.*, 573 U.S. 682, 688–90 (2014) (example of organizational standing); *Hunt v. Wash. State Apple Advt. Comm'n*, 432 U.S. 333, 335 (1977) (example of associational standing); *Singleton v. Wulff*, 428 U.S. 106, 112–17 (1976) (example of third-party standing). However, a discussion of standing is beyond the scope of this Note.

317. *See Wu, supra* note 120, at 1497.

318. *See id.*

319. *See id.*

320. *See id.*

321. *See supra* Parts I.A, I.B.1.

322. *See Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Bos., Inc.*, 515 U.S. 557, 569 (1995) (noting that a particularized message is not necessary for protection); *see also Spence v. Washington*, 418 U.S. 405, 410–11 (1974) (per curiam).

323. *See supra* Part I.C.

unprotected “functional” speech seems blurrier than at first glance. As such, the de facto functionality doctrine does not preclude AI from receiving First Amendment protection.

Since AI presents novel speech issues, the Court will also likely consider whether the theoretical foundations of free speech protection support extending free speech rights to AI.<sup>324</sup> The Court typically relies on the marketplace of ideas theory in its First Amendment cases.<sup>325</sup> Under the marketplace of ideas theory, protecting AI speech increases the total number and variety<sup>326</sup> of ideas entering the marketplace. As such, the Court will be more inclined to expand free speech protections to AI.

Even if the Court departed from the marketplace of ideas theory, AI would still have a compelling claim for free speech rights. The negative theory and democracy theory strongly support protection for AI.<sup>327</sup> Because instrumental theories argue that speech serves a functional role in communicating ideas, these theories do not focus specifically on the content of the speech or the identity of the speaker.<sup>328</sup> AI could be designed to exemplify the interests that the theories seek to advance.<sup>329</sup> For example, a democracy theory–focused AI could be designed to aid public discussion by reviewing statutes and other governmental actions for antidemocratic motives.<sup>330</sup>

Although certain variations of the autonomy theory may cut against protection for AI,<sup>331</sup> the Court rarely relies on them.<sup>332</sup> Further, even if the Court used the autonomy theory, an AI may be viewed as an expression of its creator.<sup>333</sup> Therefore, government restrictions on AI speech are effectively restrictions on the creator’s self-expression, which the autonomy theory seeks to protect. The Court would hesitate to uphold restrictions that run contrary to the theoretical foundations of free speech.<sup>334</sup> Therefore, the free speech theories all broadly favor extending free speech protections to AI.

---

324. See *Citizens United v. FEC*, 558 U.S. 310, 339 (2010) (referring to the democracy theory—“[s]peech is an essential mechanism of democracy”); *id.* at 340 (referring to the negative theory—“[p]remised on mistrust of governmental power”); *id.* at 354 (referring to the marketplace of ideas—“[regulation] interferes with the ‘open marketplace’ of ideas protected by the First Amendment”).

325. See Benjamin, *supra* note 75, at 1455; Massaro & Norton, *supra* note 19, at 1178.

326. Reviewing what patterns an AI identifies and the conclusions it reaches is another source of information that can lead to finding truth and knowledge—one that, often, will not be discovered by human thinkers. See *supra* Part I.A.

327. See *supra* Part II.A.

328. See *supra* Part II.A.

329. See *supra* notes 52–55 and accompanying text (explaining AI’s ability to address complex sources of information better than human counterparts).

330. See Massaro et al., *supra* note 120, at 2494; Massaro & Norton, *supra* note 19, at 1178.

331. See *supra* Part I.B. Other formulations of the autonomy theory would actually support extending free speech rights to AI or, at least, if not deserving on their own, as an extension of their creator’s speech rights. See *supra* Part I.B.

332. See *supra* Part II.A.

333. See Benjamin, *supra* note 75, at 1474.

334. See *Citizens United v. FEC*, 558 U.S. 310, 339–340 (2010).

The Court's decision in *Citizens United* and the free speech theories provide strong support for using the standard doctrinal approach when analyzing whether AI is entitled to free speech protection. However, there are critics, who adopt a form of the missing-something position, that argue that even if the free speech theories and First Amendment doctrine support granting protection, AI is missing a relevant trait of human cognition that is necessary for free speech protection. Therefore, the essential inquiry is whether AI possesses traits relevantly similar to protected human speech.

*B. Weak AI Has Relevantly Similar Traits to Some Protected Human Speech*

To determine whether AI possesses traits relevantly similar to protected human speech, it is critical to understand the process by which humans produce speech.<sup>335</sup> As previously discussed, the philosophical literature posits two main views surrounding the definition of human intelligence and cognition: the output-only position and the missing-something position.<sup>336</sup> The output-only position only considers the traits of the speech produced and therefore supports extending free speech protection to AI.<sup>337</sup> On the other hand, the missing-something position argues that AI does not possess certain traits present in human speech that are necessary to receive free speech protection.<sup>338</sup>

These traits fall into three general categories: (1) consciousness (recognition and identification of the subject matter), (2) intentionality (purposefully choosing to speak in a certain way), and (3) rationality (involving logical higher-order processing).<sup>339</sup> Although some scholars contend that AI is missing one or more of these traits, they presume that humans always possess these traits when producing speech.<sup>340</sup> As a result, these positions fail to fully account for the psychological research on the cognitive process of human speech and human decision-making. This section, using Professor Haidt's social intuitionist model,<sup>341</sup> responds to the missing-something position and demonstrates how some human speech lacks those traits.

According to Professor Haidt, human moral judgments are sometimes caused by intuitions.<sup>342</sup> Intuitions, as discussed above, are automatic emotional responses to a particular situation.<sup>343</sup> They are necessarily

---

335. *See supra* Part II.B.

336. *See Solum, supra* note 172, at 1262.

337. *See supra* Parts I.C.1, II.B.

338. *See supra* Parts I.C.1, II.B.

339. Each scholar that supports the missing-something position has a slightly different view on what traits of human speech are necessary for First Amendment protection. *See supra* Parts I.C.1, II.B. This Note categorizes these views into three main categories.

340. Even after identifying the trait in humans, there is still the matter of justifying why that trait should be necessary for free speech protection. *See Solum, supra* note 172, at 1262.

341. *See supra* Part I.C.2.

342. *See Haidt, supra* note 28, at 814–15.

343. *See id.*

unconscious, lack intention, and are irrational.<sup>344</sup> While humans can perform conscious, intentional, and rational decision-making, post hoc rationalization of intuition can make an individual believe that they reasoned out a judgment.<sup>345</sup> Since some human speech is purely the product of an emotional intuition, both humans and AI can produce speech lacking consciousness, intentionality, and rationality, and be unaware of how the speech was produced.

In light of this information, both the output-only and missing-something positions should agree: AI speech is relevantly similar to certain human speech. AI is actually *missing nothing*. Since AI is missing nothing, the only difference between an AI and a human producing speech is the speaker, not the speech. Since the speech, process, and speakers are comparable, following an equivalence principle, the Supreme Court should apply the same free speech analysis to AI that it does to natural persons and corporations.<sup>346</sup> With the addition of the psychological research, the jurisprudence, free speech theories, and philosophical theories all support extending free speech protection to AI that produces comparable speech.

### C. “Dogood” the AI

To further illustrate how an AI would have a plausible claim to free speech and how the Supreme Court may analyze an AI’s free speech claim, this section presents a thought experiment involving an AI program created to help with campaign advertising during an election cycle.

Imagine that Congress passes a comprehensive act entitled the “Election Integrity Protection Act” (EIPA) with a number of limits on the preelection activities of humans and corporations. EIPA provides, in part, that “any AI that influences over one million natural persons” is subject to restrictions on how it may operate on the internet, such as daily time limits on when the AI may contact voters, as well as disclosure requirements. The statute permits AI to engage in personalized interactions with voters and to obtain certain personal data. As part of EIPA, Congress permits the U.S. Department of Justice to seek an injunction against an AI program and its owner when the program “interferes with election integrity,” such as by influencing natural persons in violation of the statutory limits on internet usage. EIPA goes into effect two years before the next congressional elections.

In this hypothetical, there are companies renting out “canvassing AIs.” This type of AI interacts with people through written communication over email and on social media platforms including Twitter and Facebook. One particular canvassing AI, “Dogood,” operated by Courant, Inc., is well

---

344. *See id.*

345. *See id.* Some humans are not aware of this fact and do not attempt to change their intuition through methods such as self-reflection. *See id.*

346. *See supra* Part I.B.1 (explaining the Court’s various tests for analyzing content-based and content-neutral restrictions of speech under the First Amendment).

known for being one of the top performing AI programs.<sup>347</sup> Dogood has a history of persuasively and successfully representing various political parties and positions, such as a third-party candidate for state senator and the legalization of marijuana via a referendum. Dogood works for whoever pays Courant for its services—regular customers are politicians and groups across the ideological spectrum. As an AI, Dogood is less costly than lobbyists are in interacting with voters over email and on social media.<sup>348</sup> Using machine learning, Dogood writes persuasive emails and social media posts and adapts its approach based on the responses it receives.<sup>349</sup> Using big data, Dogood monitors trends and uses statistical information on each prospective voter to provide personalized responses to the people with whom it interacts.<sup>350</sup>

One year before the next congressional election, Courant enters into a contract with a minor third party hoping to win its first seats in Congress. The contract stipulates that Dogood will canvass for the party's candidates in a number of battleground states. Courant receives data on the candidates and legally purchases personal data of social media users in the battleground states. Dogood then gets to work, eventually reaching an audience of over one million voters. The Department of Justice receives a tip regarding the newly signed contract. After reviewing the situation, the Department of Justice seeks an injunction in federal court blocking Dogood from participating in the upcoming election due to violations of EIPA. Reviewing EIPA, the district court finds that the statute covers entities such as Dogood, the type of AI highly effective at influencing individuals. Therefore, pursuant to EIPA, the district court orders Courant to stop Dogood from engaging with voters. Courant, after appeals, presents the case to the Supreme Court, arguing, *inter alia*, that EIPA is an unconstitutional restriction on free speech.<sup>351</sup> The Department of Justice counters, arguing that Dogood does not have protected speech rights under the First Amendment, and therefore, the regulation is constitutional.

Assume the Supreme Court extended an equivalence principle as discussed above<sup>352</sup> and reviewed the regulation on AI in the same way that it reviews regulations on natural or corporate speakers. Applying its First Amendment jurisprudence, the Court would first determine whether the law was content

---

347. These names are inspired by Benjamin Franklin's pen name, "Silence Dogood." Benjamin Franklin resorted to this pen name when his older brother refused to publish letters by Franklin in his newspaper, *The New-England Courant*. See generally 1 BENJAMIN FRANKLIN & WHITFIELD J. BELL, *THE PAPERS OF BENJAMIN FRANKLIN* (Leonard Woods Labaree ed., Yale University Press 1959) (1734).

348. See *supra* Part I.A (highlighting the benefits of AI as compared to traditional computer programs and human operations).

349. See *supra* Part I.A (explaining how AI programs can operate efficiently using machine learning systems).

350. See *supra* notes 58–59 and accompanying text (explaining AI's use of big data).

351. There may be a number of procedural hurdles in this case, including standing. See *supra* note 316 and accompanying text. However, a discussion of these procedural issues is beyond the scope of this Note. This hypothetical only addresses potential free speech issues under the First Amendment.

352. See *supra* Part III.A.

neutral or content based.<sup>353</sup> Recall that a content-based regulation of speech targets either the speech's subject matter<sup>354</sup> or the speech's viewpoint or ideology.<sup>355</sup> Content-based regulations are subject to strict scrutiny.<sup>356</sup> Conversely, a content-neutral regulation of speech does not regulate content, but instead regulates conduct or the time, place, and manner of speech.<sup>357</sup> Content-neutral regulations are subject to intermediate scrutiny or the *O'Brien* test (a specific type of intermediate scrutiny).<sup>358</sup>

Here, the Court will likely characterize EIPA as a content-based regulation, as the statute generally concerns political speech—an area of speech traditionally given strong protections.<sup>359</sup> Still, there are provisions of EIPA that seem to regulate *conduct* as opposed to *content*. First, EIPA regulates preelection activities that “influence” voters. As such, these activities could touch on any topic, not necessarily political speech.<sup>360</sup> Additionally, EIPA's specific internet restrictions concern the time and manner of speech, as they limit online canvassing to certain hours each day and have disclosure requirements if one million individuals are influenced. Because EIPA contains both content-based and content-neutral aspects, this section reviews the regulations under both constitutional tests.

Starting with an analysis of content-based regulations, EIPA likely would not be upheld under strict scrutiny review. To survive strict scrutiny, the regulation must be narrowly tailored to serve a compelling government interest.<sup>361</sup> EIPA serves a government interest in election integrity. While the interest in protecting election integrity is likely compelling,<sup>362</sup> EIPA would likely fail the narrowly tailored prong. EIPA does not target the content of an influencer's speech, and in prior cases, the Court has hesitated to conclude that such a broad and sweeping regulation is narrowly tailored.<sup>363</sup> Even if Congress adopted a more restrictive regulation through EIPA, the

---

353. *See supra* Part I.A.

354. *See, e.g., Sable Commc'ns v. FCC*, 492 U.S. 115, 117–18 (1989) (invalidating a state statute criminalizing adult access to indecent phone messages).

355. *See, e.g., Rosenberger v. Rectors & Visitors of the Univ. of Va.*, 515 U.S. 819, 822–23 (1995) (invalidating a university plan that provided funding to secular newspaper publications but not to religious publications).

356. *See id.*

357. *See generally* *Ward v. Rock Against Racism*, 491 U.S. 781 (1989); *United States v. O'Brien*, 391 U.S. 367 (1968).

358. *See supra* Part I.B.1 (discussing the level of scrutiny applied to content-neutral speech regulations).

359. *See* *Citizens United v. FEC*, 558 U.S. 310, 340 (2010).

360. Whether the term “influencing” is content based is a nuanced inquiry. Previously, the Court has held that a regulation on “oral protest, education, or counseling” near a health-care facility was content neutral because it did not address the viewpoint or the message behind the protest, education, or counseling. *See Hill v. Colorado*, 530 U.S. 703, 742 (2000) (Scalia, J., dissenting). Justice Scalia disputed that such regulation was truly content neutral, even though it did not target a particular viewpoint or subject matter. *Id.*

361. *See* *Sable Commc'ns v. FCC*, 492 U.S. 115, 126 (1989).

362. *See* *Citizens United*, 558 U.S. at 341.

363. *See* *Reed v. Town of Gilbert*, 576 U.S. 155, 171 (2015); *see also* *Watchtower Bible & Tract Soc'y of N.Y., Inc. v. Village of Stratton*, 536 U.S. 150, 165 (2002) (“The mere fact that the ordinance covers so much speech raises constitutional concerns.”).

statute would still generally regulate political speech; the Court has made clear that it will closely scrutinize government restrictions on political speech.<sup>364</sup> Additionally, the act grants broad discretion to the government to determine which activities interfere with election integrity.

Turning to an analysis of content-neutral regulations, EIPA does not seem to implicate the *O'Brien* test. The *O'Brien* test covers conduct with expressive qualities.<sup>365</sup> As applied here, EIPA does not target conduct that is covered by *O'Brien*. EIPA is written to cover a form of online communication (influencing), which is not the neutral conduct that the *O'Brien* test is used for.<sup>366</sup>

EIPA regulations are not likely to be upheld as a time, place, or manner regulation. Recall that a TPM regulation must be “narrowly tailored to serve a significant governmental interest, and . . . leave open ample alternative channels for communication of the information.”<sup>367</sup> Although Congress has an interest in protecting election integrity, EIPA would likely run into problems with not being narrowly tailored. Even though EIPA regulates the time and manner of speech, it is a sweepingly broad regulation of a particular class of speakers and a significant amount of speech. This cuts against a finding that the statute is narrowly tailored.<sup>368</sup> Additionally, the Court has established strong protections for political canvassing.<sup>369</sup> An individual’s ability to “turn the AI away at the door” could well prove fatal to the regulation.<sup>370</sup> Regarding the third prong of the TPM test, there is a strong argument that EIPA severely restricts AI speech since AI is predominantly limited to online written communication.

The Dogood hypothetical illustrates how the Supreme Court might address questions regarding AI’s free speech rights. Although it is difficult to predict exactly how the Court will analyze a novel free speech issue, there are strong arguments in favor of expanding free speech rights to AI.

---

364. See *Citizens United*, 558 U.S. at 365.

365. See *United States v. O’Brien*, 391 U.S. 367, 382 (1968).

366. See *id.* (upholding restriction because it targeted the “noncommunicative” aspect of draft card burning—i.e., the disruption of the administration of the draft).

367. See *Ward v. Rock Against Racism*, 491 U.S. 791 (1989).

368. See, e.g., *Citizens United*, 558 U.S. at 340 (“Speech restrictions based on the identity of the speaker are all too often simply a means to control content.”); see also *Watchtower Bible*, 536 U.S. at 153 (striking down a regulation banning door-to-door canvassing without a permit). In addition to addressing the canvassing cases, the Court in *Watchtower Bible* stated that “[t]he mere fact that the ordinance covers so much speech raises constitutional concerns.” *Id.* at 165.

369. See *Watchtower Bible*, 536 U.S. at 156; see also *Heffron v. Int’l Soc’y for Krishna Consciousness, Inc.*, 452 U.S. 640, 654 (1981) (upholding political canvassing restrictions during a state fair due to public safety concerns); *Martin v. City of Struthers*, 319 U.S. 141, 146 (1943) (invalidating a regulation banning political canvassing because it is the individual homeowner’s choice, not the government’s, whether to turn away canvassers). In *Watchtower Bible*, the Court noted that the ability to door-to-door canvass has a long tradition of protection. See *Watchtower Bible*, 536 U.S. at 162.

370. See *Watchtower Bible*, 536 U.S. at 156. For example, this might occur if Dogood was programmed to cease communication at the request of the individual, or if they posted the digital equivalent of a “no soliciting” sign.

## CONCLUSION

Weak AI has a plausible claim for First Amendment protection. The recent extensions of free speech protection in First Amendment jurisprudence would seemingly establish a compelling framework for free speech protections for AI. The theories supporting free speech also tend to support extending protections. Philosophical and relevant psychological theories on human cognition and intelligence demonstrate that some protected human speech is relevantly similar to potential AI speech, which means that AI speech should be granted equivalent review under the First Amendment. Although the exact protection that AI might receive will depend on the features of the AI and the regulation at issue, if the AI is categorically similar to current speakers, then it is eligible for free speech protections under the First Amendment. Therefore, the Supreme Court should continue to embrace technological change and acknowledge AI's free speech rights.