Article

SIRI-OUSLY 2.0: What Artificial Intelligence Reveals About the First Amendment

Toni M. Massaro,[†] Helen Norton[††] & Margot E. Kaminski[†††]

INTRODUCTION

In the spring of 2016, Microsoft released a Twitter chatbot called MS Tay. MS Tay was described as an Artificial Intelligence (AI)—a learning algorithm—programmed to "learn" how to interact with other Twitter users and produce output delinked from the original programmers' control.[1] Within twenty-four hours, Twitter users learned how to game MS Tay's learning process. The mass mis-education of MS Tay resulted in Holocaust-denying, transphobic, and misogynistic Tweets, among others.[2] Microsoft quickly ended its experiment.[3]

† Regent's Professor, Milton O. Riepe Chair in Constitutional Law, and Dean Emerita, University of Arizona James E. Rogers College of Law.

†† Professor and Ira C. Rothgerber, Jr. Chair in Constitutional Law, University of Colorado School of Law.

††† Assistant Professor of Law, The Ohio State University Michael E. Moritz College of Law.

Our thanks to the *Minnesota Law Review*, and to Kiel Brennan-Marquez, James Grimmelmann, Meg Jones, Genevieve Leavitt, Blake Reid, Harry Surden, and Phil Weiser for their very thoughtful comments.

1. Sarah Perez, *Microsoft Silences Its New A.I. Bot Tay, After Twitter Users Teach It Racism*, TECHCRUNCH (Mar. 24, 2016), https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism; James Vincent, *Twitter Taught Microsoft's AI Chatbot To Be a Racist Asshole in Less than a Day*, The VERGE (Mar. 24, 2016), http://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist.

2. Perez, *supra* note 1. In the following Twitter exchange, TayTweets denied the Holocaust: "@ExcaliburLost: .@TayandYou did the Holocaust happen?"; "@Tayandyou: @ExcaliburLost it was made up (clapping hands emoji)." Excalibur Lost (@ExcaliburLost), TWITTER (Mar. 23, 2016, 3:25 PM), https://twitter.com/ExcaliburLost/status/712767221718773761; *see also* Shane

If MS Tay were a human speaker, there is no question that her Tweets would receive First Amendment protection. Whatever your view of the doctrine, and whatever your view of its impact on public political discourse, the United States Supreme Court has interpreted the First Amendment to protect hateful or otherwise offensive speech that does not rise to the level of incitement or "true threats."[4] Should the First Amendment similarly protect MS Tay's speech—and the speech of far more complex, more autonomous programs perhaps to come?[5]

In an earlier work, two of us explained how current free speech theory and doctrine support the claim that the First Amendment covers speech by "strong AI" (i.e., as-yet-hypothetical machines that would think and generate expressive content independent of human direction).[6] This is because First Amendment law increasingly focuses not on protecting speakers as speakers but instead on providing value to listeners and constraining the government.[7] If we take the logic of current First Amendment jurisprudence and theory to its natural conclusion, MS Tay's strong AI progeny could have First Amendment rights. *Siri-ously*.

Dingman, *How Microsoft's Friendly Robot Turned into a Racist Jerk in Less than 24 Hours*, GLOBE & MAIL (Mar. 24. 2016), http://www.theglobeandmail.com/technology/tech-news/how-microsofts-friendly-robot-turned-into-a-racist-jerk-in-less-than-24-hours/article29379054 (listing things Tay "learned," including the exchange with Excalibur Lost). MS Tay's Tweets have since been deleted.

3. *Id.*

4. *See* Snyder v. Phelps, 562 U.S. 443, 460–61 (2011) (holding that the First Amendment protects offensive speech); R.A.V. v. City of St. Paul, 505 U.S. 377, 395–96 (1992) (holding that the First Amendment prohibits the regulation of hate speech as a content-based category of fighting words); Brandenburg v. Ohio, 395 U.S. 444, 447 (1969) (holding that the First Amendment protects political speech that advocates illegal activity unless it is likely to incite imminent illegal action); Watts v. United States, 394 U.S. 705, 707 (1969) (holding that the First Amendment does not protect "true threats" from regulation).

5. We refer to these as-yet-hypothetical machines that *actually* think as "strong AIs," as opposed to "weak AI" machines that act "*as if* they were intelligent." STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 1020 (3d ed. 2010); *see also* Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 97 (2014) (describing ways in which machines can "learn" through employing heuristics and proxies "that ultimately arrive at the same or similar results as would have been produced by a similarly situated intelligent person employing higher order cognitive processes and training").

6. Toni M. Massaro & Helen Norton, *Siri-ously?: Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169 (2016).

7. *Id.* at 1175–86.

In this Article, we build on this provocative claim. In so doing, we consider important counter-arguments, and suggest ways in which the rise of AI may inspire critical engagement with free speech theory and doctrine. This Article starts by reprising our earlier observation that the United States Supreme Court now emphasizes listeners' interests in free speech outputs—rather than speakers' humanness or humanity—in ways that make it exceedingly difficult to place AI speakers beyond the First Amendment's reach. We then explain that First Amendment coverage of strong AI speakers would be consistent with the current Court's largely "negative" view of the First Amendment that focuses on the First Amendment as a means of constraining the government, rather than on protecting speech as a positive good.[8] Next, we explore a range of theoretical, doctrinal, and practical objections to the claim that the First Amendment may cover strong AI speech. Although we take these concerns seriously, we conclude that none of them eliminates the possibility of First Amendment coverage for AI speech. Finally, we suggest that current free speech law contains means by which courts could view some regulation of certain AI speech outputs as consistent with the First Amendment.[9]

Thinking about the potential emergence of strong AI speakers forces us to face complicated questions about the role of harm in First Amendment case law. This in turn drives, or really revives, discussions about whether and when speech causes harm—and whether and when the First Amendment

8. *Id.* at 1182 n.50.

9. As Frederick Schauer has observed, "The question of which forms of speech are covered by the First Amendment is . . . distinct from the question of how much protection the speech that is covered will receive." Frederick Schauer, *Out of Range: On Patently Uncovered Speech*, 128 HARV. L. REV. F. 346, 348 (2015). Some types of communication—such as a great deal of speech addressed by the law of contract, evidence, and antitrust—are "understood as having nothing to do with the First Amendment." *Id.* at 347. Such expression receives no First Amendment coverage, and thus can be regulated by the government without any free speech analysis at all. Once speech is *covered* by the First Amendment—that is, visible to First Amendment analysis at all—whether it is *protected* by the First Amendment from government regulation depends on the results of applying the appropriate level of scrutiny. As we explain below, even if AI speech is covered by the First Amendment and thus triggers First Amendment scrutiny, whether it is protected from government regulation is a separate question that turns on the application of the relevant level of review.

permits the regulation of speech to address some kinds of harms, versus others.[10]

This thought experiment highlights the centrality of the listener in ongoing as well as emerging free speech debates. If we justify expanding speech coverage on the grounds that the speech is good for listeners,[11] then we should also take seriously listeners' interests in protection from harms caused by that speech. This has implications for many current free speech problems, including the governance of information intermediaries and network neutrality;[12] of commercial speech;[13] of speech in the context of employment relationships;[14] and even of surveillance.[15]

10. *See, e.g.*, LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW § 12–2, at 580 n.9 (1st ed. 1978) (suggesting that government regulation of speech is suspect when the government "aim[s] at ideas or information," but that such regulation may be less suspect when the government aims at speech's "noncommunicative impact"); Rebecca L. Brown, *The Harm Principle and Free Speech*, 89 S. CAL. L. REV. 953, 954 (2016) (explaining that the "harm principle allows government to limit liberties as necessary to prevent harm" and that "we should be slow to assume that society is necessarily without power to protect itself from harm that expression may cause").

11. *See, e.g.*, Citizens United v. FEC, 558 U.S. 310 (2010).

12. *See* Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445, 1458 (2013) (describing how the Supreme Court's current First Amendment doctrine may apply to the substantive results produced by algorithms); James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868, 910 (2014) (detailing how multiple perspectives view the results of web-based, automatic algorithms differently under First Amendment doctrine); Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1513 (2013) (contending that more rigor and predictability ought to be injected into the method for courts' applications of First Amendment doctrine to technologically originated speech).

13. *See, e.g.*, TAMARA PIETY, BRANDISHING THE FIRST AMENDMENT 12 (2012) (evaluating and critiquing the judicial expansion of First Amendment doctrine to protect commercial entities).

14. *See generally* Helen Norton, *Truth and Lies in the Workplace: Employer Speech and the First Amendment*, 101 MINN. L. REV. 31 (2016) (critically reviewing recent trends in the ways that First Amendment protections expand in the employment context).

15. *See* Marc Jonathan Blitz, *Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right To Read, and a First Amendment Theory for an Unaccompanied Right To Receive Information*, 74 UMKC L. REV. 799, 881 (2006) (defending libraries as places where information seekers have a First Amendment right to receive information); Julie E. Cohen, *A Right To Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981, 983–84 (1996) (explaining how "copyright management" has developed to monitor and meter the use of intellectual property); Neil M. Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387, 388 (2008) (arguing that freedom from intellectual surveillance or interference is a cornerstone of First Amendment liberty because it allows citizens to freely make up their mind and develop new ideas).

Some will point out that the possibility of strong AI as yet remains—and may always remain—entirely hypothetical. We recognize that the feasibility of strong AI remains deeply contested,[16] and do not seek to resolve the debate among futurists about strong AI's likelihood. Instead, we use the possibility of strong AI speech to illuminate how the logic of free speech law and doctrine can carry us to unfamiliar and uncomfortable, or even dangerous, places.[17] Computers' growing expressive capacities cast meaningful light on puzzling and troublesome aspects of existing free speech law that may otherwise escape attention.[18]

Current free speech theory and doctrine support the extension of free speech rights to strong AI speakers (if such speakers ever come to exist), but such rights raise normative and practical difficulties. These may force reexamination of the contemporary turns in free speech law and theory that have made coverage of strong AI speakers plausible. Nothing prevents the Court from adjusting its current path, and we point out places

16. For an example of the view that the gap between capacities of human and computer speakers may be closing, see generally PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD (2015) (discussing machine learning—predictive analytics—and arguing that there may one day be a master algorithm that synthesizes the five primary schools of thought about machine learning and enables AI to teach itself from experience). Some futurists even predict that by century's end, there will be "no distinction . . . between human and machine or between physical and virtual reality." RAY KURZWEIL, THE SINGULARITY IS NEAR 9 (2005). The only unequivocally human quality will be that the human species "inherently seeks to extend its physical and mental reach beyond current limitations." *Id.*; *see also* YUVAL NOAH HARARI, SAPIENS: A BRIEF HISTORY OF HUMANKIND 407 (2015) (discussing how cyborg engineering may transform humankind such that humans would become "so fundamentally another kind of being that we cannot even grasp the philosophical, psychological or political implications"). For an example of a skeptical view, see Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 528 (2015) ("Little in the literature gives me confidence that artificial intelligence will approximate human intelligence in the foreseeable future. There are analytic and technical reasons to believe robots will never think like people.").

17. *But see* Lawrence Lessig, Commentaries, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 502 (1999) ("I am not defending the law of the horse. My claim is specific to cyberspace. We see something when we think about the regulation of cyberspace that other areas would not show us.").

18. *See generally* Jack M. Balkin, Commentary, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1 (2004) (wrestling with the implications of digital speech for his notion of democratic culture).

where it may well do so.[19] In the meantime, we start from a baseline that assumes the future Court will take current theoretical and doctrinal premises seriously, and we suggest possible paths forward.

I. JUSTIFICATIONS FOR FREE SPEECH RIGHTS FOR AI SPEAKERS

Chatbots like MS Tay are computer programs written by people. To be sure, the outputs of such algorithms may be increasingly less predictable and controllable by their creators.[20] Such algorithms are not, however, independent thinkers or beings. Contemporary discussions of whether algorithmic outputs should be protected under the First Amendment have largely focused on whether the output is functional or editorial in nature.[21] These discussions reserve, or do not address, the question of what to do about strong AI—as-yet-hypothetical machines that would actually think and generate expressive content independent of human direction—in contrast to current "weak AI" machines that only act as if they were intelligent.

As this Part explains, theoretical justifications for protecting speech under the First Amendment support protecting speech by strong AI both to provide value to human listeners and to restrain governmental excesses. The logic of the Court's current free speech doctrine also supports its extension to protecting strong AI.[22] Whether this makes normative sense is debatable. We aim to illuminate the normative tension, rather than resolve it.

We begin by revisiting the positive theoretical justifications for protecting strong AI speech first discussed in "Siri-ously?: Free Speech Rights and Artificial Intelligence" (Siri-ously 1.0).[23] We then build on that work to explore an additional thread of

19. This is far from the only area in which technological change may and perhaps should motivate doctrinal change. *See, e.g.*, Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. (forthcoming 2017) (manuscript at 9) (on file with the authors) ("Workforce analytics pose an entirely new set of challenges to [workplace] equality that calls for fundamentally rethinking antidiscrimination doctrine.").

20. Calo, *supra* note 16, at 538 (providing examples explaining how technology may change in years to come, including how technology may go awry).

21. *See supra* note 12 and accompanying text.

22. Massaro & Norton, *supra* note 6, at 1182–86 (examining the historical development of First Amendment doctrine, as well as how the First Amendment has grown to cover newly emerging technology).

23. *Id.*

First Amendment theory: the "negative" view that distrusts the government's exercise of its power to regulate speech, which lends further support for First Amendment coverage of AI speech.

A. THE "POSITIVE" VIEW: PROTECTING SPEECH TO PROVIDE VALUE TO LISTENERS

Conferring strong AI speakers with First Amendment rights is consistent with free speech theories that focus, among other things, on expression's usefulness to human listeners.[24] The elasticity of such theories makes it difficult to exclude non-human speakers entirely from their fold.[25] Courts and commentators have yet to settle on a single theory that informs the First Amendment.[26] The most influential theories have either focused on positive arguments that discuss the primary values that speech promotes—democratic self-governance;[27] enlightenment and the distribution of knowledge and ideas;[28] and individual autonomy[29]—or on negative arguments that focus on

24. *Id.* at 1183–84. In this project, we discuss U.S. freedom of expression law. In international law, readers' rights are more explicit. *See, e.g.*, International Covenant on Civil and Political Rights art. 19(2), *opened for signature* Dec. 19, 1966, 999 U.N.T.S. 171 (entered into force Mar. 23, 1976) ("Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds . . . ."); Molly K. Land, *Toward an International Law of the Internet*, 54 HARV. INT'L L.J. 393, 431 (2013) ("Article 19(2) explicitly calls for protection of the rights of individuals to receive information and expression from others, thus guarding not only the quintessential expressive activity of speaking but also the information-gathering activities that precede speech.").

25. Massaro & Norton, *supra* note 6, at 1175–76 (describing both how and why First Amendment doctrine's current breadth requires expanding the protections to technologically based speech).

26. *See, e.g.*, HARRY KALVEN, JR., A WORTHY TRADITION: FREEDOM OF SPEECH IN AMERICA 3 (Jamie Kalven ed., 1988) ("The Court has not fashioned a single, general theory which would explain all of its decisions; rather, it has floated different principles for different problems.").

27. *See generally* ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948) (arguing that freedom of speech derives from the necessities of self-governance rather than a natural right).

28. *See* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) ("[T]he best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out.").

29. *See, e.g.*, C. EDWIN BAKER, HUMAN LIBERTY AND FREEDOM OF SPEECH (1989) (emphasizing individualistic concerns and speaker liberty); RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 201–05 (1977) (focusing on speaker dignity and respect). Seana Valentine Shiffrin's work on a thinker-based First Amendment is among those that fall under the autonomy theory umbrella. *See*

the dangers of government action in this area.[30] We explain why none of these theories eliminates the possibility that the First Amendment covers AI speech.[31]

Democracy-based theories emphasize the value of speech to democratic self-governance, which usually entails focusing on public discourse rather than individual speakers.[32] Alexander Meiklejohn, often cited for developing this self-governance theory, observed that what matters for freedom of speech is not that all people speak, but that "everything worth saying shall be said."[33] Speaker identity plays little or no role in Meiklejohn's inquiry. Strong AI speech that contributes to the democratic process—i.e., that is "worth saying"—therefore may be covered.

Robert Post has set forth a theory of freedom of expression similarly based on principles of self-government. Post argues that the First Amendment is "designed to protect the processes of democratic legitimation" and fosters confidence that citizens participate in a legitimate process in which their representatives speak for them.[34] Admittedly, this version of democratic self-governance does focus on the dignity of human speakers. However, if and when they produce information useful to natural persons who seek to participate in public discourse, strong AI speakers should warrant similar protection. Likewise, if squelching AI speech draws into question the legitimacy of democratic process, then free speech values would be implicat-

Seana Valentine Shiffrin, *A Thinker-Based Approach to Freedom of Speech*, 27 CONST. COMMENT. 283, 284 (2011).

30. *See, e.g.*, Steven G. Gey, *The First Amendment and the Dissemination of Socially Worthless Untruths*, 36 FLA. ST. U. L. REV. 1, 17 (2008) (discussing a negative view of the First Amendment that "does not rest on the affirmative claim that free speech will lead to any particular social or political benefits" and instead emphasizes the dangers created "when collective entities are involved in the determination of truth").

31. For additional detail, see Massaro & Norton, *supra* note 6, at 1175–82.

32. *See* ALEXANDER MEIKLEJOHN, POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE 26 (1960) (asserting that what matters for freedom of speech is not that all speak, but that "everything worth saying shall be said"). To be sure, under this view, limits can and should be imposed where the speech does not serve this audience-sensitive value.

33. *Id.*

34. ROBERT C. POST, CITIZENS DIVIDED: CAMPAIGN FINANCE REFORM AND THE CONSTITUTION 8 (2014) (explaining that there must be "a chain of communication . . . 'sufficiently strong and discernible' to sustain the popular conviction that representatives spoke for the people whom they purported to represent" (quoting JAMES WILSON & THOMAS MCKEAN, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES OF AMERICA 30–31 (1792)).

ed.[35] That an AI rather than a human produces that information should not matter in Post's approach.

Jack Balkin goes beyond representative democracy justifications to offer a theory of free speech that uses the word "democratic" to mean "cultural participation—the freedom and the ability of individuals to participate in culture, and especially a digital culture."[36] Balkin's account gives explicit attention to speaker humanness, noting that "[h]uman beings are made out of culture. A democratic culture is valuable because it gives ordinary people a fair opportunity to participate in the creation and evolution of the processes of meaning-making that shape them and become part of them."[37] Balkin's focus, though, is not just on the humanness of the speaker; human listeners who receive the speech also matter.[38]

AI speech quite clearly can contribute to a human audience's meaning-making, and to human construction of selfhood in a cultural universe. Even AI speech aimed at another AI might, albeit indirectly, contribute to this meaning-making. Thus, Siri-ously 1.0 posited, "Balkin's democratic culture perspective . . . would not rule out cases in which strong AI speakers contribute to the democratic disco."[39] On the contrary, Balkin's perspective could afford more protection than other democratic discourse models. Where Meiklejohn's model might protect speech about elections, and Post's model might protect

35. We are not attributing this view to Post; we argue instead that it flows from his theory of democratic participation as the rationale for protecting speech. *See id.*

36. Jack M. Balkin, *Cultural Democracy and the First Amendment*, 110 NW. U. L. REV. 1053, 1059–61 (2016); *see also* Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1205 (2016) [hereinafter Balkin, *Information Fiduciaries*] (arguing in favor of expanded, quasi-fiduciary responsibilities for parties who manage and disseminate digital information); Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296 (2014) (comparing historic trends of First Amendment doctrine with more modern concerns, ultimately arguing for increased scrutiny of digitally based private speech because of its societal importance); Jack M. Balkin, *The First Amendment Is an Information Policy*, 41 HOFSTRA L. REV. 1 (2012) (distinguishing between democratic information states and authoritarian information states, discussing challenges posed by technology for the former, and arguing that thinking about knowledge and information policy and infrastructure matters as much to democratic information states as thinking about freedom of speech as an individual right).

37. Balkin, *supra* note 18, at 33.

38. *Id.* at 39 ("[P]rocesses of meaning-making include both the ability to distribute those meanings and the *ability to receive them*." (emphasis added)).

39. Massaro & Norton, *supra* note 6, at 1178.

AI contributions to the public sphere, Balkin's model also would protect AI musicians[40] and artists[41] as contributors to the culture by which human listeners and readers define themselves.

The classic marketplace-of-ideas approach to free speech questions emphasizes expression's instrumental value to listeners' knowledge and enlightenment.[42] The production of ideas and information is what matters, regardless of source. This theory presupposes that more speech best facilitates listeners' acquisition of knowledge and discovery of truth (whatever that means).[43] This "more speech beats less" justification casts an even wider First Amendment coverage net than self-governance theories. Under an enlightenment theory, speech that does not promote democratic participation still has much First Amendment value. And information that flows from nonhuman sources may have considerable value to human listeners. As long as the information contributes to the marketplace-of-ideas, its nonhuman source should not matter.

Finally, autonomy-based free speech theories—which emphasize the value of expression in furthering individual autonomy—simultaneously point in opposite directions as applied to AI speech.[44] On the one hand, if what matters is a speaker's autonomy in expressing her own thoughts and beliefs, then AI may not qualify for First Amendment coverage because even strong AI speakers are arguably not autonomous beings.[45] As Lawrence Solum has thoughtfully discussed, even strong AI might be thought to be "missing something"—souls, consciousness, intentionality, feelings, interests, and free will—in ways

40. Russell Brandom, *Google's Art Machine Just Wrote Its First Song*, THE VERGE (June 1, 2016), http://www.theverge.com/2016/6/1/11829678/google-magenta-melody-art-generative-artificial-intelligence.

41. Joseph Stromberg, *These Abstract Portraits Were Painted by an Artificial Intelligence Program*, SMITHSONIAN MAG. (Nov. 7, 2013), http://www.smithsonianmag.com/science-nature/these-abstract-portraits-were-painted-by-an-artificial-intelligence-program-180947590/?no-ist.

42. *See* FREDERICK SCHAUER, FREE SPEECH: A PHILOSOPHICAL ENQUIRY 15–34 (1982); Vincent Blasi, *Holmes and the Marketplace of Ideas*, 2004 SUP. CT. REV. 1, 33–44.

43. *See* Thomas I. Emerson, *First Amendment Doctrine and the Burger Court*, 68 CALIF. L. REV. 422, 423 (1980) (describing the values most often located at the heart of the First Amendment as including the search for truth and the discovery and dissemination of knowledge); Frederick Schauer, *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 HARV. L. REV. 1765, 1786 (2004) (same).

44. *See* Massaro & Norton, *supra* note 6, at 1178–82.

45. *Id.* at 1178.

that could be relevant to speaker-based autonomy theories.[46] We address these concerns in more detail in Part II.

On the other hand, autonomy-based theories counsel protection not just of autonomous human speakers, but also of autonomous human listeners who consume information and rely on others' speech when developing their own thoughts and beliefs.[47] To the extent that autonomy-based theories emphasize the autonomy of human listeners and readers, they support coverage of strong AI speech. Machines can and do produce information relevant to human listeners' autonomous decision-making and freedom of thought. As such, even the intuitively appealing argument that AI is still "missing something" does not eliminate the possibility of First Amendment coverage for AI speech.

B. THE "NEGATIVE" VIEW: CURBING GOVERNMENTAL POWER

The above theories all take a "positive" view of the First Amendment, which affirmatively urges that free expression provides value to individuals and communities that warrants constitutional protection.[48] "Negative" First Amendment arguments focus instead on the need to constrain the government's potentially dangerous exercise of control over expression,[49] and are rooted in distrust of the government as regulator rather than on theories that celebrate speakers.[50] The Court's increasingly negative view of the First Amendment[51] thus bolsters the above arguments for protection of AI speech.

46. Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231, 1262–76 (1992); *see also* Massaro & Norton, *supra* note 6, at 1179.

47. *See* Shiffrin, *supra* note 29 (describing a thinker-based approach to freedom of speech); *see also* Blitz, *supra* note 15 (promoting First Amendment protection of the right to receive information and ideas).

48. *See supra* notes 27–29 and accompanying text.

49. *See* Paul Horwitz, *The First Amendment's Epistemological Problem*, 87 WASH. L. REV. 445, 451 (2012) (describing a negative justification for the First Amendment as rooted "primarily on the grounds of distrust of government"); Nat Stern, *Implications of Libel Doctrine for Nondefamatory Falsehoods Under the First Amendment*, 10 FIRST AMEND. L. REV. 465, 503 (2012) ("To shelter ideas while leaving factual expression to plenary government control ignores an abiding First Amendment theme: wariness of government's capacity and motives when acting as arbiter of truth.").

50. *See* Gey, *supra* note 30, at 17 (emphasizing the dangers created "when collective entities are involved in the determination of truth").

51. *See* United States v. Alvarez, 132 S. Ct. 2537, 2574 (2012) (discussing the need for a limiting principle on the government's restriction of speech); United States v. Stevens, 559 U.S. 460, 470–72 (2010) (explaining that the

The Court currently treats the government's content-based regulation of speech as presumptively impermissible, absent a showing that the speech falls into a traditionally and historically exempted category of protected speech.[52] Even very troubling speech—animal crush videos, violent video games, and self-aggrandizing lies—ordinarily cannot be regulated in a content-specific manner without surviving the rigors of strict scrutiny.[53] The Court has justified this increasingly broad view of protection with skeptical references to the government's institutional competence and the government's limited ability to balance social costs and benefits when speech rights are at risk. Indeed, the Court's recent wariness towards speech regulations displays a negative view of the First Amendment at least as much as support for any of the positive theories discussed above. For example, in *United States v. Stevens*[54] the Court struck down on overbreadth grounds a federal law that criminalized the commercial creation, sale, or possession of depictions of animal cruelty.[55] In so doing, it rejected as "startling and dangerous" what it characterized as the government's proposed "free-floating test for First Amendment coverage . . . [based on] an ad hoc balancing of relative social costs and benefits."[56]

The negative view rests not only on concerns about shortcomings in the government's institutional competence (i.e., the government's clumsiness) but also on concerns about the government's censorial motives (i.e., its malevolence, its self-interest, or at the very least its paternalism). For example, the Court recently found in *Heffernan v. City of Paterson* that the government's improper speech-suppressing motive alone suf-

government cannot restrict speech solely because the benefits outweigh the costs); *see also* Helen Norton, *Lies and the Constitution*, 2012 SUP. CT. REV. 161, 163–92 (considering questions raised by the *Alvarez* holding, which protected false statements under the First Amendment).

52. Massaro & Norton, *supra* note 6, at 1186.

53. *See Alvarez*, 132 S. Ct. at 2551 (striking down a regulation that made false claims of receiving certain military honors illegal); Brown v. Entm't Merchs. Ass'n, 564 U.S. 786, 805 (2011) (striking down a regulation that prevented minors' access to violent interactive video games without parental consent); *Stevens*, 559 U.S. at 482 (striking down a regulation aimed at crush videos on overbreadth grounds); *cf.* Reed v. Town of Gilbert, 135 S. Ct. 2218, 2231–32 (2015) (rejecting a more nuanced, contextualized approach to a town's content-specific regulations of signs and instead applying strict scrutiny).

54. 559 U.S. 460.

55. *Id.* at 482.

56. *Id.* at 470; *see also Alvarez*, 132 S. Ct. at 2544.

ficed to establish a First Amendment violation.[57] In *Heffernan*, police department officials demoted an officer because they mistakenly believed him to support an opposition mayoral candidate.[58] The government was not insulated from liability based on its censorious motive because the employee did not actually support the candidate and had not actually engaged in political speech promoting the candidate.[59] More simply: the Court still sanctioned the government's mistaken demotion of him because of its speech-suppressing motive even though *the employee had not engaged in protected speech*.[60]

The text of the First Amendment, the Court noted in *Heffernan*, states that government "shall make no law . . . abridging the freedom of speech."[61] In so holding, the Court directly embraced the negative view of the First Amendment by emphasizing its restraints on potentially dangerous governmental power rather than positive reasons for protecting speakers or speech. The Court thus found a First Amendment violation simply because of the government's impermissible motive, even in the absence of a human speaker engaged in protected speech.

In none of these cases, of course, did the Court address the question of whether non-human speakers, or more specifically AI speakers, have First Amendment rights. But the Court's concern about the government as a bad actor deserving of constraint rather than about human speakers as deserving of protection bears directly on our question of whether the government's suppression of AI speech will be found to violate the First Amendment. As Kathleen Sullivan has explained, this negative view understands the Free Speech Clause as "indifferent to a speaker's identity or qualities—whether animate or inanimate, corporate or nonprofit, collective or individual. To the extent the clause suggests who or what it protects, it suggests that it protects a system or process of 'free speech,' not the rights of any determinate set of speakers."[62]

57. 136 S. Ct. 1412, 1418 (2016) ("When an employer demotes an employee out of a desire to prevent the employee from engaging in political activity the First Amendment protects, the employee is entitled to challenge that unlawful action under the First Amendment.").

58. *Id.* at 1416.

59. *Id.* at 1418.

60. *Id.*

61. *Id.*

62. Kathleen M. Sullivan, *Two Concepts of Freedom of Speech*, 124 HARV. L. REV. 143, 156 (2010).

A focus on the constitutional text as constraining the government *actor*—"*Congress* shall make no law"—adds to the plausibility of future First Amendment coverage of AI speech.[63] Courts that focus on curbing the government might well strike down suspect regulations because of improper government motives regardless of the speaker's humanness. Protecting strong AI speech from government regulation is thus consistent with negative theory's "deep skepticism about the good faith of those controlling the government."[64] This theory may even support coverage of future AI-to-AI speech, no less than AI-to-human speech, if government restriction of that speech were motivated by an impermissible desire to suppress the content or viewpoint of the speech. The negative view of the First Amendment is both exceptionally speech-friendly and speaker-agnostic.

The appeal of the negative view of the First Amendment may include its ability to dodge tricky definitional questions about when a speaker's own actions should count as "speech," by focusing instead on any nefarious government motive in suppressing the actions.[65] But in so doing, negative theory then creates a new line-drawing problem of its own. Negative theory offers no meaningful limiting principles that would permit government to regulate speech under certain conditions. It also does not elide the "what is speech" question entirely, as no free speech problem arises if a government motive is to regulate pure conduct and the law is applied in a speech-neutral way. Thus the definitional question does not disappear, but instead shifts from asking whether a particular activity is speech to asking whether the government intends to target speech. In any event, here we note the Court's increasing reliance on negative theory to justify its free speech decisions, and that such reliance, if taken seriously, adds support to the possibility of First Amendment coverage of AI speech.

In sum, both positive and negative theories support the idea of extending First Amendment coverage to strong AI. To make these claims less abstract, take the following hypothet-

63. U.S. CONST. amend. I (emphasis added).

64. *See* Gey, *supra* note 30, at 21; *see also* Elena Kagan, *Private Speech, Public Purpose: The Role of Governmental Motive in First Amendment Doctrine*, 63 U. CHI. L. REV. 413, 414 (1996) (emphasizing improper government motive as the central offense to First Amendment values).

65. *See, e.g.*, Wesley J. Campbell, *Speech-Facilitating Conduct*, 68 STAN. L. REV. 1, 8 (2016) ("[T]he government faces a heightened burden when it singles out speech.").

ical. A strong AI version of the Tolstoy-bot writes long, intricate, socially astute novels about the 2016 election cycle, inspired by the work of Leo Tolstoy.[66] The works cannot be traced in any way to a human author or programmer. In our hypothetical, they are the creative work solely of our Tolstoy-bot AI.

To a traditional democratic self-governance theorist, such novels would be covered by the First Amendment to the extent that they contribute to public discourse and political debate.[67] To Post, they would be covered both because they are public discourse, and because government restrictions on their publication would call into question for individual human citizens the legitimacy of that governance regime.[68] To Balkin, they would be protected because readers of the novels could use the books to dynamically construct a culturally situated self.[69] To marketplace-of-ideas theorists, the novels would be protected to the extent they contribute to their readers' search for "truth," knowledge, or enlightenment.[70] To autonomy theorists, the novels would be protected because interference in their publication would squelch readers' autonomy, impinging on freedom of information-gathering, self-construction, and thought.[71] And to those taking the negative view of the First Amendment, the novels would be protected from laws that arise from an illegitimate government motive, perhaps resulting from a desire to squelch social criticism (or a deep hatred of long-form literature).[72]

Only a theory that insists that speech has value solely insofar as it enhances human speakers' and not listeners' auton-

66. Joshua Barrie, *Computers Are Writing Novels: Read a Few Samples Here*, BUS. INSIDER (Nov. 27, 2014), http://www.businessinsider.com/novels-written-by-computers-2014-11 (describing a 2008 "320-page novel [that] is a variation of Leo Tolstoy's 'Anna Karenina,' but worded in the style of a Japanese author called Haruki Murakami").

67. *See supra* notes 32–33 and accompanying text (describing self-governance theories).

68. Our claim, not his. For a discussion of Post's theory of freedom of expression, see *supra* note 34 and accompanying text.

69. Our claim, not his. *See supra* note 36 and accompanying text (describing Balkin's theory of free speech).

70. To review the marketplace-of-ideas theory, see *supra* notes 42–43 and accompanying text.

71. *See supra* notes 44–47 and accompanying text (explaining autonomy theory).

72. *See supra* Part I.B (overviewing negative views of the First Amendment).

omy cuts against this grain.[73] The next Part therefore explores the role—if any—of speakers' autonomy, dignity, and emotion in current First Amendment law.

II. FIRST AMENDMENT LAW AND ITS INATTENTION TO SPEAKER HUMANNESS

First Amendment doctrine already protects speakers that are not conventionally human.[74] Corporations receive speech protection as legal persons,[75] and courts already have protected algorithmic speech (as the product of human programmers, but still).[76] Nevertheless, the biggest hurdle to protecting the speech rights of strong AI is the assertion that humanness matters, or should matter, to First Amendment rights.

Humans are the quintessential language animal.[77] We constantly construct ourselves out of layers of meaning, both produced and consumed.[78] In all of these respects, language makes us distinctive—it gives us a capacity to adapt, a fluidity "which has no parallel among other animals."[79] Humans are the *only* language animal, so realized. Human law, including constitutional law, was designed for *this* animal, and maps imperfectly onto other animals, let alone onto inanimate machines.

If language is so essential to our existence as human beings, how can it be that legal protection for the use of language does not turn on a speaker's humanness? This Part addresses why speaker humanness may not matter for purposes of our

73. *See* C. Edwin Baker, *The First Amendment and Commercial Speech*, 84 IND. L.J. 981, 997 (2009) (arguing that free speech protections should not apply to commercial speech because it "is not an exercise of freedom by morally significant flesh-and-blood individuals").

74. Massaro & Norton, *supra* note 6, at 1183–85.

75. *See, e.g.*, Citizens United v. FEC, 558 U.S. 310, 365–66 (2010).

76. Zhang v. Baidu.com, Inc., 10 F. Supp. 3d 433, 435 (S.D.N.Y. 2014) (characterizing search engine algorithmically produced results as "in essence editorial judgments about which political ideas to promote"); *see also* Langdon v. Google, Inc., 474 F. Supp. 2d 622 (D. Del. 2007); Search King, Inc. v. Google Tech., Inc., No. CIV-02-1457-M, 2003 WL 21464568 (W.D. Okla. May 27, 2003).

77. *See generally* CHARLES TAYLOR, THE LANGUAGE ANIMAL: THE FULL SHAPE OF THE HUMAN LINGUISTIC CAPACITY (2016) (describing how language shapes the human experience).

78. *Id.* at 30–38. We also locate emotional meanings in the body—words have meaning in part because they express a feeling that registers in our physical beings. *Id.* at 247.

79. *Id.* at 339.

contemporary First Amendment rights regime, even if one focuses on the expressive uniqueness of human beings.[80]

A. WHAT ABOUT AUTONOMY AND DIGNITY?

The intuition that a speaker must be human to trigger First Amendment coverage remains deeply felt for many reasons. If we protect speech to protect a speaker's autonomy and dignity, then it seems unimaginable to protect an AI that arguably has neither dignity nor autonomy interests. Even corporations, which fit less easily into constitutional garb, have humans within them.

Free speech theories that value speech for its role in furthering the autonomy of the speaker thus present significant barriers to coverage for strong AI speakers.[81] However, speaker autonomy arguments face several difficulties. First, speaker autonomy arguments must identify intrinsic qualities of moral personhood that are unique to humans. As AI advances, the gap between machines and humans may narrow in ways that weaken this uniqueness claim. Second, and more fundamentally, insisting that humans alone possess intrinsic qualities of moral personhood does not prove that those qualities *should* matter for purposes of conferring free speech rights.[82]

As discussed in Siri-ously 1.0, the personhood barrier for First Amendment protections could be overcome if (1) we change how we view protected "persons" for practical or theoretical reasons; or (2) AI came to function in ways that satisfied our personhood criteria.[83] We are seeing changes on both fronts. Free speech theory has marched steadily away from a construction of legal personhood that views speakers solely through an individual or animate lens, and now defines them in a practical, non-ontological sense.[84] Under this view, and as discussed

80. *See* Massaro & Norton, *supra* note 6, at 1178–82.

81. *Id.*

82. *See* Jason Iuliano, *Do Corporations Have Religious Beliefs?*, 90 IND. L.J. 47, 71 (2015) ("Two competing theories of personhood have dominated the philosophical literature. The first account maintains that persons are distinguished by certain intrinsic characteristics, that there is some innate substance that captures personhood . . . . The second holds that persons are distinguished by certain external characteristics. According to this account, any agent that performs in a certain manner qualifies as a person.").

83. Massaro & Norton, *supra* note 6, at 1179–82.

84. *See supra* Part I; *see also* EVAN FOX-DECENT, SOVEREIGNTY'S PROMISE 38 (2011) ("The defining hallmark of liberalism is that the ultimate unit of moral value is the individual. For law, however, the unit to which rights and

above, speech should be protected because it serves democratic self-governance, or leads to better thinking, or even better markets—not necessarily because the source of speech is a human speaker whose rights arise from innately human dignity.[85] Technological advances also may enable some AI to satisfy certain criteria for legal personhood.[86] AI enhancements can blur the line between AI rights and legal personhood by making the relevant differences between computer programs and human "programs" less stark and less significant. Speaker personhood thus is not necessarily a First Amendment trump card, even under autonomy theories.[87]

duties attach is the legal person. *The two are not the same*. An individual has legal personality, but so do a wide variety of groups, such as unions, corporations, communities, Indigenous people, and municipalities. The rule of law applies to legal persons, and not just to individuals." (emphasis added) (footnote omitted)).

85. *See supra* Part I. We note too that originalist arguments likely counsel against giving free speech protection to machines; James Madison no doubt envisioned that human speakers, not machines, would be covered by the First Amendment. Those who ratified the amendment presumably did, too. Humanness, according to this originalist argument, is a necessary predicate for First Amendment coverage. We offer two responses. First, the text of the free speech clause makes no mention of a human speaker and instead restricts only "Congress," suggesting the possibility that the original readers of the clause understood it simply to constrain government rather than to protect speakers. Second, the current Court boasts no more than two committed originalists (Justice Thomas and perhaps the newly confirmed Justice Gorsuch), and strict originalist arguments are unlikely to control its First Amendment doctrine.

86. Again, computers are increasingly self-directed or "autonomous." *See infra* note 122. Some forms of AI already are being designed in ways that may afford them enhanced emotional intelligence and other features that narrow the gap between humans and computers. *See generally* BLUEPRINT FOR AFFECTIVE COMPUTING: A SOURCEBOOK (Klaus R. Scherer et al. eds., 2010) (discussing affective computing, which is focused on theory and design of computers that can detect, respond to, and simulate human emotional states); RAFAEL A. CALVO & DORIAN PETERS, POSITIVE COMPUTING: TECHNOLOGY FOR WELLBEING AND HUMAN POTENTIAL (2014) (discussing how AI is being developed to detect nonverbal cues and otherwise be deployed to grasp and improve human self-awareness and emotional well-being); ROSALIND W. PICARD, AFFECTIVE COMPUTING (1997); THE OXFORD HANDBOOK OF AFFECTIVE COMPUTING (Rafael A. Calvo et al. eds., 2015) (providing a reference for research in the emerging field of affective computing, which addresses how affective factors influence interactions between humans and computers and how the latter may be designed with social skills). These capacities are of particular relevance to us here.

87. *See* Massaro & Norton, *supra* note 6, at 1181–82. The First Amendment is not the only area in which AI developments are pushing us to rethink settled legal expectations about human actors. Copyright protections, which would seem intuitively to vest in a human author, have long been justified in the United States predominantly by a utilitarian consideration of incentives

If anything, thinking about strong AI speech rights illustrates just how much human dignity and speaker autonomy have been downplayed or erased from the First Amendment equation. It forces us to attend to the ways in which law has ignored the "missing something" problem that drives autonomy-based objections to AI rights.

We elaborate on the potential irrelevance of speaker personhood by turning to two specific objections some might have to AI speech rights that relate to speaker humanness. First we discuss whether a speaker must possess human emotions to merit First Amendment coverage. Then we examine whether free speech rights for AI would lead inexorably to coverage of animals or other nonhuman speakers.

B. WHAT ABOUT EMOTIONS?

Emotion matters to human thought[88] as well as to legal rights and responsibilities. For example, shame is an important piece of emotional hardware that assists humans in thinking about the consequences of acts to which shame is linked.[89] Re-

for the production of creative goods. The utilitarian justification for copyright protection, combined with doctrinal development that lowers the originality threshold, may also have made room for non-human authors in that area of law (as much as the Copyright Office denies this possibility). Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, 2012 STAN. TECH. L. REV. 3, 3–9; James Grimmelmann, *Copyright for Literate Robots*, 101 IOWA L. REV. 657, 680 (2016) ("Robots that act indistinguishably from humans can also be expected to respond indistinguishably from them in response to legal pressures."). *But see* U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 306 (3d ed. 2014), http://www.copyright.gov/comp3 ("The U.S. Copyright Office will register an original work of authorship, provided that the work was created by a human being."); Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, 47 U. PITT. L. REV. 1185, 1199–1200 (1986) (arguing that computers need no incentives to generate output and thus cannot be classified as authors). Again, the entrance of AI into the legal system, like the entrance of any other disruptive technology, forces us to reconsider underlying theory and changes the conversation in ways that reveal an increasing utilitarian bent, where we look to costs and incentives or systemic governance justifications, such as restricting governmental overreach.

88. As emotion theorist Silvan Tomkins has put it, "Reason without affect would be impotent, affect without reason would be blind." SILVAN S. TOMKINS, AFFECT, IMAGERY, CONSCIOUSNESS: VOL. I: THE POSITIVE AFFECTS 112 (1962). For a compelling account of the interaction of cognition and emotion see MARTHA C. NUSSBAUM, UPHEAVALS OF THOUGHT: THE INTELLIGENCE OF EMOTIONS 454 (2001) ("Some emotions are at least potential allies of, and indeed constituents in, rational deliberation.").

89. For a discussion of the psychological literature on shame and how it may bear on law, especially on government use of shaming penalties, see

morse is a factor in determining whether a person who has violated the law should be punished, and to what extent, and to whether the harms to the victim have been properly acknowledged.[90] As Martha Nussbaum has said, "[L]aw without appeals to emotion is virtually unthinkable . . . . [T]he law ubiquitously takes account of people's emotional states."[91] Fear of repercussions—loss of status, compromised interpersonal bonds, economic losses, freedom, self-respect, even death—all factor into how law works to constrain human behavior.

Computers, at present, lack these and other relevant emotional capacities in ways that make them awkward legal rights bearers at best. AI speakers are incapable of assuming emotional responsibility for the harms that they cause. Even the most emotionally unintelligent humans surpass AI in this respect.

Emerging developments in affective computing, though, may challenge the casual assumption that AI lacks feelings.[92] For one thing, human emotions are, according to some theorists, themselves "adaptations, well-engineered software modules that work in harmony with the intellect and are indispen-

MARTHA C. NUSSBAUM, HIDING FROM HUMANITY: DISGUST, SHAME, AND THE LAW (2004) (discussing the links between disgust, shame, and law); Dan M. Kahan, *What's Really Wrong with Shaming Sanctions*, 84 TEX. L. REV. 2075, 2086–91 (2006) (amending his earlier work and concluding government shaming may have too many negative consequences to warrant its use); Toni M. Massaro, *The Meanings of Shame: Implications for Legal Reform*, 3 PSYCHOL. PUB. POL'Y & L. 645, 655–73 (1997) (discussing psychological literature on shame and its relevance to government shaming); David A. Skeel, Jr., *Shaming in Corporate Law*, 149 U. PA. L. REV. 1811, 1823–36 (2001) (discussing how shaming works and the implications for shaming sanctions in corporate arenas); James Q. Whitman, *What's Wrong with Inflicting Shaming Sanctions?*, 107 YALE L.J. 1055, 1060–68 (1998) (discussing risks of government shaming, especially that government thereby assigns punishment function to mob justice).

90. *See* AARON LAZARE, ON APOLOGY 107–14 (2004) (discussing the role of remorse in accepting responsibility for misdeeds and making reparations).

91. NUSSBAUM, *supra* note 89, at 5. *See generally* THE PASSIONS OF LAW (Susan A. Bandes ed., 1999) (collecting essays on emotions and law).

92. "[A]ffective computing" describes the interdisciplinary process of designing computer systems and devices that can recognize, interpret, simulate, and process human affects. That capacity is expanding. *See* DAVID ROSE, ENCHANTED OBJECTS: DESIGN, HUMAN DESIRE, AND THE INTERNET OF THINGS 167 (2014) (explaining that computers "can sense sound, light, touch, many kinds of movement, biometric data such as heart rate and fingerprints, liquid flow, barometric pressure, radiation, temperature, proximity, and location" as emotional cues); *see also supra* note 86 and accompanying text (describing affective computing).

sable to the functioning of the whole mind."[93] That is, the more we learn about human brains and the interplay of emotion and reason, the more humans may look like extraordinarily complex computers. Emotions also may be means to an end—"mechanisms that set the brain's highest-level goals."[94] If thinking and feeling go hand in hand in this way for humans, then a brilliant thinking machine also may be able to "feel" in ways akin to humans.

More centrally, this thought experiment reveals that speaker human attributes—including speaker emotions—are not necessary to contemporary First Amendment law. Free speech law pays extraordinarily little attention to speakers' emotional intelligence. The freedom of speech umbrella covers the least empathic speaker among us, no less than the exquisitely sensitive. Members of the Westboro Baptist Church are protected when they picket the funerals of soldiers and display signs proclaiming "God Hates Fags" and "Thank God for Dead Soldiers."[95] First Amendment protections apply to racist, homophobic, sexist, blasphemous, or otherwise cruel postings on Facebook pages or other social media sites, despite the grave emotional harms they may inflict.[96]

Courts likewise do not inquire into the emotional capacities of a corporation when it exercises its First Amendment freedoms.[97] And when the law focuses solely on bad government motives regarding speech suppression rather than on whether protected speech in fact occurred, the speaker's actual autonomy and human attributes drop out of the picture altogether. The primary focus is on the government and its antagonism towards speech, not on any actual human speaker herself or itself.

93. STEVEN PINKER, HOW THE MIND WORKS 370 (1997).

94. *Id.* at 373; *see also* George Johnson, *Consciousness: The Mind Messing with the Mind*, N.Y. TIMES (July 5, 2016), https://nyti.ms/29quHMK (referencing Princeton neuroscientist Michael Graziano as suggesting that "consciousness is a kind of con game the brain plays with itself. The brain is a computer that evolved to simulate the outside world. Among its internal models is a simulation of itself—a crude approximation of its own neurological processes").

95. Snyder v. Phelps, 562 U.S. 443, 448 (2010).

96. *See* DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE (2014) (discussing cyberharassment); *see also* Danielle Keats Citron & Helen Norton, *Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age*, 91 B.U. L. REV. 1435, 1437–39 (2011) (discussing cyber harassment's move to popular social media sites).

97. *See, e.g.*, Citizens United v. FEC, 558 U.S. 310 (2010) (making no inquiry into the emotional capacities of the Citizens United organization).

The darker aspect of a free speech theory that emphasizes boundless expressive autonomy—i.e., that it places a constitutional right in the hands of some speakers who use the right recklessly, stupidly, immaturely, coarsely, even insanely—has not gone unnoticed.[98] On the contrary, the unlikeable, irresponsible speaker who hurts others feelings is historically central to American free speech law. Early cases such as *Chaplinsky v. New Hampshire*, which defines the unprotected class of "fighting words,"[99] developed out of a sense that in the speech realm at least, hurt feelings are subordinated to both cultural pluralism and individualism.[100] We protect the speaker who feels no compunction about his wounding words because line-drawing around emotional harms is seen as anti-pluralistic.

Even if strong AI proves to be better than the worst of human speakers, and even if strong AI never matches the best of human speakers in terms of emotional intelligence and grace, AI's free speech rights may not hinge on either capacity. What will matter—at least under current theory and doctrine—is whether AI says something listeners should hear, or something that government should not be allowed to silence.

C. WHAT ABOUT OTHER NONHUMAN "SPEAKERS"?

Extending free speech coverage to AI raises powerful slippery slope concerns.[101] Would the extension of First Amendment

98. As Zechariah Chafee once said, free speech not only protects the noble dissenter, but also "loud-mouthed unattractive men whose evidence and ideas are rather worthless." ZECHARIAH CHAFEE, JR., THE BLESSINGS OF LIBERTY 113 (2d ed. 1956); *see also* Toni M. Massaro, *Significant Silences: Freedom of Speech in the Public Sector Workplace*, 61 S. CAL. L. REV. 3, 63–64 (1987) (describing some plaintiffs in public employee free speech cases as "whistle blowers, whiners, and weirdos" who not only violate workplace discourse norms in ways that prompt employer discipline, but then take the more aggressive steps of suing and claiming retaliation).

99. 315 U.S. 568, 572 (1942).

100. *See, e.g.*, Robert C. Post, *Cultural Heterogeneity and Law: Pornography, Blasphemy, and the First Amendment*, 76 CALIF. L. REV. 297, 299, 305 (1988) (characterizing a First Amendment question as "a question that involves the obligations of a legal order in a heterogeneous society comprised of diverse and competitive groups," and identifying three kinds of law: "assimilationist law strives toward social uniformity by imposing the values of a dominant cultural group; pluralist law safeguards diversity by enabling competing groups to maintain their distinct perspectives; individualist law rejects group values altogether in favor of the autonomous choices of individuals").

101. A related slippery slope concern asks what constitutional rights other than free speech might extend to AI. Again, we take care to focus only on the possibility of free speech rights for strong AIs, and not on any other set of con-

coverage to AI similarly require coverage of other nonhuman speakers?[102] At one discussion of AI speech rights, an audience member asked whether these arguments had a logical stopping point that would exclude her cat from free speech coverage.[103] Cats and other domestic animals communicate with their humans in multiple ways that may qualify as an intent to communicate reasonably understood as such, especially with respect to needs that they want the humans to fulfill. Similar and perhaps stronger arguments attach to the communication of dolphins and nonhuman primates.[104]

We offer several observations in response to these slippery slope concerns. First, the ease with which autonomous AI speakers fit into current free speech theory and doctrine says more about First Amendment theory and doctrine than it does about changes in technology, or our understanding of animals. Extending rights to AI thus might force courts and other decisionmakers to reexamine arguments that free speech and other constitutional rights should apply to other potential rights bearers, including animals.[105] We look forward to that provocative conversation, should it occur.[106]

---

stitutional rights. Indeed, not all rights are, or should be, necessarily available to all legal persons. For example, that a legal person has the right to sue and be sued—or to speak—does not necessarily mean that it has, or should have, the right to vote or a right to privacy. *See* Kent Greenfield, *In Defense of Corporate Persons*, 30 CONST. COMMENT. 309, 321 (2015) ("Of course corporations are not genuine human beings and should not automatically receive all the constitutional rights that human beings claim. At the same time . . . it is similarly obvious that corporations should be able to claim *some* constitutional rights. So which ones, and when?").

102. *See generally* FRANS DE WAAL, ARE WE SMART ENOUGH TO KNOW HOW SMART ANIMALS ARE? (2016) (discussing ways in which many animals have vast, still underappreciated, capacities that may bear on how law—including free speech law—should treat them). We write this mindful of his arresting examples of sweet potato-washing Japanese macaques, bottlenose dolphins that buoyed a stunned fellow dolphin by carrying her on their bodies and submerging their own blowholes beneath the water surface until she revived, and the surprising braininess of the octopus. *Id.* at 52, 133, 246–49.

103. Thanks to Sue Glueck for raising this question.

104. Ryan Calo asked whether arguments that the First Amendment covers AI speech also suggest First Amendment coverage for random acts of nature, like rogue waves that refigure sand on the beach to shape words that government might want to erase. Calo posed the free speech hypothetical in which waves on a beach (weirdly) formed symbols or words that humans interpreted as deeply offensive or frightening (or perhaps inspiring), and wondered whether a local government's erasure of such markings to prevent potential negative impacts would present a First Amendment problem.

105. *See* Tilikum v. Sea World Parks & Entm't, Inc., 842 F. Supp. 2d 1259, 1262 (S.D. Cal. 2012) (concluding that the Thirteenth Amendment applies

Second, a decision to extend First Amendment coverage to AI speech does not inexorably lead to coverage of other nonhuman speakers. Animal communication skills are not evolving nearly as rapidly as AI skills.[107] Nor is the evolutionary arc of animal language as amenable to human direction or crafting in the very direct ways that the arc of AI language is. AI is human-designed and profoundly and exclusively human-centered in terms of the needs it seeks to address. AI expression arguably will benefit human audiences more pointedly, pervasively, and profoundly than a cat's meows. Moreover, AI communication is deliberately supplanting human communication at nodes (think finance, telecommunication, transportation, energy, computer-assisted research, health care, and defense) that matter greatly to human well-being, safety, and knowledge.

A cat that is not in the mood to purr and leaves the room with rump aloft will have little impact on humans' fundamental information bearings. Depending on the context, a computer that refuses to interact or goes down unexpectedly may damage human interests in disastrous ways. AI communication is often designed to serve very central human information needs. To

solely to persons, not animals such as orcas); Nonhuman Rights Project v. Presti, 124 A.D.3d 1334 (N.Y. App. Div. 2015) (denying petition for writ of habeas corpus on behalf of chimpanzee on ground that the writ is only available when it would lead to immediate release from custody in New York); Nonhuman Rights Project v. Lavery, 124 A.D.3d 148, 151 (N.Y. App. Div. 2015) (rejecting argument for legal personhood for chimpanzees); Richard L. Cupp, Jr., *Cognitively Impaired Humans, Intelligent Animals, and Legal Personhood*, 68 FLA. L. REV. (forthcoming 2017) (considering but rejecting legal rights for animals on the ground that legal personhood is based on respect for dignity interests of persons as members of the human race, not on cognitive abilities); *see also* Justin Marceau, *If a Monkey Snaps a Selfie, Does He Own the Rights to His Own Photograph?*, QUARTZ (Aug. 26, 2016), http://qz.com/767163/naruto-monkey-selfie-peta; David Post, *The Monkey Selfie Is Back!*, WASH. POST: VOLOKH CONSPIRACY (Sept. 1, 2016), https://www.washingtonpost.com/news/volokh-conspiracy/wp/2016/09/01/the-monkey-selfie-is-back/?utm_term=.b360f38cf8b0 (discussing the monkey "selfie" and why he cannot be declared the "author" of the image under copyright law). *See generally* STEVEN M. WISE, RATTLING THE CAGE: TOWARD LEGAL RIGHTS FOR ANIMALS (2000) (arguing that animals are more appropriately persons, not things). For a thoughtful critique of Wise's approach that touches on the conceptual and practical difficulties of arguing by analogy from traditional rights bearers to new ones, see Richard A. Posner, *Animal Rights*, 110 YALE L.J. 527, 532 (2000) (reviewing STEVEN M. WISE, RATTLING THE CAGE (2000)) ("[C]ognition and rights-deservedness are not interwoven as tightly as Wise believes . . . .").

106. *See* DE WAAL, *supra* note 102, at 52.

107. *See* NICK BOSTROM, SUPERINTELLIGENCE 2–3 (2014) (discussing how the future rate of growth of AI technologies is unclear, and that current growth models may greatly underestimate pace of change).

take but a few compelling examples, computer-generated information is crucial to air traffic control, a great deal of medical care, and the functioning of the stock market. Much as humans might like to know what cats, dolphins, octopuses, or capuchin monkeys are really telling them, humans are not literally lost or gravely imperiled if these meanings remain mysteries.[108]

AI is designed to speak our language, and increasingly to do so in forms that look like us, walk like us, and talk like us. Animals, by contrast, generally do not participate in what humans understand to be genres of human communication.[109] If courts and other decisionmakers justify the protection of AI because of human listeners, then the value of AI speech to human listeners is arguably higher than the value of animal "speech."[110] A decision to extend First Amendment coverage to AI speech thus may not require a decision to extend such coverage to other nonhuman speakers.

Another way to understand these distinctions is to return to the doctrine. From a doctrinal perspective, extending First Amendment coverage to AI speech does not necessarily require the coverage of much of animals' output, because what animals produce often does not qualify as "speech" under the First Amendment.[111] The doctrine on expressive conduct, arising in cases about burning flags and destroying draft cards, protects conduct performed with an intent to create a particularized message that is likely to be understood.[112] Animals arguably

108. *Cf.* DE WAAL, *supra* note 102, at 234 (rejecting strong versions of human exceptionalism, and noting that "there is sound evidence that mental processes associated with consciousness in humans, such as how we relate to the past and future, occur in other species as well" but conceding that we do not yet know enough to map the consciousness differences precisely).

109. Animals have their own distinctive features, which may overlap with human characteristics in ways that make them rich subjects of study as humans continue the hard work of understanding themselves. *See* MARTHA C. NUSSBAUM, POLITICAL EMOTIONS: WHY LOVE MATTERS FOR JUSTICE 137–60 (2013) (discussing differences and similarities between animal and human compassion).

110. *See* Robert Post, *Encryption Source Code and the First Amendment*, 15 BERKELEY TECH. L.J. 713, 717 (2000) ("If the state were to prohibit the use of [film] projectors without a license, First Amendment coverage would undoubtedly be triggered. This is not because projectors constitute speech acts, but because they are integral to the forms of interaction that comprise the genre of the cinema.").

111. Similarly, if a wave washing up on the beach manages to produce random patterns, those patterns are unlikely to be considered First Amendment "speech." *See supra* note 104 and accompanying text.

112. Spence v. Washington, 418 U.S. 405, 410–11 (1974) (noting that the

rarely do this. Some animal communication, of course, is reasonably understood to create particularized messages—think: "feed me," "someone is at the door," or the more complex communication of primates. But the animal-human communication relationship remains much thinner than the AI-human communication relationship in ways that still make it easier, at least for now, to categorically exclude animal speakers from the First Amendment fold.

III. DOCTRINAL AND PRACTICAL OBJECTIONS TO AI FREE SPEECH COVERAGE

We have demonstrated the theoretical difficulty of placing AI speakers wholly outside the First Amendment. A great deal of computer speech shares similarities with the human speech that courts *already* protect, especially when we emphasize expression's value to listeners. Non-humanness does not necessarily pose any insurmountable theoretical obstacle to strong AI rights.

The law, however, is more than theory, and there are arguably a host of practical and doctrinal hurdles to protecting strong AI speech. This Part explains how both prevailing doctrine and practical concerns also fail to eliminate the possibility of First Amendment coverage for AI speech, even while they identify important challenges and questions yet to be addressed.

A. CAN AI SPEAKERS HAVE CULPABLE MENTAL STATES?

First Amendment law sometimes requires intent to cause harm (or some other culpable mental state on the part of the speaker) as a condition of imposing liability for speakers' harmful expression.[113] Courts justify such a requirement in an effort to protect valuable speech from the possible chilling effects of over-regulating less culpably motivated speakers.[114] But this

symbolism inherent in an object, the context of the situation, and an intention of expression can combine to create a protected form of expression).

113. Massaro & Norton, *supra* note 6, at 1172–85.

114. *See* Brandenburg v. Ohio, 395 U.S. 444 (1969) (holding that advocacy of illegal conduct is protected unless intentionally directed to inciting imminent illegal action); *see also* N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279–80 (1964) (holding that false assertions of fact regarding public officials are protected absent the speaker's malicious mental state); Leslie Kendrick, *Speech, Intent, and the Chilling Effect*, 54 WM. & MARY L. REV. 1633, 1640 (2013) (arguing that a stricter standard of liability would cause over-deterrence of speech).

creates a problem if the doctrine were to insulate AI speakers (but not human speakers) from liability because they lack provably culpable mental states.

Say, for example, a computer produces defamatory speech—i.e., false factual claims that damage its target's reputation—about a public official. First Amendment doctrine currently requires a showing of the speaker's actual malice before even demonstrably false political attacks may become actionable.[115] How might a court determine whether an AI speaker acted with knowledge of or reckless disregard for the falsity of its assertions, or that it "entertained serious doubts as to the truth of [its] publication"?[116] If the culpable intent cannot be shown, then liability arguably does not attach.

Rather than insulating AIs' defamatory speech from liability altogether, courts could manage these complexities by altering the doctrine to prevent an AI windfall or otherwise mitigate the harmful effects of defamatory AI speech.[117] Even contemporary free speech doctrine, despite its growing emphasis on formalism over nuance, offers ways to address important regulatory concerns.[118]

115. *See Sullivan*, 376 U.S. at 279–80; *see also* Meg Leta Ambrose & Ben M. Ambrose, *When Robots Lie: A Comparison of Auto-Defamation Law*, *in* IEEE WORKSHOP ON ADVANCED ROBOTICS AND ITS SOCIAL IMPACTS 56, 57 (2014), http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7020980 ("[R]obots also serve as information and communication participants that may cause social unrest and individual harm which current legal regimes will find challenging. . . . [E]merging robotic systems pose novel issues concerning defamation due to their unprecedented ability to experience the world and potential to communicate that experience to humans.").

116. St. Amant v. Thompson, 390 U.S. 727, 731 (1968).

117. For an analysis of the difficulties that already exist with mapping defamation law onto new social media, see LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE 164–85 (1999); *see also* Lyrissa Barnett Lidsky & RonNell Andersen Jones, *Of Reasonable Readers and Unreasonable Speakers: Libel Law in a Networked World*, 23 VA. J. SOC. POL'Y & L. 155, 156 (2016) (discussing the unique implications that social media has for the malice rule); Frank Pasquale, *Reforming the Law of Reputation*, 47 LOY. U. CHI. L.J. 515, 527–38 (2015) (discussing implications of new technologies for reputation law); Robinson Meyer, *Did Facebook Defame Megyn Kelly?: Which Is Another Way of Asking: Can a Bot Commit Libel?*, THE ATLANTIC (Aug. 30, 2016), http://www.theatlantic.com/technology/archive/2016/08/did-facebook-defame-megyn-kelly/498080 (discussing Facebook's potential liability for a shift in the algorithm for its trending feature, which promoted a fake story that claimed Megyn Kelly endorsed Hillary Clinton for President).

118. *See* Toni M. Massaro, *Tread on Me!*, 17 U. PA. J. CONST. L. 365, 369–82 (2014) (discussing ways in which the Roberts Court has moved toward greater formalism in its approach to free speech, but noting the many ways in which existing doctrine still offers judges significant and necessary flexibility

Indeed, our legal tradition has long had to find ways to manage the challenges presented by new machines and other technology. Copyright law provides one such example of law's adaptation to technology at the doctrinal edges. The concept of vicarious liability, with its focus on a potentially liable person's benefit from and control of a copyright infringer, was initially poorly suited for the age of mass filesharing.[119] Software distributors benefit from filesharing in more indirect ways than their physical world counterparts: through online advertisements, for example, instead of rent or direct kickbacks. The element of control presents similar doctrinal challenges: software distributors can either easily control all users, exposing themselves to broad liability and accruing what can be high monitoring costs, or can create situations of willful blindness. Over time, courts have (admittedly with serious struggles) figured out doctrinal ways to ascribe secondary liability to software distributors.[120] The sheer scale of filesharing, balanced against fears of chilling technological development, resulted not in a refusal to apply copyright law to new technologies, but in doctrinal development that made room for complicated debates about overarching policy concerns. Courts similarly may adapt First Amendment doctrine on its edges to address the challenges posed by AI speakers' defamatory speech, even absent a provably culpable mental state.

to address context-specific concerns).

119. Fonovisa, Inc. v. Cherry Auction, Inc., 76 F.3d 259, 261–62 (9th Cir. 1996) (explaining that the "concept of vicarious copyright liability . . . [is] an outgrowth of the agency principles of respondeat superior").

120. *See* Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005) (holding filesharing software company Grokster liable for inducing copyright infringement); Sony Corp. v. Universal City Studios, Inc., 464 U.S. 417, 442 (1984) (explaining that for its maker to escape secondary liability, a technology must be "capable of commercially significant non-infringing uses"); A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001) (holding filesharing software company Napster liable for user copyright infringement). *But see* Am. Broad. Co. v. Aereo, Inc., 134 S. Ct. 2498, 2516 (2014) (Scalia, J., dissenting) ("It will take years, perhaps decades, to determine which automated systems now in existence are governed by the traditional volitional-conduct test and which get the Aereo treatment. (And automated systems now in contemplation will have to take their chances)."); Bruce E. Boyden, Aereo *and the Problem of Machine Volition*, 2015 MICH. ST. L. REV. 485, 499–505 (discussing how technology muddies issues of responsibility in copyright infringement cases).

B. CAN AIS SUE OR BE SUED?

Some practically minded skeptics have pointed out the difficulties in determining how an AI speaker could sue to protect its free speech rights, or be sued and punished for its impermissibly harmful speech.[121] Here we offer an introduction to how an AI might sue or be sued, though we recognize that significant challenges remain to be addressed.

Samir Chopra and Laurence White have contributed greatly to efforts to think about how to operationalize legal rights and duties for AI.[122] We again draw on their work in considering key practical questions that would arise should courts extend First Amendment coverage to AI speakers.[123] Chopra and White note, and we agree, that AI need not have identical rights and obligations to humans.[124] Additionally, legal frameworks can be developed that would enable AI to sue and be sued. The existing category of "legal persons" already includes entities that hold a variety of legal (including constitutional) rights and duties—such as corporations, unions, municipalities, and even sailing vessels.[125] These rights-holders are legal persons, but do not possess human status.

121. *See, e.g.*, Solum, *supra* note 46, at 1248–52.

122. *See generally* SAMIR CHOPRA & LAURENCE F. WHITE, A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS (2011) (suggesting a framework in which artificial agents are given legal personhood).

123. *See* Massaro & Norton, *supra* note 6, at 1180–81.

124. For example, courts have treated corporations as holding "derivative" First Amendment rights to speak in ways that inform and benefit natural persons, rather than holding rights for their own sake. *See, e.g.*, Margaret M. Blair & Elizabeth Pollman, *The Derivative Nature of Corporate Constitutional Rights*, 56 WM. & MARY L. REV. 1673, 1678 (2015). Courts might similarly treat strong AI speakers only as derivative rights-holders rather than holding rights indistinguishable from those held by natural persons.

125. CHOPRA & WHITE, *supra* note 122, at 157; *see also* Massaro & Norton, *supra* note 6, at 1180–81 (discussing rights and duties of entities that are not moral, natural, or human persons); Solum, *supra* note 46, at 1238–39 (explaining how the term "person" is often used as synonymous with "human" for many purposes, but "[t]he question whether an entity should be considered a legal person is reducible to other questions about whether or not the entity can and should be made the subject of a set of legal rights and duties. The particular bundle of rights and duties that accompanies legal personhood varies with the nature of the entity. Both corporations and natural persons are legal persons, but they have different sets of legal rights and duties"). For interesting work speculating about the future of Bitcoin and how "independently wealthy software" might sue and be sued, see Shawn Bayern, *The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems*, 19 STAN. TECH. L. REV. 93 (2015); Shawn Bayern, *Of Bitcoins, Independently Wealthy Software, and the Zero-Member LLC*, 108 NW. U. L. REV. ONLINE 257 (2014).

As applied to AI, the first step in the progression of legal rights might be to treat AIs as dependent, not independent, legal persons. This would permit their owners (or those who direct or are assigned responsibility for them) to be sued for their allegedly harmful or otherwise illegal actions. As Chopra and White observe, "If legal systems can accord dependent legal personality to children, adults who are not of sound mind, corporations, ships, temples, and even idols, there is nothing to prevent the legal system from according this form of legal personality to artificial agents."[126] Courts and regulators could follow precedent that currently applies to corporations to establish these legal arrangements.

The more complex and controversial step would be for AI to be granted wholly independent legal personhood. Chopra and White outline the criteria for independent legal personhood as follows: an independent legal person must have intellectual capacity and rationality; the ability to understand legal obligations; susceptibility to punishment and enforcement; the ability to form contracts; and the ability to control money.[127] Sophisticated AI could satisfy these criteria.[128] Moreover, satisfaction of the theoretical criteria for independent legal personhood would enable those injured by AI speakers to pursue legal remedies against AI.

But this does not mean these notions are static or that AI may not close the gap between human cognition and machine thinking and planning. Law could take into account the inanimate nature of the legal actor in crafting liability rules, in framing the elements of torts such as defamation, or in defining constitutional defenses to liability.[129] With such adjustments,

More of these issues surely are nigh, as blockchain technology (on which Bitcoin depends) makes its way to trade finance and elsewhere. *See, e.g.*, Kim S. Nash, *Dun & Bradstreet Tests Blockchain for Trade Finance*, WALL ST. J. (Oct. 14, 2016), http://on.wsj.com/2ekRlct.

126. CHOPRA & WHITE, *supra* note 122, at 160. For example, AI might be required to register and meet capital requirements necessary to meet its financial duties and enable those who contract with AI to be on fair notice of AI's economic capacities. *Id.* at 161–62.

127. *Id.* at 162–63.

128. *Id.* at 162–71.

129. *See id.* at 153–71, 186–91 (discussing various means by which strong AI could be held legally accountable, including for damages); *see also* Ambrose & Ambrose, *supra* note 115 (discussing possible adjustments in defamation law as AI becomes more autonomous). The European parliament is ahead of the curve on these matters, and has suggested regulations be drafted that may include a form of "electronic personhood" for AI. *See* Alex Hern, *Give Robots*

AI could be held legally accountable for its harms. Moreover, as we already are seeing with weak AI that can generate news stories that include false statements,[130] the risk of defamatory harm from strong AI is likely to be profound. Courts will be motivated to redress the harms and to make doctrinal moves necessary to do so. How might judicial remedies be crafted and enforced in such cases? AI cannot be imprisoned; but of course neither can corporations.[131] Instead, an AI might be physically disabled, barred from future participation in certain economic transactions, deregistered, or have its assets seized.[132] Shallow or empty-pocket AI actors may exist and thus limit the usefulness of damage remedies, but this is true of many judgment-proof human actors as well. Injunctive and declaratory relief could be invoked against the AI speaker, just as it is invoked against human actors. Again, adjustments within the doctrine on speech protection could be made to account for the fact that the speaker is a computer, not a human being.

As for rights enforcement, we can imagine scenarios—especially given the interdependence of AI rights and listener-based interests—in which AI legal rights organizations or other humans interested in protecting valuable AI communication would step in to assist AI with assertion of AI rights. Specialized AI lawyers could help implement legal rights and remedies. Third-party standing rules that apply in federal court might allow interested human parties to assert the AI rights, along with their own. The more important the information produced and controlled by AI becomes, the more likely it will be that legal means of enforcing AI speech rights and responsibilities will be developed.

The question in all of these cases would be how to respond to the rights and remedies needs, given the functionality and capacities of the AI speaker. As the latter evolve, this would affect how rights and remedies are crafted and enforced. In fact, considering the range of available punishments for AI might lead to discussions of the theoretical purposes of punishment—

"Personhood" Status, *EU Committee Argues*, THE GUARDIAN (Jan. 12, 2017), https://www.theguardian.com/technology/2017/january/12/give-robots-personhood-status-eu-committee-argues.

130. *See* Tim Adams, *And the Pulitzer Goes to . . . a Computer*, THE GUARDIAN (June 28, 2015), https://www.theguardian.com/technology/2015/jun/28/computer-writing-journalism-artificial-intelligence.

131. *See* CHOPRA & WHITE, *supra* note 122, at 167 (comparing punishment of AI to that available for a corporation).

132. *Id.* at 167–68.

retribution, rehabilitation, constraint—much as considerations of AI speech rights have led us to reexamine free speech theory here.

Key to these practical concerns is that humanness is not essential to *legal* personhood—even if, as surely will be the case, human needs may inspire the move to AI legal personhood, and even if humans may be necessary aids to legal enforcement of AI rights. AI rights and remedies thus may be enforceable.

IV. POTENTIAL LIMITS ON AI FREE SPEECH COVERAGE

Among the most powerful objections to the notion of extending speech protections to strong AI is that such extension lacks limiting principles. We agree that the hardest problem lies here: If AI speech is covered by the First Amendment, how can the government perform important regulatory functions in the name of humans without running afoul of the First Amendment?

This puzzle produces the greatest intellectual yield of the AI thought experiment. The claims we advance are evidence of an *existing* slippage problem: the Court's contemporary free speech theory and doctrine *already* make it difficult to articulate convincing limiting principles. By emphasizing either a negative view of the First Amendment that seeks to constrain the government's dangerous ability to regulate speech[133] or positive views that emphasize the value of speech to its listeners,[134] current doctrine supports the coverage of speech regardless of its nontraditional source or form. Increasingly expansive First Amendment theories and practice have already evolved in a manner that permits further coverage extensions that may

133. United States v. Alvarez, 132 S. Ct. 2537, 2543–44 (2012) (restating the rule from *Ashcroft v. ACLU* that the First Amendment prohibits government restriction of speech); United States v. Stevens, 559 U.S. 460, 468–70 (2010) ("[A]s a general matter, the First Amendment means that government has no power to restrict expression because of its message, its ideas, its subject matter, or its content.").

134. *See* Citizens United v. FEC, 558 U.S. 310, 392–93 (2010) (Scalia, J., concurring) ("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 . . . ."); First Nat'l Bank of Bos. v. Bellotti, 435 U.S. 765, 777 (1978) ("The inherent worth of the speech in terms of its capacity for informing the public does not depend upon the identity of its source, whether corporation, association, union, or individual.").

seem exceedingly odd, counterintuitive, or even dangerous.[135] Current free speech law—which many now argue has invited deployment of the First Amendment as an antiregulatory tool[136]—is the source of this challenge, rather than any technological change past, present, or future.

The choice to extend the First Amendment to cover new sources and styles of speech inevitably involves both promise and peril. Extending constitutional protection to private activity is often a double-edged liberty sword: the greater the power of the private actor, the greater the risk that freedom for that actor will constrain the freedom of other, less powerful actors.[137]

If courts do extend First Amendment coverage to AI speakers, this move surely will invite calls for limiting principles. As this Part explains, contemporary free speech doctrine contains means by which courts could respond to dangers posed by AI speakers, though each presents its own difficulties. In short, in-

135. *See* Christopher Robertson, *The Tip of the Iceberg: A First Amendment Right To Promote Drugs Off-Label*, 78 OHIO ST. L. REV. (forthcoming 2017) (discussing slippery slope implications of the increased use of free speech law to block FDA legislation); Frederick Schauer, *The Politics and Incentives of First Amendment Coverage*, 56 WM. & MARY L. REV. 1613, 1629 (2015) ("The opportunistic lawyer or client seeking a way of fighting against some form of regulation or prosecution can now have increased confidence that an argument from the First Amendment will not be received with political scorn or doctrinal incredulity."); Philip J. Weiser, *Law and Information Platforms*, 1 J. TELECOMM. & HIGH TECH. L. 1, 33–35 (2002) (discussing concerns about how information industries may invoke free speech coverage in ways that may undermine regulatory goals).

136. For a sampling of the many thoughtful commentators who have discussed this trend, see Leslie Kendrick, *First Amendment Expansionism*, 56 WM. & MARY L. REV. 1199, 1200 (2015) ("[W]hat has been called First Amendment opportunism, where litigants raise novel free speech claims that may involve the repackaging of other types of legal arguments [lead to] First Amendment expansionism, where the First Amendment's territory pushes outward to encompass ever more areas of law."); Robert Post & Amanda Shanor, *Adam Smith's First Amendment*, 128 HARV. L. REV. F. 165, 167 (2015) ("It is no exaggeration to observe that the First Amendment has become a powerful engine of constitutional deregulation. The echoes of *Lochner* are palpable [and likely to be heard] in matters ranging from public health to data privacy."); Rebecca Tushnet, *COOL Story: Country of Origin Labeling and the First Amendment*, 70 FOOD & DRUG L.J. 25, 26 (2015) ("[T]he First Amendment has become the new *Lochner*, used by profit-seeking actors to interfere with the regulatory state in a way that substantive due process no longer allows.").

137. *See, e.g.*, Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search*, 93 CORNELL L. REV. 1149, 1191 (2008) (discussing the role search engines play as choke points to access to information).

terpreting the First Amendment to cover strong AI speakers would not necessarily mean that human needs no longer matter to First Amendment analysis. In fact, such an interpretation may inspire more careful reflection about how to define and mitigate the harmful effects of covered speech, while preserving its manifold benefits.

A. COVERING AI SPEECH BUT NOT ITS CONDUCT

The First Amendment covers speech, but not conduct. Even under the most generous listener-centered view of the First Amendment, government regulation of pure conduct triggers no freedom of speech problem, and typically triggers mere rational basis scrutiny. The government's regulation of conduct that also has expressive qualities normally triggers a form of intermediate scrutiny.[138] Current doctrine also poses a speaker intent threshold that must be crossed: in general, only conduct that is intended by the actor to communicate and that is reasonably understood by onlookers to communicate normally qualifies as "speech" for constitutional purposes.[139]

AI behavior—like human behavior—may fall under all of these categories: it may be speech, expressive conduct, or pure conduct. First Amendment protections also extend beyond speech and expressive conduct, to acts integral to well-accepted communications media,[140] and to corollary or penumbral rights necessary for fundamental First Amendment protections.[141] Which category applies to a particular AI behavior will depend both on the specific behavior in question and the context in which government seeks to regulate it.[142]

138. *See* United States v. O'Brien, 391 U.S. 367, 381–82 (1968).

139. Spence v. Washington, 418 U.S. 405, 410–11 (1974) (per curiam). We say "normally" because we recognize that strong versions of a listener-centered approach to speech threaten to swallow up *Spence*'s speaker intent requirement. If the coverage focus is *solely* on output usefulness to the listener, then it becomes harder to see why speaker intent (like speaker shame, speaker dignity, or speaker humanness) matters at all. All useful information, however created, should under current theory trigger the First Amendment. The AI problem, though, does not *create* these issues; it *highlights* them and thus might prompt courts to walk back this potentially boundless aspect of speech coverage, or seek to mitigate its effects in the ways we identify, or in other ways we may have missed.

140. *See generally* Post, *supra* note 110 (discussing how First Amendment protections could extend to encryption source code used to communicate ideas).

141. Griswold v. Connecticut, 381 U.S. 479, 482–83 (1965) ("Without those peripheral rights, the specific rights would be less secure.").

142. *See* Balkin, *Information Fiduciaries*, *supra* note 36, at 1194 (noting

Because the speech/conduct distinction may provide a reason to deny First Amendment protection to some of what algorithms and machines produce, it already has triggered significant scholarly examination. Scholars have concluded that while some machine outputs produced by present-day AI and other technologies constitute speech or should otherwise receive protection, other outputs do not.[143] Strong AI information practices will similarly vary, and thus so will the determination of when AI information products (from computer code to 3D-printed objects to music to carrying the information content of others) are speech versus conduct.[144]

that free speech treatment of information practices might vary depending on whether the practice involves data collection, analysis, use, disclosure, or sale); Margot E. Kaminski, *Drone Federalism: Civilian Drones and the Things They Carry*, 4 CAL. L. REV. CIR. 57, 62–63 (2013) (discussing constitutional implications of data recording); Seth F. Kreimer, *Pervasive Image Capture and the First Amendment: Memory, Discourse, and the Right To Record*, 159 U. PA. L. REV. 335, 408–09 (2011) (arguing that the First Amendment protects a right to record).

143. *See* Massaro & Norton, *supra* note 6, at 1186–88 (summarizing these arguments). *But see* NEIL RICHARDS, INTELLECTUAL PRIVACY: RETHINKING CIVIL LIBERTIES IN THE DIGITAL AGE 84–90 (2015) (critiquing the "data is speech" argument on grounds that it asks the wrong question and risks dooming too much worthy regulation); Kyle Langvardt, *The Doctrinal Toll of "Information as Speech*,*"* 47 LOY. U. CHI. L.J. 761, 790–801 (2016) (critiquing the doctrinal trend of treating computer code and other code-dependent technologies as speech for First Amendment purposes on the ground that this coverage will have the eventual effect of diluting protection of core First Amendment speech). This obviously is not a new challenge, as free speech theorists have long struggled over the evasive (and ultimately vanishing) line between speech and conduct, even as they recognize the line still matters. The "computer outputs-as-speech versus conduct" debate provides a modern application of an enduring and well-rehearsed characterization problem that is central to free speech theory and practice.

144. An example of how context matters arose in the recent controversy over whether broadband providers could be treated as common carriers. U.S. Telecomm. Ass'n v. FCC, 825 F.3d 674 (D.C. Cir. 2016). The court rejected the compelled speech objections of the providers on the ground that the providers covered by the regulation exercised no editorial control over the content, and thus were not engaging in speech themselves. *Id.* at 741. The court stated as follows:

> The Supreme Court has explained that the First Amendment comes "into play" only where "particular conduct possesses sufficient communicative elements," that is, when an "intent to convey a particularized message [is] present, and in the surrounding circumstances the likelihood [is] great that the message would be understood by those who viewed it," [t]he absence of any First Amendment concern in the context of common carriers rests on the understanding that such entities, insofar as they are subject to equal access mandates, merely facilitate the transmission of the speech of others rather than engage in speech in their own right.

We do not distinguish among these products here. We merely note that covering AI speech under the First Amendment will not insulate all of its outputs as speech, any more than treating humans as rights-bearing speakers converts all human behavior to speech, or insulates all of their speech outputs from government regulation.

B. PROTECTING LISTENERS AND OTHER HUMANS FROM HARMFUL AI SPEECH

Even when an AI information product is covered as speech, the government still may be able to regulate in the name of certain kinds of harms. In other words, even when courts recognize speech as *covered* by the First Amendment, this does not always mean that courts will *protect* such speech from the government's regulation.[145] Courts will likely bring the constitutional hammer down differently on some AI informational products than on others, based on the type of information product, context, and the nature of the harms at stake.[146] Courts might be persuaded that government regulation of covered AI speech meets intermediate and sometimes even strict scrutiny where the harms of AI speech appear grave enough and the regulations are sufficiently tailored.

A number of thoughtful commentators have documented the potential harms of algorithmic speech products, due to computers' jaw-dropping speed and reach.[147] These harms include deception, coercion, and discrimination. Strong AI will only deepen these concerns.

Id. (citations omitted); *see also* Stuart Minor Benjamin, *Transmitting, Editing, and Communicating: Determining What "Freedom of Speech" Encompasses*, 60 DUKE L.J. 1673, 1689–91 (2011) (arguing that mere transmission is not speech).

145. *See supra* note 9 and accompanying text (discussing the distinction between speech that is covered and how much protection speech will receive).

146. *See, e.g.*, Margot E. Kaminski, *Privacy and the Right To Record*, 97 B.U. L. REV. (forthcoming 2017) (manuscript at 4) (on file with authors) ("My central claim is that the contours of the protected right to record are defined by the privacy harms that the right potentially causes. Understanding the right to record is possible only by properly articulating the privacy interests at stake. This claim stands even as the Supreme Court in recent cases has repeatedly disavowed balancing speech against other nonspeech harms.").

147. *See, e.g.*, Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 790–96 (2015) (describing how robots' speech may include fraud, manipulation, and invasions of privacy); Wu, *supra* note 12, at 1496–1503 (describing the range of potential harms of computer-generated speech that invite regulation).

For one thing, strong AI will have social valence: AI may appear to be a real social actor in ways that trigger and manipulate intrinsic human reactions.[148] At the same time, even strong AI will likely still lack human emotions, sensitivity to social nuance, or the ability to feel shame. A listener may intuitively trust AI as though it were a human actor, without understanding the differences or risks at hand. The field of Human Robot Interaction (HRI) specifically studies how humans react to robots, and how such reactions may be deliberately triggered by machine design.[149]

As discussed in Part II.B, an AI speaker's lack of emotions may not prevent its coverage under the First Amendment.[150] Nevertheless, AIs' inability to internalize complex social cues or feel shame or equivalently binding emotions might lead to a greater government interest in regulating AI speech, based on a greater likelihood, or greater impact, of its harms to listeners. When Hello Barbie[151] tries to get a child to buy her a new outfit, the government's regulatory interest might well take into account both Barbie's canny ability to manipulate the child's attachment and emotions, and Barbie's inhuman inability to feel any guilt about doing so. Combined, these features may mean that Hello Barbie as a speaker may threaten significant and categorically distinctive harm to a child, and that the harm she causes over time may be different in degree or even—if such interactions, for example, prove to be physiologically addictive by

148. Calo, *supra* note 16, at 545 (discussing how humans react to anthropomorphic robots).

149. Laura Dattaro, *Bot Looks Like a Lady*, SLATE (Feb. 4, 2015), http://www.slate.com/articles/technology/future_tense/2015/02/robot_gender_is_it_bad_for_human_women.html; Margot E. Kaminski, *What the Scarlett Johansson Robot Says About the Future*, SLATE (Apr. 7, 2016), http://www.slate.com/articles/technology/future_tense/2016/04/what_the_scarlett_johansson_robot_says_about_the_future.html ("Robot designers know we respond to anthropomorphic features . . . . They study the ways, both for good and for bad, that robot design can affect or elicit human behavior. In one study, men were more likely to donate money to a female robot. In another, users disclosed more or less information about dating, based on whether a robot was male or female. This is no doubt true of race, as well; most robots currently have a Eurocentric design.").

150. *See supra* notes 88–100 and accompanying text (discussing the fact that emotional capacity of the speaker is not part of First Amendment analysis).

151. *See* Irina D. Manta & David S. Olsen, *Hello Barbie: First They Will Monitor You, Then They Will Discriminate Against You. Perfectly*, 67 ALA. L. REV. 135, 136–37, 179 n.232 (2015) (discussing how built-in hardware and software permits these devices to interact with other devices).

design—in kind from those that a human speaker might inflict.[152]

Of course, were courts to consider seriously this harm-on-stilts aspect of AI speech, they also may be compelled to revisit their profound agnosticism about comparably devastating effects when the speaker is a shameless or exceptionally cruel human. This is yet another way in which AI speech problems can shed new light on current doctrine and theory.

Adding to the potential harms of AI speech, computer programs are known to be far from neutral sources or decision-makers. Bias and discrimination may be built into programs, whether intentionally or accidentally; technology is not value-neutral, but value-based.[153] The possibility that AI speech threatens significant harms to human listeners through such discrimination suggests a strong governmental interest in regulation. Much as free speech doctrine has made room for the regulation of harassing speech that deprives its targets of employment and educational opportunities, so too may it make room for regulation of AI speech to further the end of preventing discrimination.[154] Judicial limits on AI speech will depend on what AI speech does (and how, and to whom).

Courts already attend to the harms that expression poses to listeners, as current doctrine recognizes a number of speech environments in which listeners' First Amendment interests are paramount. Listeners' interests can justify expression's

152. *See, e.g.*, Zoe Kleinman, *Are We Addicted to Technology?*, BBC (Aug. 31, 2015), http://www.bbc.com/news/technology-33976695.

153. *See* Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330 (1996); *see also* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 29–35 (2015) (expressing concerns about data-driven society, given the opacity of data's origins and destinations, the risk of bias and error within them, and possible cascade effects if information in one piece of software is repeated in systems throughout the economy); Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677 (2016); Kim, *supra* note 19 (noting that data and algorithms can discriminate); Kate Crawford, *Artificial Intelligence's White Guy Problem*, N.Y. TIMES (June 25, 2016), http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html ("Sexism, racism and other forms of discrimination are being built into the machine-learning algorithms that underlie the technology behind many 'intelligent' systems that shape how we are categorized and advertised to.").

154. *See, e.g.*, Jack M. Balkin, *Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2298 (1999) (discussing how the law should adapt to incentivize employers to make the workplace less hostile to less-powerful employees).

First Amendment coverage, where such coverage serves listeners' autonomy, enlightenment, and self-governance interests. But sometimes listeners' interests call for the regulation—not just the protection—of speech. Courts thus might uphold narrowly tailored restrictions on AI speech to privilege human listeners' interests in informed choices, or in avoiding the harms of coercion, deception, or discrimination.[155] If coverage of AI speech is based in large part on listeners' interests in that speech, then the government may be able to regulate the speech to protect human listeners when the interests of the AI speaker and human listener conflict.

As an example, recall that the Court has explained that commercial speech is worthy of First Amendment protection because of its informational value to consumers as listeners.[156] Commercial speech that is false, misleading, or related to an illegal activity frustrates listeners' informational interests and thus can be banned outright without running afoul of the First Amendment.[157] For related reasons, courts have permitted the government substantial leeway to require commercial speakers to make truthful disclosures because such disclosures often serve listeners' informational interests.[158]

As another example, courts often permit the regulation of professionals' speech to their patients and clients by prohibiting lies and misrepresentations to such listeners and by requir-

155. *See* Norton, *supra* note 14, at 55–60 (explaining how First Amendment theory and doctrine often support the content-based regulation of speech in certain relationships where listeners experience information or power disadvantage or are otherwise comparatively vulnerable).

156. Zauderer v. Office of Disciplinary Counsel of the Supreme Court of Ohio, 471 U.S. 626, 651 (1985) ("[T]he extension of First Amendment protection to commercial speech is justified principally by the value to consumers of the information such speech provides."); Va. State Bd. of Pharmacy v. Va. Citizen Consumer Council, Inc., 425 U.S. 748, 763–64 (1976) (emphasizing the value of "the free flow of commercial information" to individual consumers and the public more generally).

157. Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n of N.Y., 447 U.S. 557, 562–64 (1980).

158. *Zauderer*, 471 U.S. at 651 (holding that disclosure requirements that are reasonably related to the State's interest in preventing consumer deception do not violate commercial speakers' First Amendment rights). The Court applies intermediate scrutiny to laws regulating truthful and non-misleading commercial speech on the premise that such speech—although still of relatively low value—can helpfully inform individuals about their choices in the commercial realm. *Id.*

ing truthful disclosures of those listeners' options and risks.[159] A number of commentators have proposed similar approaches in the information technology context, suggesting the regulation of those information entities to which internet audiences entrust important information or functions.[160] Courts also have protected listeners under the "captive audience" doctrine, which applies when listeners cannot avoid or escape speech.[161]

Listener-based justifications would not mean limitless government power to adopt the kinds of speech regulation the Court historically has found most suspect. The government might be tempted to censor critical or inconvenient AI speech for malign motives that have long invoked First Amendment concerns. In those cases, listener interests and negative theory would align to bar the government's efforts to target AIs' viewpoint for censorious purposes, as is the case when the government conducts viewpoint-based regulation of human speech.

Drawing the line between these two scenarios—when listener and AI speaker interests align, and when they are in tension—will at times be difficult. Indeed, our AI thought problem exposes this difficulty vividly. Contemplating the possibility of

159. *See* PAUL HORWITZ, FIRST AMENDMENT INSTITUTIONS 248–50 (2013). The role of the First Amendment as a brake on occupational licensing has been a source of considerable recent scholarly attention. *See, e.g.*, Marc Jonathan Blitz, *Free Speech, Occupational Speech, and Psychotherapy*, 44 HOFSTRA L. REV. 681 (2016); Claudia E. Haupt, *Professional Speech*, 125 YALE L.J. 1238 (2016); Paul Sherman, *Occupational Speech and the First Amendment*, 128 HARV. L. REV. F. 183 (2015); Eugene Volokh, *Professional-Client Speech and the First Amendment*, WASH. POST: VOLOKH CONSPIRACY (Nov. 20, 2015), https://www.washingtonpost.com/news/volokhconspiracy/wp/2015/11/20/professional-client-speech-and-the-first-amendment.

160. *See, e.g.*, Balkin, *Information Fiduciaries*, *supra* note 36, at 1186 ("Because of their special power over others and their special relationship to others, information fiduciaries have special duties to act in ways that do not harm the interests of the people whose information they collect, analyze, use, sell, and distribute. . . . And because of their different position, the First Amendment permits somewhat greater regulation of information fiduciaries than it does for other people and entities."); Grimmelmann, *supra* note 12, at 903–04 (describing the "inescapable information asymmetry between users and search engines" that should be understood to trigger duties of loyalty and care); Neil Richards & Woodrow Hartzog, *Taking Trust Seriously in Privacy Law*, 19 STAN. TECH. L. REV. 431, 462 (2016) ("Trust in information relationships requires an affirmative obligation of honesty to correct misinterpretations and to actively dispel notions of mistaken trust.").

161. Caroline Mala Corbin, *The First Amendment Right Against Compelled Listening*, 89 B.U. L. REV. 939, 946 (2009) ("[T]he question is not whether the audience can avoid the message by leaving a particular location . . . but whether they should have to.").

AI speakers may lead to a more careful identification of the role of listeners' interests in all free speech puzzles, not just those in which AI is the speaker.

Finally, our AI thought experiment illuminates a long-central and recurring problem in First Amendment doctrine: how to distinguish speech-related from non-speech-related harms.[162] In the case of expressive conduct, such as draft card-burning, the government can often regulate expression as long as it aims the regulation at a non-speech-related harm, and adequately tailors its attempt.[163] For example, the government cannot regulate flag-burning to prevent "the expression of an idea simply because society finds the idea itself offensive or disagreeable," but could regulate burning flags and other objects where fire causes physical harm to others.[164]

Many of the information products that AI produce likely will be characterized not as pure speech, but as expressive conduct or something similar, much as computer code has been subjected to forms of intermediate scrutiny to distinguish protected expression from its regulable non-speech impacts.[165] If the government attempts to regulate such mixed-impact information products, then the question becomes what counts as a speech-related (and thus content-based and censorious) motive, and what does not. Does privacy count as a speech-related interest? Does cybersecurity? Can the government prevent Hello Barbie from doing a particular dance, or making particular

162. *See* TRIBE, *supra* note 10; Brown, *supra* note 10; *see also* Texas v. Johnson, 491 U.S. 397, 407 (1989) ("[W]here 'speech' and 'nonspeech' elements are combined in the same course of conduct, a sufficiently important governmental interest in regulating the nonspeech element can justify incidental limitations on First Amendment freedoms, we have limited the applicability of *O'Brien*'s relatively lenient standard to those cases in which 'the governmental interest is unrelated to the suppression of free expression.'" (citations omitted)).

163. United States v. O'Brien, 391 U.S. 367, 381–82 (1968).

164. *Johnson*, 491 U.S. at 414; *id.* at 403 ("If the State's regulation is not related to expression, then the less stringent standard we announced in *United States v. O'Brien* for regulations of noncommunicative conduct controls. If it is, then we are outside of *O'Brien*'s test, and we must ask whether this interest justifies Johnson's conviction under a more demanding standard." (citations omitted)).

165. *See, e.g.*, Andrea M. Matwyshyn, *Hacking Speech: Informational Speech and the First Amendment*, 107 NW. U. L. REV. 795, 816 (2013) ("[C]ourts have sometimes turned to an intermediate standard of scrutiny arising out of *United States v. O'Brien* as the basis of an analytical framework in cases involving code.").

movements, or require particular elements in her design in the name of preventing injury?[166]

This is, again, not an AI-specific problem. Recent cases have complicated the line between speech-related (or content-based) regulation and permissible non-speech-related regulation.[167] Yet the government clearly can regulate expressive conduct under intermediate scrutiny to prevent physical harms to others. Extending First Amendment coverage to AI speech thus would not disable the government from regulating such speech to address physical harms, such as crashing into others, causing particularly loud noises, or otherwise putting users at physical risk. Current doctrine suggests, however, that naming the subject of regulation (e.g., "this regulation applies to pharmaceutical companies' AI") poses the risk that the regulation will be characterized as content-based, and thus inappropriately censorious.[168] Other recent doctrine, however, suggests the government can name and thus target particular physical spaces for regulation (like the physical areas around abortion providers or embassies), when expressive conduct or even pure speech negatively impacts the behavior of others in that space for reasons unrelated to the content of the speech.[169]

Distinguishing regulations that address speech-related harms from those that address non-speech-related harms is a central free speech question of the algorithmic age that our thought experiment helps illuminate. Extending speech protection to strong AI does not make all information that AI produces immune from regulation. It instead highlights the line-drawing work courts already must do to distinguish appropriate government regulation from censorious attempts.

Contemporary free speech law offers courts opportunities for limiting the negative impact of AI expression even if they wrap AI speakers into the constitutional fold. Yet the first

166. *See* City of Erie v. Pap's A.M., 529 U.S. 277, 289–96 (2000) (applying the test from *O'Brien* to a city ordinance prohibiting nude dancing because nude dancing can be a form of expression).

167. Reed v. Town of Gilbert, 135 S. Ct. 2218 (2015) (invalidating an Arizona law that set restrictions on sign size based on whether the signs were "Temporary Directional Signs," "Ideological Signs," or "Political Signs" as content-based); Sorrell v. IMS Health, 564 U.S. 522 (2011) (subjecting a Vermont statute, which prevented pharmacies from selling client data to markets, to heightened scrutiny due to its burden on speech).

168. *IMS Health*, 564 U.S. at 563–71 (finding that the particularity of the regulation contributed to the finding that it was content-based).

169. Kaminski, *supra* note 146, at 53–56.

move of granting AI speakers First Amendment coverage would be a powerful one. It would place AI speech into the constitutional rights box, where government regulations typically face elevated judicial scrutiny. Full protection of such speech from the government's regulation may not inexorably follow, but the government would bear the weighty burden of justifying its regulation.

Our point is this: insofar as a primary basis for protecting AI speech rests on the value of expression to human listeners, free speech protection for strong AIs need not rob the First Amendment of a human focus. Absent a human speaker, government still may attend to the dangers to listeners and other humans. First Amendment coverage need not mean protection in all instances. The prospect of free speech rights for strong AI speakers might encourage useful clarification of the roles of human listeners and of speech harms in U.S. free speech theory and doctrine today.

CONCLUSION

We have explained how foundational free speech theory and doctrine present surprisingly few barriers to First Amendment coverage of strong AI speech. In so doing, we also have considered the powerful counterarguments to such coverage.

The assertion that strong AI speakers might be covered by the First Amendment makes many uncomfortable, perhaps justifiably so. Some worry that in a system of rights, humanness does—or at least should—matter. Others fear that a failure to insist on humanness means that free speech rights must also extend to our cats and dogs. More practically minded critics point out the difficulties in determining how an AI could sue (to vindicate its free speech rights) or be sued (for the harm inflicted by its regulated speech). Still others observe the challenges that would arise in areas of First Amendment law that focus on a speaker's mental state in determining coverage or liability.

We take these concerns seriously. Although we conclude that none of them settles the matter, we also agree that uncertainties will and should remain.

Failing to extend the First Amendment to cover AI speech invites the risk that government will suppress such speech in ways that deprive human listeners of valuable expression or otherwise compromise important free speech interests. At the same time, AI speech can pose substantial dangers to those

same human listeners through its potential power, speed, and scope. Free speech theory and doctrine can thus provide support for the regulation, as well as the protection, of strong AI speakers. Courts may craft adjustments to doctrine to address such expression's potential harms, as well as its value, to human listeners. Such challenges are neither altogether new nor necessarily insuperable. Legal regimes have long adapted to the pressures of emerging technologies.

We close by turning the AI thought experiment on its head: What does the prospect of First Amendment coverage of strong AIs' speech teach us about the limits and possibilities of current free speech law?

Among other things, it forces us to revisit questions about whether and when courts should permit government to address important interests in compensating and preventing the regulable harms of expression without unduly treading on free speech toes. The steady extension of free speech coverage to areas once thought to be immune from First Amendment scrutiny has exacerbated this longstanding challenge, and our AI thought experiment shows some of the challenges yet to come. The hardest judicial work likely will lie here.

The challenges posed by coverage of AI speech may encourage the Court to build on listener-centered rationales to derive limiting principles within zones where free speech coverage applies. If it cares about listeners—and this Court plainly does—it may consider how information-regulation policies can both enhance and compromise listeners' autonomy, enlightenment, and self-governance interests long thought to lie at the core of the First Amendment.

Finally, we have described the logical implications of the Court's existing First Amendment framework, but have not insisted on its inevitability or even its wisdom. The normative and practical difficulties in extending free speech rights to strong AI speakers may force reexamination of the contemporary turn in free speech law and theory that makes future coverage of AI speech plausible. As AI evolves, human judges and policymakers must muddle through the unfolding challenges and find means to address them. Future constitutional law scholars then will debate the merits of their handiwork (perhaps aided by AI enhancements in conducting their analyses).[170] However this work develops, it must be mindful of the

170. Stacy Liberatore, *Your AI Lawyer Will See You Now: IBM's ROSS Be-*

escalating power of AI over information that humans need to survive. The information stakes already are extremely high, and quickly mounting. We should start thinking—Siri-ously—about the implications of AI advances for a variety of legal domains, including freedom of expression.

comes World's First Artificially Intelligent Attorney, DAILY MAIL (May 13, 2016), http://www.dailymail.co.uk/sciencetech/article-3589795/Your-AI-lawyer-IBM-s-ROSS-world-s-artificially-intelligent-attorney.html.