March 2020

Saving the Marketplace from Market Failure: Reorienting Marketplace Theory in the Era of AI Communicators

Jared Schroeder

Follow this and additional works at: https://scholarship.law.wm.edu/wmborj

Part of the Comparative and Foreign Law Commons, European Law Commons, and the First Amendment Commons

Repository Citation

Copyright c 2020 by the authors. This article is brought to you by the William & Mary Law School Scholarship Repository.
https://scholarship.law.wm.edu/wmborj
SAVING THE MARKETPLACE FROM MARKET FAILURE:
REORIENTING MARKETPLACE THEORY IN THE
ERA OF AI COMMUNICATORS

Jared Schroeder*

Artificially Intelligent (AI) communicators represent a new type of actor within public discourse. These entities have played influential roles in recent elections in the U.S. and Europe.\(^1\) This Article examines expression rights for AI actors through the lenses provided by the foundational assumptions of the marketplace of ideas theory and existing free-expression-related rationales regarding non-human actors in the U.S. and European legal systems. The Article contends that the fundamental assumptions of the marketplace model must be revised to focus on the flow of information, the \textit{development} of truth, rather than the more Enlightenment-oriented competition of ideas that leads to the \textit{discovery} of truth. Such a shift would allow limitations on AI that harm the flow of ideas, but otherwise protect AI expression that contributes to democratic discourse.

\textbf{INTRODUCTION}

Artificially intelligent communicators represent a fundamentally new type of actor within public discourse. Such actors—algorithms, AI agents, chatbots—are conveying significant amounts of information in virtual spaces, as well as making decisions about the information citizens do and do not encounter.\(^2\) Importantly, when it comes to how we rationalize freedom of expression, these actors simply were not on the minds of those who provided the foundational assumptions upon which we have come to build our understandings of democratic discourse.\(^3\) Just the same, AI communicators

\* Jared Schroeder is an Assistant Professor of Journalism at Southern Methodist University. He is the author of \textit{The Press Clause and Digital Technology’s Fourth Wave}.  
\(^1\) See infra note 9 and accompanying text.  
\(^3\) For examples of the types of concerns that were on the minds of the thinkers who significantly influenced U.S. and European thought during formative periods, see John Locke, \textit{The Second Treatise of Civil Government and a Letter Concerning Toleration}
have arrived as part of the Fourth Wave of networked communication, a wave that is defined by increasingly meaningful interactions between humans and computer programs. Our existing theoretical and legal models are relatively unprepared for how these AI actors, which are characterized by fundamentally different natures and characteristics than human communicators, influence societal discourse. This concern has been particularly important since the 2016 United States presidential election, as computer programs of varying complexity have been used to leverage the non-physical, largely anonymous nature of online communication to increase the frequency and salience of certain ideas. This concern, along with others, strikes at the heart of marketplace theory, a tool that the U.S. Supreme Court has used to rationalize expansive protections for free expression for a century and European courts have, at times, turned to as well. AI actors, which also capitalize on the significant fragmentation and polarization that comes with the choice-rich nature of information and weaker ties that individuals share online, can flood virtual spaces with certain information and ideas, communicate threatening and hateful speech, and exploit and mislead those who receive their messages. While human actors engage in these same behaviors in virtual spaces, legal systems throughout the world have yet to resolve whether AI entities, which vary significantly in complexity and purpose, should receive the same or similar freedom of expression safeguards as humans.


4 Web 1.0 was characterized by massive information searchability and availability. Web 2.0 included social media and widespread creation of user-created content. Web 3.0 involved growing intrinsic relationships between information and knowledge. See Sareh Aghaei et al., Evolution of the World Wide Web: From Web 1.0 to Web 4.0, 3 INT’L J. WEB & SEMANTIC TECH. 1, 1 (2012); Nupur Choudhury, World Wide Web and Its Journey from Web 1.0 to Web 4.0, 5 INT’L J. COMPUTER SCI. & INFO. TECHS. 8096, 8099 (2014).


From involving themselves in political conversations during recent elections in France, the United Kingdom, and the United States, to discussions regarding school shootings in the United States, debates about the release of political documents, and terror attacks in London, more and more complex computer programs are contributing important facts, advocating on behalf of the interests of those who programmed them, and interjecting false and misleading information into human interactions in virtual spaces. In the final days of the French presidential election in May 2017, for example, bot-based Twitter accounts blasted out tens of thousands of negative messages about centrist candidate Emmanuel Macron. Just one of these accounts produced nearly 1,700 tweets about the French election in twenty-four hours, highlighting crucial concerns when it comes to how we understand AI communicators. First, they are a fundamentally global concern, with out-of-state actors often

---


12 See Chhor, supra note 9.

deploying AI communicators in efforts to influence democratic processes in certain nations. Second, these entities often manipulate the marketplace of ideas by artificially boosting certain ideas that might otherwise fail to gain traction or acceptance and, at the same time, pushing other, human-generated ideas, from the marketplace. In this sense, the AI communicators have come to exploit and exacerbate fundamental problems that scholars have long identified regarding the information marketplace as Enlightenment thinkers such as John Milton and John Locke conceived it, and how it has been interpreted as a theory of the First Amendment in the United States and, to a lesser extent, employed by European courts. Finally, the courts in both the European and United States legal systems appear ill-prepared to consistently apply existing, human-communicator-focused precedent and legal understandings to AI actors that lack the characteristics that define the “human condition,” as thinker Hannah Arendt conceptualized it.

These communicators do not labor or make decisions based on the inescapable knowledge of their own mortality. As Arendt explained, “Trust in the reality of life, on the contrary, depends almost exclusively on the intensity with which life is felt, on the impact with which it makes itself felt.” This emerging class of communicators does not function based on such a reality. They also do not sleep, vote, have families, grow tired of repeating messages, or become emotional. Of course, to simply state that they are not human and therefore ineligible for any kind of free expression safeguards oversimplifies the problem. Scholars have contended that AI actors, which in this Article is used as a blanket term that includes algorithms, AI agents, and chatbots, can contribute valuable information to democratic discourse. It would also be a mistake to conclude that AI communicators are simply another tool that humans use to communicate, such as telephones or email accounts. While some AI communicators do little more than act as vehicles for human-created information, increasingly complex computer programs are taking part in human discourse, and prominent scholars within the robotics law field have concluded it would be a mistake to attribute all AI actors’ actions as extensions of their creators. As legal scholar Jack Balkin explained, “AI agents create problems for law because one cannot always predict what

---

14 See, e.g., O’Connor & Schneider, supra note 11.
15 See id.; Woolley & Howard, supra note 11.
16 LOCKE, supra note 3, at 5–10; MILTON, supra note 3, at 45.
18 Id. at 120.
19 Jack M. Balkin, Lecture, The Three Laws of Robotics in the Age of Big Data, 78 OHIO ST. L. J. 1217, 1219 (2017). Balkin drew these different iterations of computer programs that replicate human behavior together, reasoning that they represent similar concerns when it comes to the foundational laws that should govern them. Id.
20 Norton, supra note 8, at 1145–46.
they will do when they interact with their environment."\textsuperscript{22} Similarly, robotics law scholar Ryan Calo concluded, "We can see already how robots begin to blur the line between people and instrument, and how faulty assumptions about robots lead jurists to questionable or contradictory results."\textsuperscript{23} The European Parliament conveyed similar concerns in a 2017 resolution regarding recommendations from the Commission on Civil Law Rules on Robotics, stating:

\textit{[T]he more autonomous robots are, the less they can be considered to be simple tools in the hands of other actors (such as the manufacturer, the operator, the owner, the user, etc.); whereas this, in turn, questions whether the ordinary rules on liability are sufficient or whether it calls for new principles and rules to provide clarity on the legal liability of various actors concerning responsibility for the acts and omissions of robots where the cause cannot be traced back to a specific human actor . . . .} \textsuperscript{24}

Thus, the growing complexity of AI communicators within democratic discourse raises questions about the extent to which the U.S. and European legal systems are prepared to handle the opaque realm that exists between computer programs that act as human tools, much like a megaphone, and AI that include some level of unpredictability and agency, such as bots that draw content from continuously changing pools of data or those that emulate human-like interactive experiences with the people who use them, such as the Lucia, a Spanish-speaking Facebook Messenger bot that was created in the fall of 2018 to help those who were in the immigrant caravan that was in Mexico and nearing the U.S. border.\textsuperscript{25} As political-oriented bots generated about sixty percent of the discourse regarding the heavily politicized caravan on Twitter, Lucia engaged those seeking to navigate the U.S. immigration system in a conversation that was tailored to their responses.\textsuperscript{26}

Lucia highlights another challenge that legal systems face regarding safeguarding public discourse while at the same time protecting freedom of expression. AI

\begin{thebibliography}{9}
\bibitem{22} Id. at 51.
\bibitem{26} Id.; Issie Lapowsky, \textit{Here’s How Much Bots Drive Conversation During News Events}, WIRED (Oct. 30, 2018, 2:00 PM), https://www.wired.com/story/new-tool-shows-how-bots-drive-conversation-for-news-events/ [https://perma.cc/KU84-4CZ6].
\end{thebibliography}
communicators can contribute important information, a crucial ingredient to democratic discourse. The FOIA // Feed (@FOIAFeed), for example, is a Twitter bot run by the Freedom of the Press Foundation that tweets about reporting that incorporates data gathered using Freedom of Information Act requests. Both the Athens and Frankfurt airports have Facebook Messenger–based bots that allow travelers to ask them questions and receive information about tickets, shops, and airport security. Whatever their contributions, the presence of such non-human messengers within virtual spaces represents a substantial challenge to traditional legal systems regarding freedom of expression, particularly in Europe and the United States, where long-held constructs regarding the free flow of information in democratic society struggle to account for an emerging class of communicators that certainly convey information but, quite fundamentally, lack personhood. For these and other reasons, the European and American legal systems provide foundational building blocks regarding how such communicators can be understood from a law and policy standpoint. Their similarities and differences can be instructive as we consider an emerging concern that is fundamentally international in nature.

This Article examines the question of freedom of expression rights for AI actors through the lenses provided by existing U.S. and European free-expression precedents regarding non-human actors, as well as the marketplace of ideas theory. While several scholars are addressing this emerging concern—and coming to a variety of conclusions—the Article draws together international and theoretical building blocks to contribute another potential approach to how such actors’ rights can be understood. To be clear, the Article’s foundational concerns relate with the question of how these legal systems, and others around the world, can safeguard the type of human interaction—the exchange of ideas and flow of information—that is fundamental to societal discourse from being overwhelmed and dominated by AI communicators, while at the same time protecting freedom of expression, which generally requires the government to remain clear of deciding who can and cannot communicate in most instances. In doing so, the Article explores legal rationales for freedom of expression rights for AI communicators, ultimately outlining a revised, artificial marketplace approach to conceptualizing the flow of information.

29 See, e.g., Norton, supra note 8, at 1150–51.
30 See id. at 1149; see also European Parliament Resolution, supra note 24.
The Article consists of four parts: Part I defines artificial intelligence and considers the types of AI communicators that are involved in human discourse. It also explores the connective nature of virtual spaces. Part II examines the foundational assumptions of the marketplace of ideas theory, as well as discursive thought, a philosophical approach to freedom of expression in democratic society that is uniquely positioned to account for the collective and community-based properties of networked technologies and the AI communicators that have found homes in online spaces and are changing the way the marketplace of ideas functions. Part III considers how the European and American legal systems have handled questions regarding the rights of non-human communicators in the past and how international statements about freedom of expression might influence future decisions. Part IV concludes by drawing from these and other considerations to outline the artificial marketplace approach to understanding the rights of AI communicators.

I. AI COMMUNICATORS & VIRTUAL SPACES

AI and other non-human actors created by people, such as Frankenstein’s monster, Maschinemensch in *Metropolis*, Hal from *2001: A Space Odyssey*, and the operating system in *Her*, have long captured human imaginations.32 In fact, Isaac Asimov’s fictional, genre-defining “three fundamental Rules of Robotics” have played an outsized role in legal discussions about how we should understand questions about the rights of robots and computer programs.33 Legal scholars have paid increasingly careful attention to robotics law in recent years.34 They have done so as more human activities are automated, and online networked technologies—and the virtual environments they create—have provided ideal environments for computer programs to communicate with humans and to determine the information humans do and do not encounter.35 Scholars within the robotics law field have considered a variety of concerns, such as civil and criminal liability for the physical or emotional harms such entities might create, or privacy and copyright questions that emerge in an era when computer

---


33 ASIMOV, supra note 32, at 37 (1. “[A] robot may not injure a human being, or, through inaction, allow a human being to come to harm.” 2. “[A] robot must obey the orders given it by human beings except where such orders would conflict with the First Law.” 3. “[A] robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.”).

34 See generally Calo, supra note 31; Cofone, supra note 31; Norton, supra note 8.

35 See, e.g., Lapowsky, supra note 26; Sillers, supra note 28.
programs are scraping information and creating new works. This Article focuses on rights and freedoms for AI actors, rather than the types of question that arise when these entities break existing laws. Addressing the question of free expression for AI requires that we establish the nature of such communicators and the networked environments in which they engage with people.

A. Definitions

The term “artificial intelligence” was coined during a conference that was organized by Defense Advanced Research Projects Agency (DARPA) and conducted at Dartmouth College in the summer of 1956. Those who proposed the conference, including Claude Shannon, contended “that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.” In the ensuing decades, scholars have defined AI as systems that can think and act rationally in ways that mirror human behavior and intelligence. This Article uses AI as a term that engulfs computer programs that have the capacity to take action in their environment, algorithms, and machine learning tools. While these terms certainly can be distinguished from one another, the central concern of this Article is how computer-based, non-human actors are influencing human relationships, community, and discourse, as well as how we should conceptualize their nature and the extent to which they can communicate freely in democratic society. Thus, AI, for present purposes, refers to computer-based communicators, including algorithms, AI agents, and chatbots.

U.S. and European courts have done little to guide efforts to agree upon distinct legal definitions for AI and related technologies. When the courts have referred to AI, they have primarily done so in reference to algorithms, computer programs that filter online content or sift through data to return search results. Computer scientists Stuart Russell and Peter Norvig characterized AI systems as being on a spectrum from weak to strong, concluding that strong systems can replicate human-like consciousness, while weaker systems are limited to merely executing programmed
Weak AI include Apple’s Siri and Amazon’s Alexa, which simply transcribe verbal requests for information by humans and then produce auditory versions of the search results. They represent a technological shift in the availability of information, but do not represent a dynamic AI entity that can replicate human actions or consciousness. Strong, or full, AI is more theoretical at this point, as no computer program has been able to fully implement the human brain or human actions. Google’s DeepMind, however, is working on numerous projects in which advanced algorithms learn to replicate limited aspects of human behavior, such as teaching itself how to excel at chess or predicting eye diseases and providing recommendations for treatment. As far as communicators go, advanced chatbots, such as Mitsuku and Rose, have won the Loebner Prize for their ability to replicate human interaction to the extent that other humans have no idea whether they are communicating to with a person or a bot. The potential applications for such communicators, from automating most commercial transactions to making political arguments that are triggered by the use of certain hashtags or keywords, have significant implications for discourse in society, even if they are limited in their ability to interact organically with others. The implications are particularly important as AI actors become more nuanced and complex, and the nature of the virtual environments in which they encounter human communicators continue to change the way individuals communicate and understand both themselves and others.

B. The Nature of Virtual Spaces

The emergence of networked communication tools and the virtual realms for discourse that they create have led to substantial changes for human and AI entities. Most crucial among these changes, however, has been the emergence of a space in which the two can coexist within the same, virtual environment. While human-like walking, talking robots—as seen in the movies—remain more imagination than reality,

42 See id.; Massaro & Norton, supra note 5, at 1192–93.
43 Balkin, supra note 19, at 1220.
47 See Massaro & Norton, supra note 5, at 1170–71.
the relatively minimal entry requirements in virtual spaces allow computer programs, absent physical forms, to capitalize upon anonymity and lack of presence to engage with humans on a somewhat level playing field.\textsuperscript{48} When combined, the introduction of human and non-human communicators into common, interconnected, and generally anonymous spaces is fundamentally changing the way democratic discourse occurs in the twenty-first century.\textsuperscript{49} Crucially, the growing presence of AI entities in such spaces means that the influence of networked tools on discourse can no longer be considered without accounting for AI communicators. Networked technologies have substantially lowered the hurdles that once limited individuals—and AI entities—from communicating messages to audiences.\textsuperscript{50}

Discourse now takes place within a choice-rich, deeply fragmented environment.\textsuperscript{51} In such an environment, individuals can select the people, organizations, and sources of information with which they would like to interact, ultimately limiting the scope of potential ideas they encounter.\textsuperscript{52} Such a development has led individuals to construct intentional communities in which they surround themselves with generally like-minded individuals and sources of information that reinforce their pre-existing ideas.\textsuperscript{53} Legal scholar Cass Sunstein explained that “[m]embers of a democratic public will not do well if they are unable to appreciate the views of their fellow citizens, if they believe ‘fake news,’ or if they see one another as enemies or adversaries in some kind of war.”\textsuperscript{54} Individuals in such environments generally can only become more extreme in their views, since they primarily encounter ideas that reinforce their existing beliefs and, conversely, few messages or individuals who present opposing perspectives.\textsuperscript{55} Such a process of community formation and construction of realities is fundamentally different than how such information flowed and interactions occurred in preceding eras. Generally, in the twentieth century, traditional news organizations provided relatively similar sets of information to mass audiences.\textsuperscript{56} Since

\textsuperscript{48} See Persily, supra note 6, at 70.

\textsuperscript{49} See CLAY SHIRKY, HERE COMES EVERYBODY: THE POWER OF ORGANIZING WITHOUT ORGANIZATIONS 17 (2008); SHERRY TURKLE, ALONE TOGETHER: WHY WE EXPECT MORE FROM TECHNOLOGY AND LESS FROM EACH OTHER 11 (2011); Massaro & Norton, supra note 5, at 1172.

\textsuperscript{50} See supra notes 27–28 and accompanying text.


\textsuperscript{53} For network analysis of this phenomenon, see Itai Himelboim et al., Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter, 18 J. COMPUTER-MEDIATED COMM. 154, 166–71 (2013).

\textsuperscript{54} SUNSTEIN, supra note 52, at IX.

\textsuperscript{55} See SUNSTEIN, supra note 51, at 76–77.

audiences had a limited range of sources from which to choose, citizens generally received relatively similar baselines of ideas from which to enter into interactions with others.\textsuperscript{57} Furthermore, when individuals entered into discourse with others, they almost always had to be physically present, which brought with it certain dynamics that created social capital and strengthened the ties that bound individuals together in a community.\textsuperscript{58}

In this regard, networked communication scholars define a “tie” as a relationship “between communicators wherever they exchange or share resources such as goods, services, social support or information.”\textsuperscript{59} People who have long-standing relationships in which they share ideas and communicate often maintain deeper bonds than those who do not.\textsuperscript{60} Scholars have concluded that online interactions are inherently less likely to involve strong ties or the exchange of meaningful amounts of social capital between participants.\textsuperscript{61} This conclusion aligns with sociologist and psychologist Sherry Turkle’s conclusion that, “in person, we have access to the messages carried in the face, the voice, and the body. Online, we settle for simpler fare: We get our efficiency and our chance to edit, but we learn to ask questions that a return email can answer.”\textsuperscript{62} Thus, individuals not only have significant choice regarding the facts and \textit{truths} they encounter in their intentionally formed communities, they also fundamentally do not form close relationships with others in virtual forums. The result of such an environment, one in which individuals within a society do not share common factual basis for interactions with others, is a sort of fragmented multiverse of public spheres. These shifts in human communication—how individuals communicate and form relationships—have opened the door for AI agents to succeed in having their messages accepted and shared within virtual communities.\textsuperscript{63} In both the Brexit vote and the 2016 U.S. presidential election, for example, bots were found to be most influential when they reinforced existing opinions.\textsuperscript{64}

Within each fragmented, ideologically formed public sphere, the truths that emerge from interactions among individuals in that community might be substantially different than those in other online spaces. Therefore, in some communities, the \textit{truth} is that former United States President Barack Obama was born in Kenya, while in others

\textsuperscript{57} See id. at 70.
\textsuperscript{59} Caroline Haythornthwaite, \textit{Strong, Weak, and Latent Ties and the Impact of New Media}, 18 INFO. SOC’Y 385, 386 (2002); see also Putnam, \textit{supra} note 58, at 18–19.
\textsuperscript{60} Haythornthwaite, \textit{supra} note 59, at 386.
\textsuperscript{61} \textit{Sherry Turkle, Reclaiming Conversation: The Power of Talk in a Digital Age} 23 (2015).
\textsuperscript{62} \textit{Id.}
\textsuperscript{63} See \textit{id.}; Bessi & Ferrara, \textit{supra} note 6.
such a conclusion is preposterous. The emergence and massive spread of persistently believed, yet directly conflicting realities such as these requires a different conceptualization of the flow of information and the nature of understanding in democratic society. In particular, the marketplace metaphor, which has dominated American understandings of freedom of expression for nearly a century and has historically been associated with European-founded Enlightenment understandings, must be reconsidered in light of the loss of shared realities and the fragmentation into communities of generally like-minded individuals. While the marketplace theory is discussed in more detail later in the Article, it suffices at this point to conclude that we must consider a fundamentally different type of information marketplace in the twenty-first century.

This change in the flow of information has altered how people interact with others and AI communicators. Individuals are forming intentionally crafted personalities that are substantially influenced by the nature of the tools they use to communicate. The relative anonymity and global nature of discourse in virtual spaces, along with the lack of strong bonds between individuals, is influencing how individuals choose to represent themselves. While people are limited regarding self-representation in physical environments, virtual spaces present opportunities and challenges for identity formation and enactment. Communication scholar Zizi Papacharissi explained that “the individual must then engage in multiple mini performances that combine a variety of semiological references so as to produce a presentation of the self that makes sense to multiple audiences.” Thus, while the lack of physical presence and the global nature of virtual spaces allow individuals to largely construct their identities online, the form of the tool also plays a role in stripping away nuance. Therefore, individuals limit the amount of subtlety in their messages as they seek to mitigate the potential for misunderstanding and to reach the different constituencies within their intentionally formed communities. The result is generally more strident, emboldened self-representations. Individuals trade the types of surgical, strategic efforts they might have made in a physical environment for more blunt projections of meaning to difficult-to-identify online audiences.

---


66 See Jared Schroeder, Shifting the Metaphor: Examining Discursive Influences on the Supreme Court’s Use of the Marketplace Metaphor in Twenty-First-Century Free Expression Cases, 21 COMM. L. & POL’Y 383, 426–30 (2016); see also infra Section II.B.

67 See Schroeder, supra note 66, at 430.

68 See, e.g., TURKLE, supra note 49, at 11.

69 Zizi Papacharissi, A Networked Self: Identity Performance and Sociability on Social Network Sites, in FRONTIERS IN NEW MEDIA RESEARCH 207, 209 (Francis L.F. Lee et al. eds., 2013).

70 See id. at 209.

71 See id.

72 See SCHROEDER, supra note 56, at 69–74.

73 Id. at 69–70.
Each of these concerns—the choice-rich environment, human-relationship-limiting tools, fragmented virtual communities, limited bonds and social capital, and the intentional and limited nature of self-representation—all not only substantially influence human discourse, but enable AI communicators to thrive. While AI entities that truly learn from experience and can interact with their environment, such as Google’s DeepMind project, are still being developed, computer-based communicators are becoming more and more adept at mimicking human communication. The Cleverbot, a program that has been learning from its exchanges with humans for more than a decade, references more than 180 million lines of previous interactions each time it engages in a conversation. The Picture Description Bot selects images from Wikimedia Commons and considers them using Microsoft’s Cognitive Services. The service attempts to describe what is in the picture and then posts the image and description on Twitter (@picdescbot). While the descriptions are correct as often as they are hilariously wrong, these types of communicators continue to develop human-like capabilities. The lack of physical form, and the ways that individuals interact and create communities online, has opened the doors for AI communicators to easily integrate themselves into human discourse. While they might struggle to mimic the more intricate aspects of human behavior, the often blunt, ideologically focused messages humans share in their intentional communities online and the general lack of meaningful relationships allow such non-human communicators to fit into existing narratives and integrate themselves into human discourse.

C. AI and Political Discourse

Many AI communicators automatically post at certain times or are set up to respond to certain hashtags or key words. Such actors operate about 15% of Twitter’s

---

74 See id. at 71.
78 Id.
79 See id.
80 See, e.g., Hill et al., supra note 76, at 249.
320 million active monthly users. Similarly, as many as 24 million of Instagram’s 700 million accounts are run by non-human users. During the most recent election in the United Kingdom, partisans created bots on Tinder, the dating and hook-up app, to strike up conversations with younger voters. The bots communicated more than 30,000 messages, swiping “yes” on users and, if the other person swiped “yes” back, they engaged them in a discussion about the upcoming election rather than a conversation about getting together. Several months beforehand, in November 2016, about one-quarter of all of the tweets about the United States presidential election were generated by bots. In the aftermath of the November 2016 election and the Brexit referendum beforehand that June, a group of researchers suggested that bot activity on Twitter led to a 0.6% increase in votes for Trump and a 0.54% swing for “leave” votes in the Brexit referendum. The authors emphasized that the AI communicators’ impact was strengthened by the speed at which information travels online and the fragmented nature of online communities. Many of the bots, in these instances and in matters that have since arisen, communicate false or misleading information. The AI agents are especially impactful during politically charged events. AI communicators dominated online discussion of the immigrant caravan, which was central to the discourse during the run-up to the U.S. midterm elections in 2018. In the week before the election, they accounted for 40 to 60% of the online discussion about the caravan. Of course, many of the human account holders who encountered the AI-authored messages had no idea that the ideas they were repeatedly encountering were not an example of widespread public support. They were, instead, artificially boosted ideas.

83 Alexandra Ma, Millions of Instagram Users Are Just Spambots, HUFFPOST (July 2, 2015), http://www.huffpost.com/entry/instagram-spambot_n_7708550.html [https://perma.cc/5BG8-E8H9].
86 Kollanyi et al., supra note 9, at 4.
87 Gorodnichenko et al., supra note 64, at 27.
88 Id. at 1.
90 See Lapowsky, supra note 26.
91 Id.
Choice-rich virtual environments have created ideal spaces for AI communicators to reinforce the strongly held, and seldom-challenged, narratives that are found within fragmented, intentional communities and to mimic the self-representation approaches that are used by many human account holders. The bot-based activities in recent elections in France, the United Kingdom, and the United States, as well as evidence that foreign bots have engaged in issue-based campaigns, clearly illustrate that the combination of emerging AI communicators and the nature of virtual environments have raised substantial questions about the place of such entities within democratic discourse throughout the world.92

II. THE MARKETPLACE AND DISCURSIVE THOUGHT

Questions regarding the rights of AI communicators almost inherently find themselves caught up in concerns about the flow of information and their potential influence on meaning making among citizens in democratic society. These entities lack the characteristics that have long been associated with being human, particularly in regard to conscience and empathy, but are endowed with the ability to share discourse-enabling information and to interact with humans. Thus, these entities bring the potential for significant interruption, and even market failure, into the flow of ideas and interactions among citizens in democratic society. Such a concern inherently relates to the marketplace of ideas theory of the First Amendment, the dominant justification for freedom of expression in the United States.93 The theory assumes that in a relatively free communication environment, rational individuals are capable of discerning truth from falsity.94 Thus, truthful ideas will succeed and false ideas will fail in the open market. The theory does not purely pertain to the United States legal system.95 The European Court of Human Rights (ECtHR) has cited the theory in several freedom-of-expression-related decisions.96 This Section examines the basic assumptions of marketplace theory and its shortcomings, which are particularly apparent when it comes to AI communicators. This Section furthermore considers how the theory has been used in ECtHR decisions before outlining the foundational assumptions of discourse theory and the approach’s potential for reinvigorating marketplace theory in the twenty-first century.

92 See Abeshouse, supra note 89.
A. The Holmsian Marketplace

Justice Oliver Wendell Holmes introduced the marketplace of ideas theory of the First Amendment in his dissent in Abrams v. United States in 1919.97 Jacob Abrams and four other Russian immigrants who distributed leaflets in New York City to protest the U.S. military’s actions on Russian soil toward the end of World War I were convicted of violating the Espionage Act of 1917.98 The law, which was signed soon after the United States entered the war, criminalized any communication that could be understood as interfering with military operations, supported enemy activities, encouraged members of the military to be disloyal, or obstructed the draft.99 Justice Holmes had written for a unanimous Court in three cases that were substantially similar to Abrams in the previous term, only eight months earlier, concluding in each case that the Espionage Act did not violate freedom of expression.100 The Court came to the same conclusion in Abrams, but Justice Holmes, joined by Justice Louis Brandeis, dissented.101 In doing so, he constructed a much different argument.102 He explained,

[W]hen men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market . . . .103

He continued by concluding that his finding was “the theory of our Constitution.”104 While it was a dozen more years before the Supreme Court struck down a law based on its interpretation of the First Amendment,105 the dissent created a lasting rationale for freedom of expression.106

Justice Holmes’s dissent in Abrams was substantially based on his understanding of the nature of truth.107 Importantly, however, his conclusions in this regard

97 250 U.S. 616, 630 (1919).
98 Id. at 616–19.
101 Abrams, 250 U.S. at 624–31 (Holmes, J., dissenting).
102 Id.
103 Id. at 630.
104 Id.
106 See Howard C. Anawalt, The Right to Communicate, 13 DENV. J. INT’L L. & POL’Y 219, 227 (1984) (“Holmes states that there is a market place of ideas where all communications are free to circulate.”).
were substantially different than the assumptions that have generally come to be associated with the theory. He did not believe in objective truth, a conclusion he communicated in many of his legal and personal writings. In a letter ten years after he wrote the Abrams dissent, he declared “[a]bsolute truth is a mirage.”\footnote{108} In another letter, this time in 1912, he explained to a friend that facts change and can only be taken for what they are at the moment they are needed.\footnote{109} Similarly, in his book, The Common Law, which was published early in his career in 1881, he explained that “[t]he life of the law has not been logic: it has been experience.”\footnote{110} Ultimately, rather than an objective truth, Justice Holmes contended he was a “bettabilitarian,” a term he created to describe his belief that the best anyone can do when making a decision is to \textit{bet} using the facts as they know that at the time and their personal experiences.\footnote{111} He also rejected those who claimed to have ownership of absolute truth, contending that such unmoving positions were what led the U.S. Civil War, in which he was shot on three different occasions.\footnote{112} In “Natural Law,” which was published a year before the Abrams dissent, he explained, “[W]hen differences are sufficiently far reaching, we try to kill the other man rather than let him have his way. But that is perfectly consistent with admitting that, so far as appears, his grounds are just as good as ours.”\footnote{113} Thus, the Holmsian marketplace, as introduced in Abrams and contextualized in his other legal and personal writings, contended that, since no one could know that absolute truth, the government should generally avoid limiting expression.\footnote{114}

B. The U.S. Supreme Court’s Marketplace

Since Justice Holmes’ use of the marketplace metaphor in 1919, the Supreme Court has consistently employed it as a tool to rationalize decisions to protect freedom of expression.\footnote{115} Justices from a variety of ideological perspectives have used the metaphor and it has remained popular, particularly since the 1970s.\footnote{116} Furthermore, justices have used the metaphor in nearly every area of First Amendment law, including prior

\textsuperscript{109} Id. at 7.
\textsuperscript{113} Oliver Wendell Holmes, Natural Law, 32 HARV. L. REV. 40, 41 (1919).
\textsuperscript{114} For a more detailed look at Justice Holmes’ understandings of truth and his personal, scholarly, and legal writings in that regard, see Schroeder, supra note 111, at 169–203.
\textsuperscript{115} W. Wat Hopkins, The Supreme Court Defines the Marketplace of Ideas, 73 JOURNALISM & MASS COMM. Q. 40, 42 (1996).
\textsuperscript{116} See id.
restraint, defamation, commercial speech, and obscenity. Though Justices have never explicitly defined how they understand the marketplace of ideas and Justice Holmes’s use of the marketplace in *Abrams* did not include any footnotes or in-text references to sources that might have influenced his decision and his intended meaning, the metaphor has come to be associated, by Justices and legal scholars, with Enlightenment-related assumptions about truth, the rationality of citizens, and the flow of information in society. Such a connection between Enlightenment thought and the marketplace of ideas can be seen in the Court’s decision in *Red Lion Broadcasting v. FCC* in 1969. Justice Byron White, in writing for the Court in its decision to uphold the Fairness Doctrine, explained, “It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail . . . .” Importantly, Justice White’s reasoning was constructed to rationalize upholding requirements that essentially forced broadcasters to allow a diversity of perspectives. In other words, the Court justified upholding the doctrine because Justices understood it as preserving the marketplace. Similarly, in 1953 in *United States v. Rumely*, which dealt with political activist Edward Rumely’s fight with a House committee’s demand that he provide the names of those who purchased his books in bulk, Justice William Douglas wrote, the “publisher bids for the minds of men in the market place of ideas. The aim of the historic struggle for a free press was ‘to establish and preserve the right of the English people to full information . . . .’” Justice Douglas characterized these conclusions as “the tradition behind the First Amendment.” In this sense, he associated the marketplace with the right to communicate and receive information.

The landmark 1974 defamation decision *Gertz v. Robert Welch, Inc.* indicates similar connections between truth, the First Amendment, and the marketplace rationale. Justice Lewis Powell, in writing for the Court, explained, “However pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries but on the competition of other ideas.” Finally, in relation to the Internet, Justice John Paul Stevens concluded the Court’s 1997 decision in *Reno v. ACLU*, which established that online communication would receive significant First Amendment

---


120 Id.

121 Id. at 389–90.

122 Id.


124 Id.


126 Id. at 339–40.
protection, with the finding that government regulation of online content “is more likely to interfere with the free exchange of ideas than to encourage it.”127 He continued, “The interest in encouraging freedom of expression in a democratic society outweighs any theoretical but unproven benefit of censorship.”128

In each of these instances, the Justices communicated a concern for the flow of information and the danger of the government influencing the ideas that can and cannot be communicated in the marketplace of ideas.129 In doing so, the Justices rationalized their uses of the marketplace metaphor with fundamental assumptions regarding Enlightenment thought. In particular, Enlightenment thought generally conceptualizes truth as being objective and universal, something that must be found rather than created.130 It also assumes that individuals are generally rational and, given access to unfettered ideas, are capable of discerning truth from falsity.131 Such thinking can be seen in British philosopher and central Enlightenment figure John Milton’s conclusions in Areopagitica in 1644.132 Milton, in arguing against licensing requirements and censorship, concluded, “Where there is much desire to learn, there of necessity will be much arguing, much writing, many opinions; for opinion in good men is but knowledge in the making.”133 He continued that, in an open contest, truth would vanquish falsity and limits on publication halt the pursuit of truth.134 Nearly fifty years later, British thinker John Locke communicated similar ideas, explaining that truth and reason were gifts from God.135 He found that human nature was governed by reason, which “teaches all mankind[ ] who will but consult it.”136 Finally, Scottish thinker Adam Smith, whose ideas about the free market were published in The Wealth of Nations in 1776, connected Enlightenment assumptions about human rationality with economic decisions.137 All of these influences can be seen in the ways that the Court has come to understand the marketplace approach. This is particularly

128 Id.
131 Peter J. Gade, Postmodernism, Uncertainty, and Journalism, in Changing the News: The Forces Shaping Journalism in Uncertain Times 64 (Wilson Lowrey & Peter J. Gade eds., 2011); Siebert, supra note 130, at 40–41.
132 See Milton, supra note 3, at 45.
133 Id.
134 Id. at 50.
135 Locke, supra note 3, at 5.
136 Id. at 4.
apparent in regard to Justices’ assumptions about the nature of truth and the likelihood that truth will succeed and falsity will fail in a free exchange of ideas.\(^{138}\)

**C. The ECtHR and the Marketplace of Ideas**

Though the ECtHR has employed marketplace-related rationales for freedom of expression far less often than the U.S. Supreme Court, judges in several decisions have carefully considered and sought to apply the theory in article 10–related decisions.\(^{139}\) In *Mouvement Raëlien Suisse v. Switzerland*, a 2012 ECtHR Grand Chamber judgment, judges narrowly upheld a Swiss administrator’s right to reject a poster that sought to bring awareness to the Raelien Movement.\(^{140}\) The poster read “The Message from Extraterrestrials” and included a phone number and website address.\(^{141}\) The group, as the Court outlined, believes in creating a superior genetic version of humanity and has been connected with pedophilia.\(^{142}\) The Movement contended that halting it from displaying its poster violated its article 10 rights under the ECHR.\(^{143}\) The Grand Chamber upheld the limitations on the public display of the poster but, in its reasoning, contrasted public parks and billboards with online content.\(^{144}\) The court reasoned, “If streets and parks of a city are the historical quintessential public fora, the Internet is today’s global marketplace of ideas.”\(^{145}\) In doing so, the court indicated that different forums, online versus physical, should receive different levels of protection under article 10.\(^{146}\) In two extensive footnotes that accompanied the reference, the court attributed the marketplace approach to Justice Holmes’s dissent in *Abrams*, but went a step further, associating the theory not only with Enlightenment thought, including Milton’s *Areopagitica*, but also with German thinker Jürgen Habermas, whose work (discussed later in this Article) operates using fundamentally different assumptions about truth and the flow of information.\(^{147}\) Though the U.S. Supreme Court has repeatedly associated the marketplace with Enlightenment thought,\(^{148}\) it has never

---


139 Convention for the Protection of Human Rights and Fundamental Freedoms art. 10, Nov. 4, 1950, 213 U.N.T.S. 2889 [hereinafter ECHR]. This Article promises “[e]veryone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.” Id. The second part of the Article, however, indicates that these safeguards come with “duties and responsibilities.” Id. These “may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society.” Id.


141 Id. at 382.

142 Id. at 382–83.

143 Id. at 390.

144 Id. at 434 (Pinto de Albuquerque, J., dissenting).

145 Id.

146 Id. at 434 & n.2.

147 Id. at 430 n.2.

148 See Dickerson, supra note 118, at 368.
connected it with more discursive ideas, such as those put forth by Habermas. Similarly in Delfi v. Estonia, a 2015 ruling by the Grand Chamber of the same court, justices concluded that despite Delfi’s efforts to take down offensive comments that were made by users, a balance must be found between article 10’s protections for expression and limiting immoral, hateful, and offensive expression. In the final sentences of the dissent, judges drew together Areopagitica and the marketplace to caution lawmakers against using the case to support wider limitations on online expression. The court explained that Areopagitica sought to convince Parliament “that licensing had no place in the free pursuit of truth. It argued that an unlicensed press would lead to a marketplace of ideas in which truth might prevail. It could not undo the bigotry of Parliament. We hope that it will have more success today.”

In four other cases that did not address online messages, jurists considered government efforts to limit potentially dangerous speech. In two cases from Turkey, the court found that the communicator’s free expression rights were violated when officials limited their expression. In dissents in both cases, judges qualified the marketplace of ideas in contending not all speech should be protected. In Karatas v. Turkey, a 1999 judgment that involved a Kurdish author’s poetry, titled “The Song of Rebellion,” dissenting judges explained, “Unlike the advocacy of opinions on the free marketplace of ideas, incitement to violence is the denial of a dialogue, the rejection of the testing of different thoughts . . . .” Similarly, in Sürek and Özdemir v. Turkey five years earlier, two judges dissented when the court found government officials had violated the newspaper publishers’ article 10 rights. The dissenting judges contended the marketplace was reserved for peaceful expression, not threats of violence. More recently, in Taranenko v. Russia in 2014, a protestor who contended her article 10 rights were violated after she was arrested for joining a group that occupied a government building in protest had her conviction upheld.

150 Id. at 413–14 (Sajo & Tsotsoria, J.J., dissenting).
151 Id. at 414.
156 See id. at 43–44.
The court reasoned that “[e]xpression in the marketplace of ideas is only possible where no violence is incited, threatened or exerted. Where there is violence, there is no communication.” 158

In each of these instances, judges used marketplace theory to distinguish communication from violent action, something that is not a concern regarding AI agents in online environments. In the Internet-related cases—Mouvement Raëlien Suisse and Delfi—however, judges both associated the marketplace approach with Enlightenment ideas and used it to rationalize expansive protections for online expression. 159 These Enlightenment-related assumptions about truth and information, however, are central to consistent concerns regarding the marketplace of ideas foundational assumptions. These concerns are particularly relevant in the twenty-first century.

D. The Marketplace’s Shortcomings

Despite the theory’s place as the U.S. Supreme Court’s most commonly used tool for rationalizing safeguards for freedom of expression and its use by the ECtHR in article 10 cases, legal scholars have consistently questioned the fundamental assumptions that have come to be associated with the marketplace approach. First Amendment scholar Jerome Barron, in laying out his access-focused theory, contended “[o]ur constitutional theory is in the grip of a romantic conception of free expression, a belief that the ‘marketplace of ideas’ is freely accessible.” 160 In rejecting the marketplace approach, he contended that, if such a space once existed, it no longer does. 161 Instead, news organizations limit the range and scope of ideas that are accessible to the public. 162 Barron’s concern about access to ideas gets at one of the fundamental assumptions of the marketplace approach—that rational individuals will encounter the same or similar ideas as they seek to come to conclusions about truth and falsity. 163 Legal scholar Stanley Ingber communicated similar concerns. He characterized the marketplace as “more myth than reality.” 164 He explained that political, economic, and social perspectives firmly influence the information and frequency people encounter and their personal conclusions regarding truth. 165 He explained, “[T]he inevitable differences in these perspectives caused by the vastly differing experiences among

---

158 Id. at 28 (Pinto de Albuquerque, J.; Turković, J. & Dedov, J., concurring).
161 See id. at 1641, 1678.
162 See id. at 1641.
163 See id. at 1678.
164 Ingber, supra note 94, at 48.
165 See id.
individuals make resolution of disagreement through simple discussion highly unlikely.” Similarly, Frederick Schauer explained,

[O]ur increasing knowledge about the process of idea transmission, reception, and acceptance makes it more and more difficult to accept the notion that truth has some inherent power to prevail in the marketplace of ideas, or that the distinction between truth and falsity has much explanatory power in telling us which ideas are likely to be accepted by the public and which are likely to be rejected.

Each of these scholars communicated concerns about the fundamental inner workings of the theory. They questioned, in concerns that are particularly related with AI actors, how the marketplace can function if generally rational individuals do not encounter the same ideas and, if they do so, those ideas are presented differently or encountered with greater or lesser frequency. These concerns are of particular importance when considered within the context of the unique nature of AI communicators and their ability to convey false information, to drown out human ideas in the marketplace, and to reach individuals in intentionally formed, ideological communities online.

Of course, at the heart of the theory, and of these concerns, is the discovery of truth. Scholars have reserved their greatest criticisms and concerns regarding the theory for its assumptions about the nature of truth. First Amendment scholar C. Edwin Baker argued, “[T]he assumptions on which the classic marketplace of ideas theory rests are almost universally rejected.” He continued that the “theory is unworkable, dangerous, and inconsistent with a reasonable interpretation . . . of the first amendment.” Baker, rather succinctly, concluded, “[T]ruth is not objective.” He recognized that many factors influence a person’s conclusions regarding what they find to be true, such as “interests, desires, or aesthetics—which guide the development of perceptions.” Similarly, legal scholar Vincent Blasi concluded that when the marketplace concept’s assumptions of truth are questioned, it “is generally found wanting as a persuasive contemporary rationale for free speech.”

---

166 Id. at 15.
167 Frederick Schauer, The Role of the People in First Amendment Theory, 74 CALIF. L. REV. 761, 777 (1986).
170 Id. at 3.
171 Id. at 12.
172 Id. at 13.
highlighted similar concerns. He explained, “In order to be discoverable, however, truth must be an objective rather than subjective, chosen concept.”\textsuperscript{174} If the foundational understandings about truth falter, the rest of the theory struggles. If truth is subjective, then it is no longer something that can be discovered via a competition with falsity in an exchange of ideas.\textsuperscript{175}

When applied to twenty-first century questions regarding AI actors, it becomes clear that the marketplace metaphor struggles as a rationale for freedom of expression. AI actors have the power to overwhelm human expression by posting countless messages. In doing so, whether these messages are truthful or not, they push certain ideas, or products, from the storefronts of the marketplace of ideas. When these concerns are drawn together with the fragmented, polarized, choice-rich nature of online communities, it is entirely possible that the marketplace of ideas will be transformed into a series of storefronts that offer only one or two products—or ideas. In this scenario, no competition between truth and falsity occurs. The decisions that humans make regarding the ideas, people, and organizations they encounter in virtual spaces, along with the power AI actors have to limit ideas in perspectives, can easily lead to competition-free marketplaces of ideas.

\textbf{E. Discursive Revisions}

When the ECtHR placed Habermas alongside Milton’s and John Stewart Mill’s ideas, as well as Justice Holmes’s conclusions from Abrams, in its discussion of the marketplace approach in Mouvement Raëlien Suisse, it drew together entirely different sets of assumptions regarding the flow of information and meaning making.\textsuperscript{176} Habermas, along with other thinkers, came to significantly different conclusions regarding the nature of truth and the flow of information in democratic society.\textsuperscript{177} These conclusions, which are generally in philosophical opposition to Enlightenment thought, are particularly relevant when we consider the nature of virtual spaces and the human and non-human actors that inhabit them. Virtual spaces, along with the interactions that occur among human and non-human communicators, are inherently discursive. The term “discursive,” in this conceptualization, refers to interactions that encourage collective, purposive discussions regarding issues and concerns that face society. Discursive thought is in many ways an antithesis to the more traditional, Enlightenment-founded rationalizations that have dominated much of European and American legal thought regarding how information flows and freedom of expression functions in democratic society.\textsuperscript{178}

\begin{itemize}
  \item \textsuperscript{174} Ingber, \textit{supra} note 94, at 15.
  \item \textsuperscript{175} See Schroeder, \textit{supra} note 66, at 393.
  \item \textsuperscript{177} See 1 JÜRGEN HABERMAS, THE THEORY OF COMMUNICATIVE ACTION: REASON AND THE REALIZATION OF SOCIETY 11–14 (Thomas McCarthy trans., 1981).
  \item \textsuperscript{178} See David A. Hollinger, \textit{The Enlightenment and the Genealogy of Cultural Conflict in...}
In particular, discursive approaches to the flow of information and the construction of meaning within society generally assume that truth is the result of interactions among free individuals. In this sense, truth emerges when individuals use information to interact with one another with the goal of coming to a mutual agreement. Such was the conclusion of American thinker John Dewey, who explained that “knowledge is a function of association and communication; it depends upon tradition, upon tools and methods socially transmitted, developed and sanctioned.” Similarly, Habermas, who read Dewey’s work, made community and collective activity central to his conceptualizations of the public sphere and, later, his theory of communicative action. He contended that “[i]n other contexts, one also speaks of ‘general presuppositions of communication,’ but I prefer to speak of general presuppositions of communicative action because I take the type of action aimed at reaching understanding to be fundamental.” In both understandings, the thinkers conceptualized truth as subjective and personal, something that was the result of interactions and ultimate agreements among free individuals. Thus, discursive thought finds itself in fundamental opposition with a central assumption of the Enlightenment-based understanding that truth is universal and pre-existent. Milton, for example, contended that “[t]ruth be in the field, we do injuriously by licensing and prohibiting to misdoubt her strength. Let her and Falsehood grapple; who ever knew Truth put to the worse in a free and open encounter?” Similarly, French philosopher Jean-Jacques Rousseau contended that “falsehood is capable of an infinite variety of combinations; but the truth has only one manner of being.” Enlightenment thought, as considered earlier, assumed that truth must be found. It is universal and pre-existent. It is simply waiting to be discovered.

The more collective, discourse-based assumptions regarding truth carry with them the requirement that individuals have access to quality information so that they can be informed and carry shared basic understandings regarding matters of public concern into discussions with others. Such an assumption was crucial to Habermas’s

the United States, in What’s Left of Enlightenment: A Postmodern Question 8–11 (Keith Michael Baker & Peter Hanns Reill eds., 2001); Schroeder, supra note 66, at 398–401.


180 See Dewey, supra note 179, at 154–55; Habermas, supra note 179, at 3–5.

181 Dewey, supra note 179, at 158.

182 See Jürgen Habermas, The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society 14–26 (Thomas Burger & Frederick Lawrence trans., 1989); Habermas, supra note 177, at 11–14.


184 See Dewey, supra note 179, at 158; Habermas, supra note 183, at 21.

185 Milton, supra note 3, at 50.

186 Rousseau, supra note 3, at 159.
conceptualization of the public sphere. The sphere, which Habermas described as “the sphere” of “public authority” and “the sphere of private people come together as a public,” flourished when newspapers provided information about matters of public concern and individuals engaged with the news and used it to enter into interactions with others with the goal of coming to a mutual agreement, or truth, regarding an issue or concern. Of course, Enlightenment thought has generally assumed, as has most consistently been stated in the marketplace-of-ideas theory, that in an open exchange of ideas, rational individuals will identify the truth, separating it from false ideas. The false ideas, therefore, will fall aside, just as poor products fail to find a consistent market. The discourse model emphasizes protecting the formation of truths that arise through interactions among informed individuals while Enlightenment-based conceptualizations of the flow of information generally understand society’s challenge as identifying the truth amidst a sea of falsities.

Of course, other important assumptions go into considering the two approaches. The Enlightenment and discourse models differ regarding how they conceptualize the place of the individual in society. Traditional Enlightenment thought has generally conceptualized society as being created to benefit the individual, while discursive approaches emphasize the value of collective action and ability of each individual to contribute to democratic deliberation. Dewey, for example, explained that “a good citizen finds his conduct as a member of a political group enriching and enriched by his participation in family life, industry, scientific and artistic associations.” Importantly, discursive thought emphasizes community because its assumptions are founded upon the importance of free individuals coming together to come to mutual agreements regarding problems and concerns. These differences between these assumptions about the nature of truth and community also fundamentally influence how freedom of expression should be understood in democratic society. Both Enlightenment and discursive thought are constructed upon the assumption that freedom of expression is fundamental to self-government. Dewey emphasized the need for “free gatherings of neighbors” and “gatherings of friends in[ ] living rooms.”

188 See id. at 11, 27.
189 See HABERMAS, supra note 179, at 400, 470.
191 See DEWEY, supra note 179, at 154–55; HABERMAS, supra note 182, at 3–9; Siebert, supra note 130, at 41–42.
192 DEWEY, supra note 179, at 148.
193 See id. at 147–48.
194 See id. at 167–68.
Similarly, Habermas contended that individuals must have access to a “diversity of independent mass media, and a general access of inclusive mass audiences to the public sphere.” Of course, Enlightenment thinkers came to similar conclusions in the seventeenth and eighteenth centuries. Milton, in *Areopagitica*, railed against governmental limitations on freedom of expression. Similarly, other Enlightenment thinkers, such as Locke and Rousseau, emphasized the retention of human liberty in the face of the institution of civil government. Rousseau, for example, contended that individuals in democratic society give up “only such part of his powers, goods, and liberty as it is important for the community to control.”

Crucially, the assumptions within these two approaches regarding how the flow of information is understood, and the very purpose behind freedom of expression, can substantially influence how protections for AI communicators should be conceptualized. This dichotomy between the theories begs the question whether expression should be protected so that fully formed truths can be separated from falsity, or if expression must be safeguarded to allow free individuals to come together to conduct discourse and come to mutual agreements, or truths, about matters of public concern. In the Enlightenment model, information flows into a marketplace, where rational individuals assess whether it is true or false and whether or not the idea should be accepted. Conceivably, this model requires examining the extent to which AI communicators can contribute to such a marketplace. Conversely, the discourse model must consider such entities based on two points along the communicative process. First, whether they can convey information that provides a baseline for individuals to enter into informed discourse with others. Second, whether they themselves can contribute to such discourse by interacting with citizens. These considerations, along with the larger question regarding the rights of AI communicators, require considering the existent legal structures regarding such entities in the United States and Europe.

## III. AI Speech Rights and the U.S. and European Systems

Neither the European nor the U.S. legal systems have directly addressed the question of freedom of expression rights for AI actors.

The ways in which the legal systems have interpreted freedom of expression questions more generally, particularly in regard to other non-human actors, provide guidance regarding how jurists might understand AI communicators’ rights. Furthermore,

---

197 MILTON, supra note 3, at 45–46.
198 See generally LOCKE, supra note 3; ROUSSEAU, supra note 3.
199 ROUSSEAU, supra note 3, at xliii.
200 See generally LOCKE, supra note 3; ROUSSEAU, supra note 3.
statements found in international treaties and covenants, primarily in regard to European courts, provide insights regarding free expression rights for AI actors.

A. The U.S. System: Human, Corporation, or Animal?

The United States’ legal system has generally remained silent about AI communicators. In *American Library Association, Inc. v. United States* in 2002, a federal district court considered the constitutionality of the Children’s Internet Protection Act of 2000. The law required public schools and libraries to place filtering software on the computers that are made available to students and patrons. Libraries and schools that failed to comply would not receive certain federal grants. The judge ruled that the law violated the First Amendment, reasoning that filtering algorithms cannot identify “many distinctions between types of content that would be obvious to a human. And of critical importance, no presently conceivable technology can . . . determine whether a visual depiction fits the legal definitions of obscenity, child pornography, or harmful to minors.” Essentially, he reasoned that the lack of human characteristics equated to a limitation on First Amendment protections for expression. In the appeal that followed, the U.S. Supreme Court found the law did not violate the First Amendment. While Justice David Souter mirrored the district court judge’s reasoning in his dissent, none of the opinions in the decision referred to filtering software as AI. Thus, the nation’s highest court has not, despite the opportunity, specifically addressed algorithms.

More recently, when hackers stole and published the names of more than thirty-seven million Ashley Madison users, the class action lawsuits that followed included references to bots. Users claimed that the website, which assists customers in their efforts to cheat on their spouses, “made extensive use of artificial intelligence ‘bots’ and other mechanisms to mimic fake users (specifically, female users) on the Ashley Madison website in order to induce actual (predominantly male) users to make purchases.” The Ashley Madison lawsuits were ultimately settled, but the question

---

203 See id.
204 Am. Library Ass’n, 201 F. Supp. 2d at 433.
205 See id. at 408, 410–11.
207 See id. at 231–43 (Souter, J., dissenting).
209 Id. at *8.
of rights for AI communicators in the United States has not been. The courts, however, have provided some important building blocks in this regard.

1. Algorithms and Expression

Perhaps the most direct considerations of AI communicators have come in a pair of federal district court decisions that considered whether the results of searches using search engines such as Google and Bing are a form of protected speech. Both cases considered how algorithms work and the corporations’ rights, rather than whether the algorithms themselves enjoyed any particular protections. In 2003, Search King sued Google after its place in the search engine’s PageRank system was downgraded. The firm contended that Google’s algorithm’s decision to reduce the site’s relative importance in queries “impacted the business opportunities available . . . by limiting their exposure on Google’s search engine.” The judge concluded that the lists of ranked websites that search engine algorithms produce are a form of opinion, and are therefore protected First Amendment speech. The judge reasoned that, “PageRanks do not contain provably false connotations. PageRanks are opinions—opinions of the significance of particular web sites as they correspond to a search query.” Quite paradoxically, the judge found the process by which the search results are ranked to be objective, but the results that come from the process to be subjective. She explained that “the result, which is the PageRank—or the numerical representation of relative significance of a particular web site—is fundamentally subjective.” Such a conclusion appears to indicate that AI actors, even those that function by relatively rudimentary computer programming, can be found to have human-like rights. Importantly, however, the question before the court in the case was focused on the corporation’s rights rather than the AI communicator’s. Interestingly, four years later, a federal court in Delaware came to a similar conclusion, but did so using significantly different reasoning. In Langdon v. Google, Inc., a North Carolina website operator contended that Google violated his First Amendment rights when it did not allow him to run advertisements for his websites and removed his pages from its potential search results. The court found that compelling Google

212 See Langdon, 474 F. Supp. 2d at 627; Search King, Inc., 2003 WL 21464568, at *1.
214 Id. at *4.
215 See id. at *3.
216 Id. at *4.
217 See id. at *3.
218 Id.
219 See id. at *3–4.
to adjust its search results and advertising policies would equate to compelling the corporation to speak.221 Citing the Supreme Court’s ruling in Miami Herald Publishing Co. v. Tornillo regarding a right-of-reply law that would have required newspapers to print responses from those they criticized, the court explained that the First Amendment does not allow such directives.222 Interestingly, the Langdon decision pays little attention to the nature of search results, never mentioning the algorithms. Instead, the court emphasized that the corporation cannot be compelled to “speak,” as the petitioner sought.223 Thus, in both algorithm-centric cases, federal judges found limiting the search outputs would violate the First Amendment. Legal scholar Eugene Volokh and attorney Donald Falk came to similar conclusions regarding these cases in a white paper they wrote for Google, contending that algorithm-based search results constitute a form of editorial judgment, which makes them analogous to newspapers or book publishers.224 Legal scholar Tim Wu, however, contended that an algorithm does not express ideas, it simply gathers existent data, or ideas, in response to a query.225 In this conceptualization, it is a vehicle, not a speaker.226 Both approaches, as well as the courts’ decisions, provide building blocks regarding how we should understand the rights of AI communicators.

2. Corporations and Blackie the Cat

Search King and Langdon, while two of the most AI-focused cases thus far in the United States court system, were not principally about the rights of such communicators.227 Thus, to address the problem posed in this Article, we must turn to cases that involve other forms of non-human actors. In particular, the courts have provided strong precedents regarding corporate speech and the rights of animals. In terms of animals, a federal appeals court ruled in 1983 that a couple suing on behalf of the First Amendment rights of their talking cat had no real claim.228 Carl and Elaine Miles had taught their cat, Blackie, to say “I love you” and “I want my Mama.”229 When the city of Augusta, Georgia, required them to buy a business permit because they were collecting money for their cat’s performances, they contended their First Amendment

221 See id. at 629–30.
222 See Langdon, 474 F. Supp. 2d at 630 (citing Miami Herald Publ’g Co. v. Tornillo, 418 U.S. 241, 256 (1974)).
223 See id. at 629–30.
225 See Wu, supra note 8, at 1529–30.
226 See id.
228 Miles v. City Council, 710 F.2d 1542, 1544 n.5 (11th Cir. 1983).
rights were violated. The circuit court was not convinced. The court concluded, in a footnote at the end of its opinion, that “[t]his Court will not hear a claim that Blackie’s right to free speech has been infringed. . . . [H]e cannot be considered a ‘person’ and is therefore not protected by the Bill of Rights.” The court explained, finally, that, “if Blackie had such a right . . . [he] can clearly speak for himself.”

More recent rulings have reinforced the court’s reasoning regarding Blackie’s rights. In 2004, animal rights groups sued the Bush Administration on behalf of the world’s whales, porpoises, and dolphins. Similarly, animal rights groups sued Sea World in 2012 on behalf of Tilikum, the focus of the award-winning documentary “Blackfish,” and four other orcas. The group contended the whales’ captivity violated the Thirteenth Amendment, which prohibits slavery. The judge, much as was the case with Blackie the Cat, concluded that the Thirteenth Amendment “applies to humans, and not orcas.”

Finally, in Naruto v. Slater, better known as the “Monkey Selfie” Case, the Ninth Circuit concluded that Naruto, the seven-year-old crested macaque that picked up photographer David Slater’s unattended camera and took several pictures of himself, could not claim authorship of the images under U.S. copyright law. The court reasoned, citing the Cetacean Community v. Bush and Tilikum v. Sea World Parks and Entertainment, Inc., that “this monkey—and all animals, since they are not human—lacks statutory standing under the Copyright Act.” Thus, these cases suggest American courts have generally not been amenable to extending human-like rights to non-humans. Fundamentally, judges communicated the expectation in these cases that those who claimed fundamental human rights be, well, human. This conclusion, however, is contradicted in one important area of the law.

---

230 Id. at 351.
231 Miles, 710 F.2d at 1544 n.5.
232 Id.
233 Cetacean Cmty. v. Bush, 386 F.3d 1169, 1171 (9th Cir. 2004).
234 Id. at 1179 (quoting Citizens to End Animal Suffering & Exploitation, Inc. v. New England Aquarium, 836 F. Supp. 45, 49 (D. Mass. 1993)).
236 Tilikum, 842 F. Supp. 2d at 1261.
237 Id. at 1264–65.
238 See 888 F.3d 418, 421 n.3 (9th Cir. 2018) (discussing PETA’s blog post which refers to the litigation as the “Monkey Selfie” Case).
239 Id. at 420.
240 386 F.3d 1169 (9th Cir. 2004).
242 Naruto, 888 F.3d at 420.
Since the 1970s, the Supreme Court has consistently concluded that corporations, another type of artificial entity, have freedom-of-expression rights similar to those enjoyed by citizens.\(^{243}\) Importantly, the Court reasoned that corporations can contribute to discourse in important ways, a consideration that was not present in the animal-related cases. In *First National Bank v. Bellotti*, a 1978 case that dealt with whether a Massachusetts law that limited corporations from taking part in discourse was constitutional, the Court concluded that, “[i]f the speakers here were not corporations, no one would suggest that the State could silence their proposed speech. It is the type of speech indispensable to decisionmaking in a democracy, and this is no less true because the speech comes from a corporation rather than an individual.”\(^{244}\) In the same passage, the Court reasoned “[t]he inherent worth of the speech in terms of its capacity for informing the public does not depend upon the identity of its source.”\(^{245}\) The Court, in striking down the law, overturned the state’s highest court’s ruling to uphold the law and limit corporations’ speech rights.\(^{246}\) The Supreme Judicial Court of Massachusetts’s ruling, which was in many ways similar to the reasoning found in the animal rights cases, explained, “It seems clear to us that a corporation does not have the same First Amendment rights to free speech as those of a natural person . . . .”\(^{247}\)

In 2010, the Supreme Court reaffirmed and expanded its conclusions from *Bellotti* in *Citizens United v. FEC*.\(^{248}\) Justice Anthony Kennedy, in writing the opinion for a deeply divided Court, explained that, “on certain topics corporations may possess valuable expertise, leaving them the best equipped to point out errors or fallacies in speech of all sorts, including the speech of candidates and elected officials.”\(^{249}\) Crucially, “artificial intelligence” could easily be substituted for “[c]orporations” in Justice Kennedy’s opinion for the Court.\(^{250}\) In his concurring opinion in the same case, Justice Scalia explained that the First Amendment was “written in terms of ‘speech,’ not speakers. Its text offers no foothold for excluding any category of speaker.”\(^{251}\) His conclusion that the courts should not concern themselves with *who* is speaking provides further support for corporate-speech-based protections for AI communicators. Using this line of precedents from the Court, the potential AI communicators have to benefit public discourse and their ability to act as speakers, despite their fundamental lack of human characteristics, could be enough for the Court to justify treating them similarly to citizens in First Amendment cases.


\(^{244}\) 435 U.S. at 777.

\(^{245}\) Id. at 777.

\(^{246}\) Id. at 767.


\(^{248}\) 558 U.S. at 340 (citing *Bellotti*, 435 U.S. at 784).

\(^{249}\) Id. at 364.

\(^{250}\) Id.

\(^{251}\) Id. at 392–93 (Scalia, J., concurring).
B. AI Speech Rights in Europe

Much like the American court system, the European Court of Justice (ECJ) and the European Court of Human Rights (ECHR) have not squarely addressed the rights of AI communicators.252 Several ECJ applications have made passing references to algorithms since the early 2000s, but the nature of such tools and stronger, higher-functioning programs, have not been central or controlling in any of the instances, particularly in regard to freedom of expression as it is outlined in article 10 of the European Convention on Human Rights (ECHR).253 The most notable discussion of algorithms came in Google Spain v. AEPD in 2014, which involved a challenge put forth by the nation’s data protection agency against Google’s practice of gathering information from those who conduct searches using its services.254 Ultimately, the Court found that Google must, under parts of a few different sections of European Parliament law and the ECHR, fulfill requests made by individuals to remove information that appears about them when their name is searched.255 The advocate general’s opinion devoted a few different passages to the nature of the algorithms that are involved in Google’s search engine.256 The advocate general wrote, “Google’s search engine’s crawler function, called ‘googlebot’, crawls on the internet constantly and systematically and, advancing from one source web page to another on the basis of hyperlinks between the pages, requests the visited sites to send to it a copy of the visited page.”257 As with the United States cases, the judgment and the opinion were not focused on the nature of AI communicators. They were instead focused on the “right to be forgotten,” which gained substantial attention after this decision.258

While the courts have not directly considered the rights of AI communicators, the Legal Affairs Committee of the European Parliament voted in January 2017 to urge the European Commission to create legislation that explicitly recognizes that “robots’ autonomy raises the question of their nature in the light of the existing legal categories—of whether they should be regarded as natural persons, legal persons, animals or objects—or whether a new category should be created.”259 The sweeping report also raised questions about liability for physical injuries caused by robots, the

252 Searches for “artificial intelligence,” “bot,” “chatbot,” “AI,” “robot,” and other key words did not return any relevant applications on the European Court of Human Rights’ case search website: https://hudoc.echr.coe.int [https://perma.cc/65ZS-XXGQ].

253 ECHR, supra note 139, art. 10(1).


255 Id. at 21.


257 Id. ¶ 73.


potential impact the technologies will have on the workforce, and copyright and data-ownership concerns. Importantly, the report further recognized that AI should be regulated based on its impact upon human rights and that, “increasing communication and interaction with robots have the potential to profoundly impact physical and moral relations in our society.” Though the report balanced substantial concerns about the impact and influence of AI communicators with recognitions that they can also be a benefit to society, it did not convey the types of understandings in the American corporate-speech cases, in which the Supreme Court rationalized its decision to extend protections to artificial entities, in this case corporations, because they could contribute meaningful information to democratic discourse. Ultimately, the European Parliament has created a robotics committee, which continues to investigate legal, ethical, and commercial concerns as they relate to the development and incorporation of AI into citizens’ lives.

The ECtHR has not been silent, however, regarding corporate speech. Precedents from this court extend similar protections to non-human communicators as compared to the American cases, but the legal questions that led to such applications arose from substantially different circumstances. While United States corporate-speech cases generally arose from campaign finance law and political speech questions, European decisions have primarily dealt with the nature of for-profit media companies. In Autronic AG v. Switzerland in 1990, a broadcast television company applied for the rights to rebroadcast content from a Russian television satellite. A Swiss agency rejected the company’s request and, ultimately, its appeal reached the ECtHR as an article 10 concern. The court ruled in favor of the corporation, reasoning, “[N]either Autronic AG’s legal status as a limited company nor the fact that its activities were commercial nor the intrinsic nature of freedom of expression can deprive Autronic AG of the protection of Article 10.” Crucially, in the same passage, the court explained that article 10 “applies to ‘everyone’, whether natural or legal persons. . . . Article 10[ ] applies not only to the content of information but also to the means of transmission or reception since any restriction imposed on the means necessarily interferes with the right to receive and impart information.” A year earlier, the court came to a similar decision in a case that involved a retail-oriented

260 Id. at 3–4.
261 Id. at 22.
262 Id. at 16.
263 See supra Section III.A.2.
265 See supra Section III.A.2.
267 Id. at 15.
268 Id. at 17.
269 Id.
print publication based in Germany. In the face of a government-imposed injunction against further publication, the company argued that its article 10 rights were violated. The government contended that the publication was a for-profit institution, not a news organization, and thus could not claim the restraint on publication was an article 10 matter. The Court found, “The applicants did not deny that they defended the interests of the specialised retail trade. . . . To restrict the freedom of expression to news items of a political or cultural nature would result in depriving a large proportion of the press of any protection.” Thus, the court constructed a similar line of reasoning in this case as the American Supreme Court did in Citizens United v. FEC and other corporate-speech-related cases. The judges reasoned that to limit this organization’s rights because it was not a traditional news organization would endanger the rights of journalists. Chief Justice John Roberts opened his concurring opinion in Citizens United with the similar concern that letting the campaign finance laws in question stand would “empower the Government to prohibit newspapers from running editorials or opinion pieces supporting or opposing candidates for office, so long as the newspapers were owned by corporations.”

Thus, when it comes to these cases, the two legal systems share a common allowance for one particular “artificial legal entity”—corporations, which receive freedom-of-expression rights that have historically only otherwise been associated with human communicators. While the corporate-speech-rights precedents are not identical, the two court systems communicated understandings that such non-human speakers should be protected because they contribute to discourse. Also, neither system has substantially considered questions that surround the rights of AI communicators. The crucial difference, however, is that the European Parliament has moved forward with a document—however limited in legislative power it might be—that recognizes many concerns and is beginning a dialogue regarding how to approach the challenges that, as Web 4.0 arrives, will likely substantially affect discourse within society.

---

271 Id. at 11.
272 See id. at 12.
273 Id.
277 Citizens United, 558 U.S. at 373 (Roberts, C.J., concurring).
278 Id. at 390 (Scalia, J., concurring).
C. International Treaties and Non-human Entities

While the European Parliament has taken steps toward understanding the growing influence of AI communicators on democratic discourse, neither it nor other governing bodies related with the European Union or the United States have created laws, guidelines, directives, or other tools regarding how communicators and jurists should understand the questions that surround the extent to which freedom of expression protections should be extended to AI communicators. As with the preceding sections, in the absence of direct statements regarding such questions, this Section examines the extent to which existent covenants, many of which have been signed by the United States and European nations, outline how such rights for non-human entities might be understood. Importantly, while many European nations’ constitutions have allowances for international law built into their rights and protections,281 the U.S. Constitution includes no such expectation. The “Basic Law for the Federal Republic of Germany,” for example, proscribes “[t]he general rules of international law shall be an integral part of federal law. They shall take precedence over the laws and directly create rights and duties for the inhabitants of the federal territory.”282 Similarly, the French and Spanish constitutions outline the role of international law in the nation’s society.283

Thus, while the U.S. and European nations have signed and ratified or voted for the United Nations’s Universal Declaration of Human Rights (UDHR),284 the International Convention on the Elimination of All Forms of Racial Discrimination (CERD),285 and the International Covenant on Civil and Political Rights (ICCPR),286 these documents are far more likely to influence European courts’ decisions than those in the U.S.287 The U.S. Supreme Court in Sosa v. Alvarez-Machain in 2004, for example, emphasized that the UDHR, though it “does bind the United States as

282 GRUNDEGESETZ [GG] [Basic Law], art. 25.
a matter of international law, the United States ratified the Covenant on the express understanding that it was not self-executing and so did not itself create obligations enforceable in the federal courts.\textsuperscript{288} Four years later, in \textit{Medellin v. Texas}, the Court repeated that, absent congressional direction regarding international law, claims based on non-domestic covenants and agreements will not be compelling.\textsuperscript{289} Despite this theme, however, the international covenants provide important perspectives regarding the potential scope and limitations of freedom of expression for non-human entities on global levels.

1. Limited Room for Human Rights for Non-human Entities

All four of these documents explicitly and specifically focus on human rights. As one scholar indicated, “[A]ll thirty articles of the Universal Declaration of Human Rights are predicated on the individual as the subject of human rights.”\textsuperscript{290} The preamble of the ICCPR, for example, states that the document focuses on “the equal and inalienable rights of all members of the human family,” characterizing them as “the foundation of freedom, justice and peace in the world.”\textsuperscript{291} Similarly, CERD emphasizes “all human beings are equal before the law.”\textsuperscript{292} Alternatively, the documents do not use wording such as “entities” or “artificial persons,” which might cause confusion about whether the promised rights extend to AI or other non-human actors. Despite this theme, courts’ interpretations of the ECHR have created some limited openings for non-human entities, primarily in the form of corporations, to receive the promised protections.\textsuperscript{293} Article 10 explains that “[e]veryone has the right to freedom of expression.”\textsuperscript{294} Its wording, at the outset, is nearly the same as the UDHR and ICCPR’s article 19 and CERD’s article 5.\textsuperscript{295} Unlike the others, however, the ECHR continues by allowing nations to require licensing for broadcast, television, and cinema “enterprises.”\textsuperscript{296} As indicated with the \textit{Autronic AG}\textsuperscript{297} and \textit{Markt Intern Verlag GMBH}\textsuperscript{298} cases, this aspect of article 10 has led to support in the

\textsuperscript{288} 542 U.S. 692, 735 (2004).
\textsuperscript{289} 554 U.S. 759, 760 (2008) (per curiam).
\textsuperscript{290} Anna Yeatman, \textit{Who Is the Subject of Human Rights?}, 43 AM. BEHAV. SCIENTIST 1498, 1505 (2000); see also UDHR, supra note 284, arts. 1–30.
\textsuperscript{291} ICCPR, supra note 286, at preamble.
\textsuperscript{292} CERD, supra note 285, at preamble.
\textsuperscript{293} See supra Section III.B.
\textsuperscript{294} ECHR, supra note 139, art. 10(1).
\textsuperscript{295} See ICCPR, supra note 286, art. 19; CERD, supra note 285, art. 5; UDHR, supra note 284, art. 19.
\textsuperscript{296} ECHR, supra note 139, art. 10(1); see also Peter Oliver, \textit{Companies and Their Fundamental Rights: A Comparative Perspective}, 64 INT’L & COMP. L. Q. 661, 677 (2015).
courts for corporate speech protections. The ECtHR has also extended article 8–based privacy protections to corporations. Article 8 ensures that “everyone has the right to respect for his private and family life, his home and his correspondence.” In Société Colas Est v. France and Niemietz v. Germany, the court interpreted article 8 as extending beyond individual citizens’ homes to businesses. In Société Colas Est, a 2002 judgment that dealt with French officials’ raids of businesses during a corruption investigation, the court reasoned that “it has previously held, the word ‘domicile’ (in the French version of Article 8) has a broader connotation than the word ‘home’ and may extend, for example, to a professional person’s office.” Perhaps more importantly, in the ensuing paragraph, the court explained that the ECHR “is a living instrument which must be interpreted in the light of present-day conditions.” Finally, article 1 of Protocol I, which was added shortly after the initial ECHR document was drafted, indicates that “[e]very natural or legal person is entitled to the peaceful enjoyment of his possessions.” Such a statement opens the door for non-human entities to be classified as a “legal person,” and to therefore receive rights.

Importantly, the preceding examples identify instances where the wording found in the ECHR left room for the courts to interpret otherwise human-focused rights as extending to non-human entities. None of these instances dealt with AI communicators. Also, the other three documents’ statements regarding human rights include little, if any, indication that any such protections, particularly in regard to freedom of expression, should be extended to AI communicators. As sociologist Anna Yeatman concluded in her analysis of human rights treaties and covenants, AI may deserve some form of rights, but “these will not be human rights, but the rights of such entities.” While this is one possible conclusion, the inherent interaction between humans and AI communicators within societal discourse makes the international human rights documents an important contributor to questions regarding free expression safeguards for AI.

2. Balancing Public Need and Free Expression Rights

The international covenants provide significant support for lawmakers and different nations’ courts to curtail the ability of AI communicators to take part in
public discourse, particularly when these non-human entities put forth unpopular ideas. Crucially, this theme turns on the substantially different constructions of freedom of expression in the United States and within the European nations. These different interpretations are reflected within the international treaties, which, as previously indicated, are substantially more aligned with traditional European legal understandings. The documents all include relatively strong statements regarding the importance of freedom of expression, which, on the surface makes them similar to the First Amendment freedoms that are outlined in the Bill of Rights. Importantly, however, the international documents generally include caveats that indicate that such expression must be done responsibly and respectfully. Article 19 of the ICCPR, for example, notes that “[e]veryone shall have the right to hold opinions without interference.” The third paragraph of the Article, however, indicates that the right to free expression “carries with it special duties and responsibilities” and “it may therefore be subject to certain restrictions.” Article 19 of the UDHR also safeguards free expression, but article 29 indicates that “[e]veryone has duties to the community” and that everyone must have “respect for the rights and freedoms of others.” The passage continues by explaining that morality, public order, and societal welfare are crucial goals to democratic society. These passages in the UDHR are mirrored in the other documents, which generally, in sections that are separate from where freedom of expression is outlined, include proscriptions against racist, immoral, or other communication that is deemed unhealthy to democratic society. CERD outlines the expectation that “all dissemination of ideas based on racial superiority or . . . against any race or group of persons of another colour or ethnic origin” must be criminalized. The ICCPR explains that freedom of expression can be limited “[f]or the protection of national security or of public order[,] or of public health or morals.” Article 10 of the ECHR includes similar concerns. It ensures that “[e]veryone has the right to freedom of expression . . . without interference by public authority and regardless of frontiers.” This passage, however, is followed

---

308 See generally ECHR, supra note 139; ICCPR, supra note 286; CERD, supra note 285; UDHR, supra note 284.

309 ECHR, supra note 139, art. 10(1); ICCPR, supra note 286, art. 19; CERD, supra note 285, art. 5; UDHR, supra note 284, art. 19.

310 See U.S. CONST. amend. I.

311 ICCPR, supra note 286, art. 19(1).

312 Id. art. 19(3).

313 UDHR, supra note 284, arts. 19, 29.

314 Id. art. 29(2).

315 See CERD, supra note 285, art. 4; ICCPR, supra note 286, art. 20(2); ECHR, supra note 139, art. 10.

316 CERD, supra note 285, art. 4.

317 ICCPR, supra note 286, art. 19(3)(b).

318 See ECHR, supra note 139, art. 10.

319 Id. art. 10(1).
by expectations that citizens have “duties and responsibilities” regarding “formalities, conditions, restrictions or penalties as . . . are necessary in a democratic society.”

Interpretations of the First Amendment generally do not include any parallels in this regard. Instead, as it is perhaps most concisely expressed in Texas v. Johnson, a 1989 U.S. Supreme Court case in which justices struck down a law that criminalized flag-burning, the government generally cannot, in keeping with the First Amendment, make content-based restrictions on expression. Justice William Brennan characterized the “bedrock principle” of the First Amendment as “the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.” Years earlier, in a defamation case, Justice Lewis Powell provided a similar interpretation by invoking the marketplace concept. He concluded, “Under the First Amendment there is no such thing as a false idea. However pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries but on the competition of other ideas.” The Court has upheld and expanded upon these rulings in more recent cases that involved actions and ideas that the international covenants almost certainly sought to discourage. In particular, the Court in 2011 upheld Westboro Baptist Church’s right to protest using messages such as “God Hates Fags” and “Thank God for Dead Soldiers.” Eight years earlier, the Court struck down a state law that criminalized cross burning, despite the practice’s long history in the United States as a tool to threaten and intimidate a group of people because of their race. The deeply divided Court allowed that cross burning with the intent to intimidate a certain group could be criminalized, but the practice in general could not be. Justice Sandra Day O’Connor, in writing for the Court, reasoned, “It may be true that a cross burning, even at a political rally, arouses a sense of anger or hatred among the vast majority of citizens who see a burning cross. But this sense of anger or hatred is not sufficient to ban all cross burnings.” Such rulings provide a substantially different approach to limitations on expression, particularly as they are outlined in international covenants and how European courts have interpreted them.

European efforts to limit the types of speech that were at the heart of U.S. cases such as the Westboro Baptist and cross-burning decisions have led to struggles regarding how to both safeguard expression while still limiting immoral, unhealthy, or racist ideas. In Mouvement Raëlien Suisse, the details of which are discussed in

320 Id. art. 10(2).
322 Id. at 414.
324 Id.
327 Id. at 366.
328 Id.
Part II, the ECtHR rationalized its decision to uphold Swiss authorities’ rights to halt an organization’s poster—not because of its content, but because of the organization’s ideas. The court agreed that the poster did not include anything “unlawful or likely to offend the general public,” but found that the local authorities had wide latitude to pursue “the legitimate aims of the prevention of crime, the protection of health and morals, and the protection of the rights of others.” The court indicated that the content of the group’s website could offend “the religious beliefs of certain persons” and, later in its reasoning, that the concept of an ideal genetic society “might offend the democratic and anti-discriminatory beliefs” on which democratic society is based. In Delfi AS, the court rationalized its decision to hold the service responsible for the comments users posted via its message boards by concluding a balance must be struck between freedom of expression and irresponsible communication. The court reasoned that “the case concerns the ‘duties and responsibilities’ of Internet news portals,” as outlined in article 10, section 2, when users “engage in clearly unlawful speech, which infringes the personality rights of others and amounts to hate speech and incitement to violence against them.” Ultimately, the court found that Delfi had not done enough to protect its audiences and that its claims that its article 10 protections had been violated were not valid.

A year later, the fourth section of the ECtHR faced a similar question regarding the extent to which forum providers are liable for how their services are used in MTE v. Hungary. Once again, the court sought to balance the free expression rights of the forum holders with concerns regarding the falsity, offensiveness, and reputational impact the ideas that were published might have. The court found that third-party providers, such as the industry self-regulating body and the for-profit company involved in the case “assume duties and responsibilities.” Ultimately, the court concluded that Hungarian courts had limited the third party providers’ rights when they held them liable for the content that was published on their forums.

---

330 Id.
331 Id. at 391.
332 Id. at 394–95.
334 Id. at 345 (quoting Declaration on Freedom of Communication on the Internet, Council of Europe, May 28, 2003).
335 Id. at 372.
336 Id. at 388.
338 Id. at 3.
339 Id. at 14–15.
340 Id. at 21–22.
While the court in *MTE* sided with free expression, it still emphasized that communicators are responsible for conveying ideas that are moral and beneficial to society. Such concerns generally have had no place in free expression cases in the United States. Finally, in 2013 the same ECtHR upheld a British administrator’s decision to reject for broadcast an animal rights group’s advertisement. The court in *Animal Defenders International v. United Kingdom* reasoned that it must balance the NGO’s rights with the nation’s ability to “mould its own democratic vision.” The court explained that it was required “to balance, on the one hand, the applicant NGO’s right to impart information and ideas of general interest . . . with, on the other, the authorities’ desire to protect the democratic debate and process from distortion.”

Beyond these balancing efforts, the court in these judgments consistently communicated that it understood the government as the keeper of the public debate. It also conveyed that it conceptualized the goal of the ECHR and similar efforts as creating a pluralist society. The state-as-caretaker role was evident in *Mouvement Raëlien Suisse*, where the court acknowledged, and ultimately upheld, the officials’ concern that the State “did not wish its name to be associated with certain non-majority but lawful ideas.” Similarly, this concern was central to the *Animal Defenders* judgment, where the court reasoned that individual nations, because of its knowledge of “societies and their needs . . . are best placed to assess the particular difficulties in safeguarding the democratic order in their State.” In several cases, the government’s role as a steward for public debate manifests itself within the “pressing social need” approach, which judges attributed to article 10, section 2. The court indicated that if the limitation on expression was found to be a “pressing social need,” it could be upheld. This was evident in *Hertel v. Switzerland*, a 1997 case that involved a journal article that made claims that damaged the appliance industry. The Court concluded that restraining the author from communicating his ideas did not represent “pressing social need.” Similarly, in *Steel and Morris v. United Kingdom*, a

---

341 *Id.* at 22.
343 *Id.* at 235.
344 *Id.*
345 See *id.* at 234.
346 See *id.* at 235.
352 *Id.* at 31–35.
2005 chamber judgment regarding whether a Greenpeace pamphlet defamed the McDonald’s Corporation, the Court turned to the “pressing social need” standard in rejecting the fast-food giant’s claims. Such a standard, along with the other efforts to balance social concerns with freedom of expression protections have led to substantially different lines of precedent between the U.S. and European legal systems.

Ultimately, the legal systems’ differing approaches to interpreting freedom of expression, particularly in light of the international covenants, create the potential for markedly different frameworks regarding emerging questions about the freedom of expression rights of AI communicators. The European system, which prioritizes cultural pluralism and seeks to balance free expression with public concerns, such as morality and health, includes numerous existing safeguards against many of the potentials harms that AI entities can inflict on democratic discourse. The U.S. system, which pays little attention to the international documents and is far more oriented toward a relatively unfettered marketplace-of-ideas approach, has fewer existing tools for limiting AI expression.

IV. AN ARTIFICIAL MARKETPLACE

The growing presence of AI entities within the global, instantaneous discourses that take place in virtual communities around the world raises fundamental questions about how the freedom of expression rights of an entirely new form of communicator should be understood. Alongside such entities’ increasing presence within our discourse, the nature of our own communication has changed. Within the virtual environments where AI entities are flourishing, people are forming weaker ties, producing less nuanced messages, and grouping into fragmented, like-minded, and intentional communities. These shifts in human discourse cannot be ignored when considering how the rights of non-human communicators should be conceptualized, since these changes have made it easier for AI communicators to take part in human discourse.

While the U.S. and European court systems diverge in fundamental and important ways regarding how they have interpreted safeguards for freedom of expression, both models are likely to struggle as AI communicators become more complex, and as they introduce novel problems into legal systems that have yet to squarely address the rights of such entities, even in their more basic forms. U.S. Supreme Court precedents, which have remained closely associated with the Enlightenment-oriented marketplace model for nearly a century, provide few avenues for lawmakers and jurists to limit the potentially harmful effects AI entities might have on democratic discourse without limiting human free-expression rights. At the same time,

354 See, e.g., Shirky, supra note 49, at 71–75; Sunstein, supra note 51, at 44; Iyengar & Hahn, supra note 51, at 20–21.
355 See, e.g., Massaro & Norton, supra note 5, at 1173–74; Schroeder, supra note 66, at 383–84, 401–02.
however, the marketplace model cannot—in its current form—account for the power of AI entities to use their fundamentally non-human qualities to artificially boost some ideas while drowning out the ability of human ideas to organically compete in the marketplace. They can, essentially create marketplaces where only one product, or truth, is available for consideration.

The U.S. court system could turn to the animal rights cases in seeking a way to delineate protected human expression from that of non-humans. The outcomes in these cases indicated that when non-human actors—animals—claimed rights that have traditionally only been made available to humans, they were consistently found to lack standing.\(^{356}\) They therefore were not afforded human-like protections. These cases, however, fail to account for the potential contributions AI entities can make to discourse. AI communicators are simply, by their nature, not like animals. Animals generally struggle to communicate messages that are created by others and are not capable of creating meaningful messages that contribute ideas to the flow of information. Some AI communicators can not only repeat messages from human actors, but also construct ideas and make them available to citizens.\(^{357}\) These capabilities should only advance over time.

Another option is to approach AI entities as being more like corporations, which have found substantial First Amendment protection in the U.S. and strong freedom of expression support in European courts.\(^{358}\) Corporate-speech protections have been substantially supported because of their abilities, despite being artificial entities, to contribute to democratic discourse.\(^{359}\) A blanket decision to protect all AI actors’ free-expression rights, based on the corporate-speech precedents, however, would open the door to crippling market failure in the marketplace of ideas. With no power to limit AI actors’ messages, these communicators could drown out human ideas, build artificial appearances that some ideas are more accepted than others, and do much of this work while remaining indistinguishable from human communicators online. Thus, a more nuanced approach is needed.

The international covenants that have guided European lawmakers and jurists, and the precedents that courts have put forth as a result of them, have provided a far wider spectrum of tools for managing the impact of AI entities on discourse. European

\(^{356}\) See, e.g., Naruto v. Slater, 888 F.3d 418, 426 (9th Cir. 2018); Tilikum v. Sea World Parks & Entm’t, Inc., 842 F. Supp. 2d 1259, 1264–65 (S.D. Cal. 2012); Cetacean Cmty. v. Bush, 386 F.3d 1169, 1179 (9th Cir. 2004); Miles v. City Council, 710 F.2d 1542, 1544 (11th Cir. 1983).

\(^{357}\) See Massaro & Norton, supra note 5, at 1171–72 (“Such computer speakers also are increasingly self-directed or ‘autonomous’—which is to say, the computer generates content further afield from human direction.”).


\(^{359}\) See, e.g., Bellotti, 435 U.S. at 777.
court decisions have balanced freedom of expression protections with concerns regarding morality, public health, and the preservation of a healthy democratic society. They have also expressed consistent concerns for safeguarding a pluralist and tolerant society. By doing so, lawmakers and courts have constructed more diverse understandings of freedom of expression. These approaches provide some promising building blocks regarding how certain types of AI communication can be limited while human expression remains protected. The difficulty with this approach, however, is that it becomes likely that individual nations will create a patchwork of varyingly rigid restrictions upon AI communicators’ expression. Such a patchwork would, in all likelihood, be relatively ineffective in addressing a challenge that is essentially borderless. Furthermore, the ECtHR’s relatively recent efforts to rule on freedom-of-expression concerns have been based more on case-by-case, balancing approaches, rather than constructing a clear precedential line that would allow lawmakers and jurists in individual nations to understand how the ECHR’s rights are to be interpreted. Indeed, the Court’s careful, deferential approach toward individual nations’ legal concerns, while understandable, works against the broader effort to remedy this global concern.

A potential avenue for addressing this global concern, however, would be to revise the fundamental assumptions of the marketplace model to emphasize the evolution of truth as the result of discourse rather than as a competition between truth and falsity. In other words, the foundational aspects of the marketplace rationale could be shifted from a focus on protecting the discovery of truth to safeguarding the development of truth. By constructing and applying a model that incorporates more discursive assumptions about truth, rather than the more traditional Enlightenment-founded building blocks, the theory helps to create a dividing line between AI that contribute to the development of truth and AI that mislead, create or convey falsity, and simply artificially overwhelm the marketplace. This foundational revision to the theory would create a stronger rationale for protecting the flow of information, rather than the competition of ideas, which in this case would mean safeguarding the potential contributions that AI entities can make to discourse. Thus, lawmakers could regulate AI-based expression that does not contribute or damage the flow of information, but the First Amendment and article 360 See Defi AS v. Estonia, 2015-II Eur. Ct. H.R. 345 (quoting Declaration on Freedom of Communication on the Internet, Council of Europe, May 28, 2003), http://hudoc.echr.coe.int/eng?i=001-126635 [https://perma.cc/DXF3-VFT6]; Animal Defs. Int’l v. United Kingdom, 2013-II Eur. Ct. H.R. 234, http://hudoc.echr.coe.int/eng?i=001-119244 [https://perma.cc/X6W3-7RBX]; Mouvement Raëlien Suisse v. Switzerland, 2012-IV Eur. Ct. H.R. 391, http://hudoc.echr.coe.int/eng?i=001-112165 [https://perma.cc/27NM-XN9U].


362 See, e.g., id. at 235.

363 See DEWEY, supra note 179, at 158; HABERMAS, supra note 179, at 3–7; HABERMAS, supra note 182, at 21.

10 of the ECHR would safeguard AI communicators’ expression on the assumption that halting their contributions would damage discourse. Such an approach aligns with the wording found in other international covenants. Article 19 of the ICCPR emphasizes that “[e]veryone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers.” The ICCPR’s wording substantially draws from the UDHR, which also highlights that people have a right to both express and receive information. Protecting AI that contribute to the flow of information aligns with the right-to-receive aspects of the articles.

Such a shift in the theory’s assumptions regarding truth would also align more with Justice Holmes’s more pragmatic approach to the nature of truth. Justice Holmes, according to his legal and personal writings, did not believe in absolute, objective truth. Instead, he understood truth as being conditional, an approach that, while it does not line up entirely with discursive thought, appears to be more amenable to it than the more Enlightenment-related assumptions that the marketplace approach has come to be associated with. Of course, even if a more discursive approach is used, such an approach still leaves unresolved the crucial question of who will decide which AI-based expression contributes to the flow of information and the eventual emergence of truth and which does not. This is not a simple question. The European courts, along with the international covenants, have established precedents regarding expression that is harmful to individuals or society. These standards were particularly on display in the Mouvement Raëlien Suisse, Delfi, and Animal Defenders judgments. While the reasoning in these judgments does not provide a clear precedent regarding the extent to which certain types of harmful expression can be limited, the rulings, along with the covenants, have at least broken trail in that direction. Similar approaches could be employed for AI communicators,

---

365 ICCPR, supra note 286, art. 19(2).
366 See id.; see also UDHR, supra note 284, art. 19.
367 See HOLMES, supra note 110, at 107; Holmes, supra note 113, at 40–41.
368 While Justice Holmes and Dewey never met in person, they read each other’s work. Justice Holmes commented to a friend that in Dewey’s Experience & Virtue, he “read sentences that I didn’t understand, for his style is horrid, but I thought that I never anywhere had read a philosophical work that felt our Universe so deeply and so widely.” Letter from Oliver Wendell Holmes to Nina Gray (Jan. 2, 1927) (on file with Harvard Law School Digital Suite), http://library.law.harvard.edu/suites/owh/index.php/item/42882378/11 [https://perma.cc/M49U-3MYN]. At the time of his death in 1935, Justice Holmes had three of Dewey’s books on his shelf. See Oliver Wendell Holmes, Jr. Digital Suite (on file with Harvard Law School Library Digital Suite).
which would provide a framework for limiting the false and misleading information that such communicators often interject into discourse.

In the U.S., however, the courts would have to take the step of giving AI communicators fewer rights than humans. The revised, more discursive marketplace approach, provides a theoretical rationale for such a change. Of course, this line of thinking would go against the precedents from the corporate-speech cases, which provided strong support for protecting artificial entities’ First Amendment rights. Of course, corporations were protected because they can contribute to discourse. They are also collections of citizens. The courts could reason that allowing the regulation of AI entities that damage the flow of information does not conflict with the corporate-speech precedents, or the First Amendment more generally, because they are doing quite the opposite regarding the Supreme Court’s rationale from cases such as *Bellotti* and *Citizens United*. These types of AI are not contributing to discourse and the development of truth that comes with it. Thus, AI, in this scenario, would receive lesser freedom-of-expression protections than human communicators, but would still receive limited, conditional safeguards. Such an approach, one that emphasizes protecting the *development* of truth, rather than its *discovery*, aligns with the more discursive marketplace approach. If courts rationalize freedom of expression protections as safeguards for the development of truth, dividing AI communicators between those that contribute to discourse and those that do not becomes a more achievable task.

---


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