Servers and Waiters: What Matters in the Law of A.I.

Ignacio N. Cofone*

21 STAN. TECH. L. REV. 167 (2018)

ABSTRACT

The European Parliament's recent declaration that robots are "electronic persons" illustrates the widespread uncertainty about how to regulate robots and artificial intelligence (A.I.) agents. This article aims to confront that uncertainty. To date, most regulations have treated robots and A.I. agents either as tools or people, making questionable assignments of rights and responsibilities. Instead, regulations should reckon that robots and A.I. agents escape this dichotomy. The law must assign rights and responsibilities for entities with characteristics that exist on a continuum between tools and people. This article describes this continuum through three characteristics that help us consistently place robots and A.I. agents along it: emergence, embodiment, and social valence. It proposes a framework for analogizing A.I. entities to existing entities that the law already understands, thereby creating a baseline for assigning rights and responsibilities for their actions.

* Research Fellow, NYU Information Law Institute. ignacio.cofone@nyu.edu. Many thanks to BJ Ard, Jack Balkin, Ryan Calo, Rebecca Crootof, Samantha Goodwin, Claudia Haupt, Klaus Heine, David Kimel, Kate Klonick, Cody Knapp, Janny Leung, Asaf Lubin, Kathy Strandburg, and the editors of the Stanford Technology Law Review for their helpful comments. Thanks also to Fabian Bargout for his excellent research assistance. The piece also benefited from an internal presentation at the Yale Law School Information Society Project.

TABLE OF CONTENTS

I. INTRODUCTION 168
II. DISRUPTIVE CHARACTERISTICS OF A.I. AGENTS 171
III. A FRAMEWORK FOR COGNIZING A.I. AGENTS: THE CONTINUUM OF ANALOGIES FROM TOOLS TO PERSONS 174
A. Identifying Analogies 174
B. Placing A.I. in a Continuum from Tool to Person 177
IV. COMPLETING THE FRAMEWORK: CHARACTERISTICS THAT MATTER FOR ANALOGIZING A.I. AGENTS 179
A. Embodiment: What is Essential is Invisible to the Eye 179
B. Social Valence as Being Sentient and Felt 181
C. Two Dimensions of Emergence: Agency and Unpredictability 183
V. APPLYING THE FRAMEWORK TO DOCTRINAL A.I. PROBLEMS 186
A. Future Cases: Science Fiction and the Intuitiveness of the Framework 186
B. Tort Law 189
C. Copyright Law 191
D. Criminal Law 194
VI. CONCLUSION 197

I. INTRODUCTION

A European Union (E.U.) commissioner once told me her greatest challenge in crafting technology legislation like the E.U. General Data Protection Regulation: most of her colleagues cannot distinguish between a server and a waiter.[1]

The problem that tortured this tech-savvy politician has never been more present than now. Earlier this year, the European Parliament came closer than ever to conflating the two categories by declaring that, for legal purposes, a robot is "an electronic person."[2]

The European Parliament report argues that

> the more autonomous robots are, the less they can be considered simple tools in the hands of other actors (such as the manufacturer, the owner, the user, etc.)[, and] this,[3] in turn, makes the ordinary rules on liability insufficient and calls for new rules which focus on how a machine can be held – partly or entirely – responsible.[4]

1. For obvious reasons, I will refrain from naming this candid politician.

2. Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics, PARL. EUR. DOC. (PE582.443v01-00) (2016) http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML%2BCOMPARL%2BPE-582.443%2B01%2BDOC%2BPDF%2BV0//EN [https://perma.cc/XCE7-8XUC].

3. *Id.* at 5, recital S.

4. *Id.*

However, the report recommends solving this simply by giving robots personhood, with its rights and obligations under E.U. law.

Beyond ambiguously declaring robots "electronic persons," the report calls for defining "smart autonomous robots," creating a European agency for robotics and artificial intelligence, and developing a reporting system for companies that use them in their work.[5] The precise consequences of these measures are still unclear. There is, moreover, a lack of identification or categorization of the newly regulated entities.

The European Parliament is not alone in its endeavor to identify and regulate artificial intelligence (A.I.) robots.[6] A.I. and robotics law is burgeoning. In the last decade, there have been extensive doctrinal and regulatory debates on how to regulate robots,[7] what entity should regulate them,[8] and what should be the scope of their civil liability,[9] criminal liability,[10] role in wars,[11] free speech rights,[12] rights of authorship,[13]

---

5. *Id.* at 7–8.

6. *See* Ryan Calo, *Robots in American Law*, DRAFT, 5 (2016) (illustrating that this phenomenon is pervasive by arguing that "jurists on the whole possess poor, increasingly outdated views about robots and hence will not be well positioned to address the novel challenges they continue to pose").

7. *See, e.g.*, Ryan Calo, *Open Robotics*, 70 MD. L. REV. 571, 573 (2011) ("In this article, I will advance several hypotheses about the commercial prospects of robotics in the United States. I will argue that to fulfill its enormous promise personal robotics must be sufficiently 'open' to third party innovation and that paving the way toward such openness may require modest legal intervention").

8. *See, e.g.*, RYAN CALO, BROOKINGS INST. CENT. TECH. INNOV., THE CASE FOR A FEDERAL ROBOTICS COMMISSION, 3 (2014) (proposing that a new federal agency should be established to address the experiences and harms that robotics create).

9. Samir Chopra & Laurence White, *Tort Liability for Artificial Agents*, *in* A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS; Curtis Karnow, *The Application of Traditional Tort Theory to Embodied Machine Intelligence*, DRAFT (2013).

10. *See* GABRIEL HALLEVY, LIABILITY FOR CRIMES INVOLVING ARTIFICIAL INTELLIGENCE SYSTEMS (2016); GABRIEL HALLEVY, WHEN ROBOTS KILL: ARTIFICIAL INTELLIGENCE UNDER CRIMINAL LAW (2013); Gabriel Hallevy, *I, Robot—I, Criminal—When Science Fiction Becomes Reality: Legal Liability of AI Robots Committing Criminal Offenses*, 22 SYRACUSE SCI. TECH. L. REV. 1 (2010); Gabriel Hallevy, *The Criminal Liability of Artificial Intelligence Entities*, 4 AKRON INTELL. PROP. J. 171 (2010).

11. *See generally* Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347 (2015); Duncan Hollis, *Setting the Stage: Autonomous Legal Reasoning in International Humanitarian Law*, 30 TEMP. INTL. & COMP. L.J. 1 (2016); Michael Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, HARV. NAT'L. SEC. J. (2013).

12. Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445 (2013); Toni M. Massaro & Helen Norton, *Siri-Ously? Free Speech Rights and Artificial Intelligence*, 110 NW. U.L. REV. 1169 (2015).

13. *See generally* Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, STAN. TECH. L. REV. 5 (2012); James Grimmelmann, *There's No Such Thing as a Computer-Authored Work, and It's a Good Thing Too*, 39 COLUM. J.L. & ARTS 403 (2016); Liza Vertinsky & Todd Rice, *Thinking About Thinking Machines: Implications of Machine Inventors for Patent Law*, 8 B.U. J. SCI. & TECH. L. 574 (2002).

and impact on privacy law.[14] A.I. robots are difficult to categorize under law, and so they disrupt legal decisions and systems that depend on such categorizations because they change the assumptions under which the law was operating.[15] Developing the norms that apply to them is as important as it is inescapable.[16]

This leaves open a wide set of doctrinal questions. Who should be responsible for accidents caused by self-driving cars? Who owns the copyright over work created by an algorithm? Do algorithms have free speech? Can an A.I. agent be responsible for a crime? Can it be an accessory to a crime? To address such regulatory questions coherently as they continually arise calls for a new framework for applying our existing legal principles to these emerging A.I. agents.

In this article, I address this gap. I do so by analyzing how the law should classify robots and A.I. agents along our existing legal categories and, therefore, how it should regulate them.[17] My framework follows Jack Balkin in addressing robots and A.I. agents together because, as I explain below,[18] the main challenges posed by these agents for the law are independent of whether they are embodied.[19] Therefore, I will often refer to robots and A.I. agents indistinctly. By exploring examples of A.I. agents' and robots' impact on relationships normally governed by tort law, I expand on the current categories of "person" or "tool" to propose a framework for determining the rights and obligations that stem from A.I. and robots' behavior.

The European Parliament and the burgeoning U.S. literature have rightfully pointed out that treating robots as either objects or subjects fails to address increasingly common scenarios involving robot entities that do not fall into either extreme. For example, Ryan Calo has demonstrated through nine different case studies that American case law treats robots either as objects or as subjects,[20] and adds that "[t]he mismatch

14. Ryan Calo, *Robots and Privacy*, *in* ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS 187 (2012).

15. *See* Jack Balkin, *The Path of Robotics Law*, 6 CALIF. L. REV. 45, 46 (2015) (demonstrating that robotics and A.I. present two central new problems for the law: how to distribute among humans rights and responsibilities that arise from actions of non-human entities, and how people will substitute A.I. agents for humans in a contextual, unstable and opportunistic way).

16. Calo, *supra* note 6, at 141 (adding that "[t]he robotics revolution will take a coordinated, global effort").

17. Jack Balkin, *The Three Laws of Robotics in the Age of Big Data*, 78 OHIO STATE LAW J. 1217 (2017).

18. See Part IV.A.

19. Balkin, *supra* note 15, at 45 ("I do not distinguish sharply between robots and artificial intelligence [AI] agents"); *Id.* at 51 ("we may be misled if we insist on too sharp a distinction between robotics and AI systems.").

20. Calo, *supra* note 6, at 33–43 (adding that "courts have struggled with the status of robots, asking how robots can be said to represent, imitate, extend, or absolve people" and that "judges have a problematically narrow conception of what a

between what a robot is and how courts are likely to think of robots will only grow in salience and import over the coming decade."[21]

However, instead of classifying robots and A.I. agents in a dichotomous scheme as either tools or persons, the law should set analogies to appropriate legal categories that turn on the degree to which each robot exemplifies the characteristics of A.I. agents that are critical to law: emergence, embodiment, and social valence.

I first address the question about the appropriate legal category for A.I. robots by recalling the conversation between Calo (identifying essential characteristics of robots) and Balkin (identifying relevant characteristics of robots for the law), in Part II.

I then show in Part III how seeing Calo and Balkin's characteristics as gradients rather than binaries enables the application of a richer set of existing legal categories to these technologies. This, in turn, informs how the law should treat each kind of A.I. (III.A). This exercise traces the continuum between tools and humans, showing where each A.I. can be placed (III.B).

With this framework in mind, Part IV reevaluates the three main characteristics of A.I. and robotics law. I argue that what legal treatment an A.I. agent should have does not depend on embodiment (IV.A), but depends on social valence (IV.B) and on emergence, with one facet of emergence being relevant now and another increasingly relevant ahead (IV.C).

Finally, in Part V, I apply the framework proposed. I first explore examples from science fiction to show the intuitiveness of the framework (V.A). I then explore contemporary issues in doctrinal discussions about A.I., focusing on tort law (V.B), copyright law (V.C), and criminal law (V.D), to show the framework's immediate utility. Building legal arguments from these analogies, rather than on a dichotomous determination of A.I. technologies a tool or a person, leads to more nuanced legal outcomes.

II. DISRUPTIVE CHARACTERISTICS OF A.I. AGENTS

"Robotics blurs the very line between people and instrument."[22] Bearing this in mind, Calo seeks to define robotics' essential qualities in

robot is.").

21. *Id.* at 44.

22. Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 515 (2015).

order to distinguish robotics law from internet law.[23] The three characteristics that he proposes are embodiment, emergence, and social valence. [24]

An entity has embodiment when it can interact with the world physically. Of course, any computer exists as a physical object, so embodiment means something more: the agent must have a physical presence that can act directly on the world according to its guiding algorithm.[25] This idea of "acting directly on the world" refers to programmability, which enables the robot to act physically—as opposed to someone physically using it as an object. Programmability, in turn, leads to the second element: emergence.[26]

Emergence is the unpredictability about how an agent interacts with the environment. It refers neither to purely predictable nor purely random behavior, but to how an entity takes input from the world, processes it, and uses it to engage in actions that produce results in the world.[27] Emergence is essential to define how much an A.I. agent's actions can be foreseen and how liable other people should be for them.

Social valence, perhaps the most open of the three elements, refers to when (real) people treat robots and other A.I. agents as human beings.[28] Robots and A.I. algorithms produce, in the average person, different reactions than do hammers, flash drives, and calculators.

These three characteristics "support a separate academic inquiry from cyberlaw, one that reflects the differences between the two technologies and the human experiences they support."[29] In other words, these characteristics justify and delineate the scope of A.I. law.

The characteristics of A.I. that matter for law are those that affect how people use technology and how people interact with each other.[30] Therefore, we should treat the essential characteristics of A.I. that guide people's interactions with it as the characteristics that determine its treatment under law. Legal categories should be determined not by qualities of the technology that do not affect how people use it but by

---

23. *Id.* at 513 ("Robotics has a different set of essential qualities than the Internet and accordingly will raise distinct legal issues.").

24. *Id.* at 532–49 (explaining each of these characteristics at length).

25. *Id.* at 532–37 (explaining embodiment). *See also* ROLF PFEIFER & CHRISTIAN SCHEIER, UNDERSTANDING INTELLIGENCE 37 (1999).

26. Calo, *supra* note 22, at 534 ("Robots thus combine, arguably for the first time, the generative promiscuity of data with the capacity to do physical harm.").

27. *Id.* at 538–45 (explaining emergence).

28. *Id.* at 545–49 (explaining social valence and arguing that people blame or praise robots, they can feel their privacy is invaded by robots, they might undergo risks to save robots, they might be hurt by losing a robot, and they might be hurt if others mistreat robots).

29. *Id.* at 562.

30. Balkin, *supra* note 17, at 85 (demonstrating how "in most cases, the problem isn't the robots. It's the humans").

how they affect the relationships between people.[31] This means that characteristics that technologists might not consider essential to a given technology may be essential to lawyers working with it. The converse is also true. Embodiment, for example, while perhaps an essential technical characteristic, becomes salient for legal conflicts involving robots because, in the United States, tort law treats economic harm and physical harm differently.[32] Embodiment is therefore a characteristic of robots that becomes salient because of a specific feature of tort law.[33]

In other words, rather than focusing on technologically essential characteristics of a certain technology, this discussion focuses on the legally essential characteristics, which are those central to problems that robots and other A.I. agents pose for law due to how they impact social relations.[34] A.I. robots do not pose problems for law independent of the social relations that they are used for. A.I. law, like any branch of law, ultimately concerns relationships between people.[35]

---

31. Balkin, *supra* note 15, at 45 ("I do not think it is helpful to speak in terms of 'essential qualities' of a new technology that we can then apply to law. On the contrary, we should try not to think about characteristics of technology as if these features were independent of how people use technology in their lives and in their social relations with others."). *See also* Balkin, *supra* note 16, at 96 ("Instead, we have to focus on the social effects of the use of a particular [A.I. agent or another] algorithm, and whether the effects are reasonable and justified from the standpoint of society.").

32. Aikens v. Debow, 541 S.E.2d 576 (W. Va. 2000). *See also* AttorneyFirst L.L.C. v. Ascension Entm't, No. 06-2320, 2007 WL 2733349 (4th Cir. Sept. 20, 2007) (describing *Aikens* as "holding that in the absence of personal or property damage or a contractual relationship, a party alleging negligence resulting only in economic loss must establish a special relationship between the plaintiff and the tortfeasor").

33. Calo, *supra* note 22, at 536–37 ("[C]ourts have invoked the economic loss doctrine to limit liability for data lost pursuant to a computer freezing, even when such data is clearly of great value to the aggrieved party. But the economic loss doctrine does not apply by its terms to physical harm."); Balkin, *supra* note 14, at 49–51 ("[T]he reason that it is salient is that, as Calo explains, from the early days of the internet, judges decided to adapt existing tort law doctrines to limit liability for purely economic losses caused by new digital technologies."). *See also* Michael L. Rustad & Thomas H. Koenig, *The Tort of Negligent Enablement of Cybercrime*, 20 BERKELEY TECH. L. J. 1553, 1578-80 (2005).

34. Balkin, *supra* note 15, at 47 ("The features that are most salient about our technology may depend on how people come to use it. Often people use technology in ways its designers did not foresee or intend."). *See also* Calo, *supra* note 22, at 516 ("[T]he better we understand how a technology will affect society, the better position we are in to integrate the technology gracefully").

35. Balkin, *supra* note 15, at 48–49 ("When we talk about 'technology,' we are really talking about [1] how people interact with new inventions and [2] how people interact with other people using those new inventions."); Balkin, *supra* note 17, at 83–86 (calling the idea that robotics law is concerned with regulating the relationship between people and machines "the homunculus fallacy").

III. A FRAMEWORK FOR COGNIZING A.I. AGENTS: THE CONTINUUM OF ANALOGIES FROM TOOLS TO PERSONS

A. Identifying Analogies

The biggest challenge of emerging A.I. law is finding the appropriate legal category for A.I. agents. As Calo puts it, "[c]ourts have struggled for the proper metaphor to apply to the Internet and will struggle anew with Robotics."[36] Legal reasoning often proceeds by analogy. But which category should we apply here?[37] Are A.I. robots (most similar to) adult humans, children, animals (domesticated or wild), corporations, tools, or something else?

Analogical reasoning is the most conventional method of legal reasoning in the common law.[38] Under analogical reasoning, decision-makers develop low-level reasons to identify two or more cases as being relevantly similar without requiring a full theory of decision.[39] Unless a particular regulation mandates some specific treatment,[40] the analogy

---

36. Calo, *supra* note 22, at 516.

37. Emily Sherwin, *A Defense of Analogical Reasoning in Law*, 66 CHICAGO L. REV. 1179, 1179 (1999) ("The analogical method, as commonly practiced, works something like this: confronted with an unsettled question, the judge surveys past decisions, identifies ways in which these decisions are similar to or different from each other and the question before her, and develops a principle that captures the similarities and differences she considers important"). *See also* Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923, 938-63 (1996) (comparing analogical reasoning with Rawlsian reflective equilibrium).

38. Cass Sunstein, *On Analogical Reasoning*, 106 HARV. L. REV. 741, 741 (1993) ("Reasoning by analogy is the most familiar form of legal reasoning"); *id.* at 790 ("analogical reasoning is the conventional method of the lawyer"); Brewer, *supra* note 37, at 925 ("reasoning by analogy is one of the most frequently used techniques of legal argument"). *See also* LLOYD WEINREB, LEGAL REASON: THE USE OF ANALOGY IN LEGAL ARGUMENT (2016) (arguing that analogy is the essence of legal reasoning).

39. CASS SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT (2018) at 61-69 (adding that it allows a society to agree on the legal treatment of a phenomenon while there is moral controversy in the background).

40. At a policy level, analogical reasoning allows for low level agreement upon disagreement about comprehensive theories. CASS SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT (2018) at 61-69. This is relevant for regulating A.I. agents, at least until there is agreement on how a new and different legal treatment should look like. See also Sunstein, *supra* note 38, at 791 (arguing that analogical reasoning "does not require people to develop full theories to account for their convictions; it promotes moral evolution over time; it fits uniquely well with a system based on principles of stare decisis; and it allows people who diverge on abstract principles to converge on particular outcomes."); Sherwin, *supra* note 37, at 1186-94 (arguing that there are several epistemic and institutional benefits to analogical reasoning: it produces data for decision-making, it represents a collaborative efforts of different legal agents over time, it tends to correct biases, and it helps the law develop at a gradual pace). But see Richard Posner, *Reasoning by Analogy*, 91 CORNELL L. REV. 761, 765 (2006) (arguing that "reasoning by analogy tends to obscure the policy

chosen to describe A.I. robots will determine how they are treated by the law.[41]

Choosing an analogy for A.I. robots from among the categories that exist in our legal system can, moreover, help the law cope with yet unforeseen challenges in technology. The law can apply settled doctrinal and regulatory debates from the analogized category to the new technology, rather than rehash them with each technology anew.[42]

More helpful than choosing any one analogy, however, is developing a framework for selecting the right analogy in any given scenario. Each A.I. technology presents different levels of embodiment, emergence, and social valence. The real question is, therefore, not finding the proper analog for A.I. in general but finding the most appropriate analog for each A.I. agent. A framework can help in this endeavor.

The sense-think-act paradigm, which builds on the idea that "robots are mechanical objects that take the world in, process what they sense, and in turn act upon the world,"[43] describes characteristics that exist in each technology. These characteristics exist in varied degrees. We can think of full emergence as autonomy, full embodiment as a humanoid form, and full social valence as an emotional being who produces empathy in others by performing a specific social role. But these characteristics do not exist in the world in binary states. To find the adequate legal category for each kind of A.I., we must stop seeing A.I.'s characteristics as either off or on, and instead see them as inhering in A.I. in amounts along a continuum.

Moreover, while the three traits are related, the appearance of any one within a technology does not always trigger the appearance of any other. A robot with a body (embodiment) and independence (emergence) may be more easily treated as a humanoid (social valence). But an A.I. agent could alternatively be both incorporeal (not embodied) and autonomous (emergent), or be neither incorporeal nor humanoid, as is a Roomba.[44] The A.I. that present the most puzzling doctrinal questions, such as self-driving cars, Siri, writing algorithms, and drones, present some but not all of these disruptive characteristics. For this reason, we should evaluate the characteristics independently.

---

grounds that determine the outcome of a case, because it directs the reader's attention to the cases that are being compared with each other rather than to the policy considerations that connect or separate the cases").

41. *See, e.g.*, Chopra and White, *supra* note 9.

42. The risk introduced is tying the legal outcome to the analog too tightly. However, this is only an undesirable risk if, in expectation, the undesirable results produced by being tied too closely to the background story outnumber the undesirable results of having no background at all.

43. Calo, *supra* note 22, at 529 (demonstrating that there is consensus over the sense-think-act paradigm).

44. *See generally Home Robots: Robot Vacuums*, IROBOT (2017), http://www.irobot.com/For-the-Home/Vacuuming/Roomba.aspx [https://perma.cc/K5WY-5L27h].

The following table orders the possible legal categories mentioned earlier in this Section along a spectrum that ranges from tool to person. It also evaluates how much of each of the salient characteristics of A.I. is present in each of the analogs that conforms to a legal category. The legal categories are ordered to reflect the varying implications for issues like supervision and control, foreseeability, and the reactions we expect people to have towards the technology, which are the main themes along which the disruptive characteristics present doctrinal challenges.[45]

EMBODIMENT	EMERGENCE	SOCIAL VALENCE	ANALOG
Very High	Very High	Very High	Adult Human
Very High	High	Very High	Child
High	Low	High	Domesticated Animal
High	Very High	Low	Wild Animal
Very Low	Low	Very Low	Corporation
Low	Very low	Very Low	Tool

Table 1: A.I. Robotics' Analogs

The grid set out in Table 1 fits different technologies into existing categories of liability. It allows one to examine an A.I. technology, existing or still unforeseen, by evaluating the extent to which each characteristic is present and then identifying the most adequate analog among law's existing categories.[46]

These analogies are not simply comparisons *per se* between A.I. robots and the object of comparison but comparisons with regards to the characteristics that are relevant for liability. If it looks like a dog, walks like a dog, and barks like a dog, it might still not be (like) a dog for normative purposes. The relevant analogies are comparisons in the eyes of the law between the most significant legal problems generated by the technology in question and those generated by the analog with which the technology shares salient characteristics.

Moreover, as with any analogy, its purpose is not to treat the analogized object exactly as its analog, but to bring an underlying legal regime on which to build regulations for its specific differences. Because anal-

45. *See infra* Part IV.

46. Note that, by categorizing analogs in these three dimensions, I necessarily refer to averages. In certain situations, wild animals can be highly predictable, children highly unpredictable, and adults may be either very predictable or unpredictable, while corporations can be as predictable as their corporate officers. It is only on average that we find children's behavior easier to predict than that of their parents, and our neighbors' behavior easier to predict than that of park squirrels.

ogies, in other words, aim to compare one object with another with respect to some principle, analogs are not to be taken as an indivisible set of normative consequences.[47] Embodiment, emergence, and social valence are not ideas that the law is seeking *per se*, but they reflect different problems that A.I. robots can pose for the law.[48]

The adequate analog, then, will oftentimes depend on the regulatory question asked. Some simple technologies might escape this specificity. If a technology has very low embodiment, very low emergence, and very low social valence, the law can generally treat it as a tool, independent of the specific regulations that it might require. Consider, for example, 3D printing. Issuing specific regulations, such as a prohibition on printing weapons, is necessary. But the problems 3D printers trigger in different areas of the law can be dealt with adequately by treating it as a tool.

Most technologies under A.I. law, however, will not escape some analysis before the use of an analogy. Depending on the doctrinal question at hand, we might examine the foreseeability of an A.I. agent's actions or its ability to deceive people. The specific legal question will determine the most adequate analogy. Some examples of these regulatory questions are dealt with below.[49] Before addressing them, though, we should evaluate how these characteristics interrelate in the continuum from tool to person.

B. Placing A.I. in a Continuum from Tool to Person

These characteristics work under archetypes.[50] Near-term technologies do not resemble the A.I. robot archetype that the European Parliament had in mind when issuing its report: an autonomous, humanoid, and empathic robot. Nor will these technologies present such straightforward doctrinal and regulatory challenges as archetypical A.I. robots will present in the future. As Calo writes, "we can see already how robots begin to blur the line between people and instrument, and how

47. It would be unhelpful to break down the characteristics into their relevant elements just to regroup them differently.

48. Robotic analogies, like any others, face two limitations. First, there are sometimes multiple analogies for the same object and people might disagree on which is the most appropriate story. Second, each analogy will inevitably be limited in its power to regulate a new technology because it may lead one to see the problem only in one way and cause one to misfire in the possible regulatory scheme—one could use the story as a crutch and misunderstand the underlying reality. *See* Balkin, *supra* note 17.

49. *See infra* Part V.

50. In some way, this logic of archetypes is similar to that of property law in the bundle of rights represented by sticks that is often used as a metaphor for property rights. Like ownership is a property archetype, but there are many legal situations with only some aspects of it, such is the robot archetype, but most (and so far all) A.I. agents have only some aspects of it.

faulty assumptions about robots lead jurists to questionable or contradictory results."[51] For that reason, we must go beyond the dichotomy between the archetypical A.I. robot and the common household appliance.

Accordingly, based on the mapping of technologies to legal categories I proposed above, we can instead construct a continuum showing the full range of expression for each characteristic. We can then place A.I. agents on this continuum based on their central characteristics, revealing their status from tool to person, as Calo suggested should be done.[52] This sorting generates tailored analogies for A.I. that help show how to treat each technology in the midst of doctrinal and regulatory uncertainty. Because the analogies presented in the last section depict the legal categories that would exist in this continuum, we can now construct it.

The following figures illustrate the proposed gradients.

Tool – Corporation/Dom. Animal – Child – Wild Animal/Adult

Figure 1: Emergence

Tool/Corporation – Wild Animal – Dom. Animal – Child/Adult

Figure 2: Social Valence

This proposal helps determine the type of treatment that should be granted to a technology, particularly when technologies do not fit into any of our existing legal categories. The approach also focuses on the legally relevant questions behind these analogies, addressing the concerns presented above about analogical reasoning.[53] After evaluating the three characteristics mentioned, the following sections present examples for technologies closer to the "human" side of the spectrum[54] and for others closer to the "tool" side, the latter posing more present doctrinal questions.[55]

As Balkin has said, "[w]e might identify many different features of a technology as its key or essential characteristics, but the real issue is always *why* we care about them. How we define the central features of a technology depends on what our definition is *for*."[56] When we evaluate

51. Calo, *supra* note 6, at 5.

52. *See* Calo, *supra* note 22, at 530–32 (arguing that "each of these characteristics of sensing, processing, and acting exists on a spectrum").

53. *See supra* note 48.

54. *See infra* Part V, Section A.

55. *See infra* Part V, Sections B, C, and D.

56. Balkin, *supra* note 15 at 51 (emphasis in the original) (adding that "we might identify different features of a technology as its key or essential characteristics, but the real issue is always why we care about them. How we define the central features

which characteristics are the most disruptive for the law, rather than evaluate those which are essential to a technology, we can question to what extent these can define where a technology enters in the spectrum from tool to person.

IV. Completing the Framework: Characteristics that Matter for Analogizing A.I. Agents

A. Embodiment: What is Essential is Invisible to the Eye

Embodiment is, actually, programmability.[57] The mere fact of having a physical body contributes to a robot's emergence and social valence but it does not by itself present robotics' most challenging legal questions: "what is important is not that robots have a physical form—so do toasters—but that their interactivity creates particular social cues in human beings."[58] It could be concerning how a robot with a realistically humanoid plastic body could deceive people into thinking she is more human than she actually is, but this is an issue of social valence affected by the robot's body, not an issue of embodiment.

For this reason, while embodiment can present interesting doctrinal questions in specific cases, it does not reveal where to place a technology in the proposed continuum in general,[59] nor does it help determine the most appropriate legal analog for a technology.[60] Focusing instead on emergence and social valence allows us to see the legal issues of the technology.

This goes back to the consideration made in the introduction in which, following Balkin, I said we should treat A.I. robots and disembodied A.I. algorithms indistinctly.[61] Irrespective of whether we consider embodiment determinant of legal analogies, defining embodiment as programmability as Calo does helps us avoid unhelpful distinctions, such as the one between robots and self-learning algorithms: "[s]elf-learning algorithms can raise or lower temperatures in a house, turn on appliances, lock or unlock gates, and notify security services. Algorithms can buy and sell securities; they can create holographic projections that look and act like people; they can threaten, entertain, copy,

---

of a technology depends on what our definition is for").

57. *See id.* at 49; Calo, *supra* note 22, at 533–35 (explaining embodiment and noting further at 536 that "embodying data also disrupts a more basic distinction between informing and acting").

58. Balkin, *supra* note 15, at 49–50 (adding that "A robot's ability to cause physical injury is not really an 'essential' characteristic of robotic technology. It is a particularly *salient* feature of robotics for lawyers.").

59. *See supra* Part III, Section B.

60. *See supra* Part III, Section A.

61. *See supra* note 17.

defame, defraud, want, console or seduce."[62] These consequences are as physical as it gets, albeit produced by an agent without embodiment.[63]

Embodiment is said to be centrally relevant because tort law treats economic harm differently than physical harm.[64] Yet as we just saw, mere operating systems can cause physical harms. Moreover, treating physical and economic harm differently is not essential for tort law, as most civil law jurisdictions do not do so. As economic harm becomes increasingly relevant—especially if A.I. agents continue to create situations in which harm is not physical but some form of redress is appropriate—this distinction could change in the future.[65] It could also become less consequential, as some comparativists argue is already taking place.[66]

This brings us back to how robotics law is concerned with relationships between people.[67] As Balkin argues, one of the central legal problems that robots (defined by the characteristics mentioned) present is how to distribute the rights and duties arising when robots behave among people.[68] The second central challenge is what Balkin calls the substitution effect: the extent to which people are willing to substitute robots for animals or humans in certain contexts.[69] The first issue tracks with emergence and the second with social valence.

---

62. Balkin, *supra* note 15, at 50–51.

63. The same can be said about making sense of the so-called "chatter robots," which in the most literal sense have no embodiment at all. Problems such as algorithmic speech and automated patent infringement "further undermine the assumption that a capacity to cause physical injury is an essential characteristic of the technologies we are interested in." *Id.* at 52.

64. Calo, *supra* note 22, at 533 ("While robotics also relies on data, the mainstreaming of robots signals a shift back to the physical."); Balkin, *supra* note 15, at 51 ("The problem of 'emergence' . . . cross-cuts with the problem of physical harm.").

65. For a similar situation, imagine Congress passed a tax on metal. Being made of metal would not become a central characteristic of robots. It would be one salient characteristic of some robots for tax purposes, but it would not determine how we think about how the law should treat that technology.

66. *See* Gerhard Wagner, *Comparative Tort Law*, *in* THE OXFORD HANDBOOK OF COMPARATIVE LAW 1004, at 1020 (Mathias Reimann and Reinhard Zimmerman eds., 2006) (admitting that there are substantive differences between legal traditions and observing that "[i]t is, however, safe to contend that there is much less diversity in outcomes than one would expect when looking at the several systems of tort law or delict in the abstract and in isolation"); *id.* at 1018 (observing that one of the reasons of this convergence is that "the most important cases have been taken care of by regulatory law, which brought with it statutory heads of liability"); *id.* at 1020 (noting that, when not covered by statute, pure economic loss is often covered by tort law under a different category (mostly when there is quasi-contract or negligent misrepresentation), so the problem becomes solely the coverage of those cases that fall outside those categories).

67. Balkin, *supra* note 15, at 48–49.

68. *Id.* at 46.

69. *Id*. Note that this is different than the economic concept with the same name.

Embodiment is a relevant feature of many robots—as it may enable them to fly, drive fast, be bullet-proof, or bake *quiches*—but it does not define the rights and obligations that these robots should have.

B. Social Valence as Being Sentient and Felt

Because law concerns relationships between people, we must examine how A.I. agents provoke psychological, social, and economic changes in human relationships, as these will affect their treatment under law.

The social valence of current A.I. changes how people interact with it. On the one hand, for example, soldiers in war situations sometimes take risks to save A.I. robots with high social valence.[70] On the other, soldiers do not think twice before shooting down a drone.[71] This dramatic difference in human response turns on the two robots' different levels of social valence, not on their abilities.

Much of social valence is about *seeming* rather than about *being* emergent. Therefore, confusion between them is frequent: high social valence might lead us to miscategorize a robot in terms of its emergence—a separate quality that will be analyzed in the next Section. Social valence, in other words, is not about what actions are foreseeable for the technology but about what reactions are foreseeable from other people with respect to the technology. Because an entity's social valence affects victim vulnerability, for instance, it is relevant for people's ability to injure others through the torts of misrepresentation, fraud, and seduction.

Depending on their social valence, A.I. agents could also deceive people into buying more products than they want or should buy. The key problem here is not "too many products" but "deception." A risk that technology is close to introducing is that a humanoid robot with a developed A.I. could pretend to care for our interest while caring for the commercial interests of other people. The social role that the A.I. agent is perceived to be performing may be different from the social role that it is actually performing.[72]

---

70. *See generally* Ryan Calo, *The Drone as Privacy Catalyst*, 64 STAN. L. REV. ONLINE 29 (2011).

71. Calo, *supra* note 6, at 23.

72. This idea relates to fiduciary duties. In particular, it relates to how the social role of professionals, which implies people putting trust on their advice, places limits on what behavior is expected and acceptable. *See* Jack Balkin, *Information Fiduciaries and the First Amendment*, 49 UC DAVIS L. REV. 1183 (2016) (introducing the concept of information fiduciaries as those agents in the digital age who have become so central in human interactions that must be considered to have fiduciary duties in their roles); Kiel Brennan-Marquez, *Fourth Amendment Fiduciaries*, 84 FORDHAM L. REV. 611 (2015) (calling for an expansion of the concept of information fiduciaries to re-think how the law, and particularly the Fourth Amendment, deal with shared information); Claudia E. Haupt, *Professional Speech*, 125 YALE L.J. 1238,

Because human beings can deceive people too, one could think that this risk is not new. However, while people have ways of guarding against human deception, they do not yet have ways of guarding against being deceived by their own devices. If a robot appears to be relatable or at least loyal (i.e., serving a person's own interests) but, in fact, it is programmed to serve a third party's interests (or its own), the robot can be exploitative.

Social valence is also relevant to plaintiffs' expectations of harm. If my car is approaching another one, I can expect the driver to swerve to avoid colliding with me, but I should not expect this if my car is approaching a rock. So, when approaching an object, my behavior depends on what I believe that object is.

As A.I. technology advances, using social valence as a proxy for measuring a robots' deserved empathy will not be sufficient. At some point, the law may need to give A.I. agents rights irrespective of how people choose to treat those A.I. agents in individual cases. In other words, one cannot normatively answer how society should treat a certain robot simply by empirically investigating how people do (or how society does) treat that robot; this would be a naturalistic fallacy.[73] While current technology does not yet present this question, technology eventually will.

How people react to a technology is relevant to determining the protections that should be established around it. But, if used to determine the rights that the technology should have, it misses the mark. For this, we will need to measure something different. What social valence will turn to represent then—what we will need to ask—is whether the A.I. system is sentient, or even conscious.[74]

This question will determine whether a robot should be treated as a moral patient.[75] Most of us treat animals differently than we do plants

---

1271 (2015) (providing a comprehensive account of the doctrinal and theoretical bases of professional speech); Claudia E. Haupt, *Unprofessional Advice*, 19 U. PA. J. CONST. L. 671 (2017) (providing a theoretical foundation to distinguish professional from unprofessional advice).

73. *See generally* Thomas Hurka, *Moore's Moral Philosophy*, *in* STAN. ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2015), (July 18, 2015), https://plato.stanford.edu/archives/fall2015/entries/moore-moral [https://perma.cc/QN24-BR5W] (explaining the naturalistic fallacy)); Michael Ridge, *Moral Non-Naturalism*, *in* STAN. ENCYCLOPEDIA OF PHILOSOPHY, *supra id.* https://plato.stanford.edu/archives/fall2014/entries/moral-non-naturalism [https://perma.cc/VAF5-FZLX].

74. See Carissa Véliz, *The Challenge of Determining Whether an A.I. Is Sentient*, SLATE, (Apr. 14, 2016, 8:30 AM), http://www.slate.com/articles/technology/future_tense/2016/04/the_challenge_of_determining_whether_an_a_i_is_sentient.html [https://perma.cc/D9E7-TC79].

75. Initially, whether a robot is sentient will depend on the programmer. We feel pleasure and pain because we have been "programmed" to do so over millions of years as we evolved. This allows us to judge value—we like hugs but not insults, just as we like ice cream but not sharp objects on our tongue. A robot, divorced from

or objects because of their ability to feel emotions such as pain (though most of us still do not care enough to stop eating them).[76] We also treat different animals differently depending on how sentient we believe them to be: we tend to respect insects less than mammals. If conscious, the robot agent will deserve a still higher moral status: in most cultures, people are not comfortable killing a being conscious of its own existence.[77]

In sum, social valence as sentience or consciousness is relevant for assigning rights to A.I. robots, while social valence as socially induced empathy is relevant for assigning people rights against A.I. robots (even in the absence of the robots' sentience or consciousness). Since we tend to empathize with what is sentient, and to a much larger degree with what is conscious, the two types of social valence appear along the same scale.

C. Two Dimensions of Emergence: Agency and Unpredictability

Emergence presents a similar dichotomy to social valence, but with different scales. There are two independent lines of emergence: unpredictability and agency. These relate to the law differently. The law assesses foreseeability (and control) to determine direct and vicarious liability for an entity's actions. The law also assesses agency to determine whether it makes sense to attribute liability to someone (or something) who caused the harm.

Currently, A.I. agents lack agency; it would be unhelpful to make a robot civilly or criminally liable for its actions. A robot's emergence, on the other hand, is crucial to determining whether other people could have foreseen its actions. Foreseeability will, in turn, determine the human liability for the robot's actions;[78] it is crucial to determining

---

the physical perception of pain and pleasure, may find it difficult to make emotional judgments of pain and pleasure. For her to do so, we would have to program her with packets of information that make, for example, the input information "knife in chest" produce something more than the output information "human cease to function." Only after we understand how to program sentience into her, and succeed in doing so, will her self-learning mechanism play a role in making moral judgments. If the robot becomes conscious, this might change. Consciousness could create a notion of good independent from her programming that relates to her self-preservation.

76. PETER SINGER, ANIMAL LIBERATION (2d ed. 1990) (arguing that any entity that can suffer deserves moral consideration).

77. Perhaps we will eventually recognize the rights of robots not when they begin to remind us of ourselves but when they understand and demand rights. This has not been how we have come to recognize rights for other beings on our planet but, unlike those beings, robots might be able to develop their ability to understand and demand rights within a lifespan.

78. *See, e.g.*, *Devlin v. Smith* 89 N. Y. 470 (1882) (extending the scope of foreseeable victims of a contractor's negligence to include not only the purchaser but

whether the programmer or the owner are liable for the harms caused by the robot.

An example of how emergence relates to unpredictability absent agency, recall *Tay*, an A.I. chatting bot that Microsoft released in April 2017.[79] *Tay* had to be shut down after sixteen hours because it became racist and sexist, denied the holocaust, and supported Hitler.[80] Had *Tay's* mainframe been in Germany instead of the United States, what she did would have been considered a criminal offense—as Germany differs strongly in its approach to this type of speech.[81] The next A.I. agent could be based in a country where this type of speech is a crime, or could commit a different crime in the United States.[82] If this occurs, we would be unsure how to respond legally. Because *Tay's* behavior would have been difficult to anticipate, we would know to treat *Tay* differently than a common tool used by a person to produce the same result directly. *Tay's* programmers (or Microsoft) would perhaps be considered in some way responsible. But, in recognizing this lack of foreseeability, the situation should be evaluated differently than if they had typed and sent those messages themselves.

This disaggregates two types of emergence. The first is emergence from the perspective of the human viewer, which is the one on which Calo and Balkin focus: unpredictability.[83] The second is emergence

---

the purchaser's employees); *Heaven v. Pender* 11 Q.B.D. 503 (App. Ct. 1883) (establishing early on in English law that a duty to take reasonable care exists whenever a person of ordinary sense would recognize that the conduct would endanger a person or property). Foreseeability was key to establishing liability for products, including for a milestone case in 20th century tort law: *MacPherson v. Buick Motor Co.* 217 N.Y. 382 (1916). In the same way, algorithmic liability, which is likely to be the milestone of tort law in our time, will depend on foreseeability.

79. Sarah Perez, *Microsoft's new AI-powered bot Tay answers your tweets and chats on GroupMe and Kik*, TECHCRUNCH (Mar. 24, 2016), http://social.techcrunch.com/2016/03/23/microsofts-new-ai-powered-bot-tay-answers-your-tweets-and-chats-on-groupme-and-kik [https://perma.cc/M7PX-HFVZ].

80. [A]dezero, *Microsoft Creates AI Bot - Internet Immediately Turns It Racist*, SOCIAL HAX (Mar. 31, 2016), https://socialhax.com/2016/03/24/microsoft-creates-ai-bot-internet-immediately-turns-racist [https://perma.cc/8ZGD-JR2P]. Sarah Perez, *Microsoft Silences its New A.I. Bot Tay, After Twitter Users Teach It Racism*, TECHCRUNCH (Mar. 24, 2016), https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism [https://perma.cc/E9K5-T2QB].

81. Strafgesetzbuch §130. This would also be the case in Israel (Denial of Holocaust Law, 5746-1986, art. 2) and in other European countries such as Austria (Verbotsgesetz 1947 §3h), Belgium (Act of 23 March 1995 on punishing the denial, minimisation justification or approval of the genocide perpetrated by the German National Socialist Regime during the Second World War, art. 1), and France (Law 90-615 to repress acts of racism, anti-Semitism and xenophobia (Gayssot Act), art. 9).

82. Especially if, embodied or not, she is a robot that can do physical damage.

83. *See* Calo, *supra* note 22, at 538–45 (explaining emergence as unpredictabil-

from the internal perspective of the robot itself: agency.[84] The first type of emergence will determine the extent to which an A.I. agent's behavior is foreseeable to people who have free will and who live under the rule of law—which is fundamental to determining liability under tort law. The second type of emergence can be viewed as higher order intentionality:[85] it determines the extent to which an A.I. agent could respond to legal incentives directly. Agency, as such, determines the extent to which A.I. could disrupt tort law in the future.

Foreseeability operates as a limit to responsibility for one's own actions: one is rarely responsible for what one cannot foresee.[86] As for control, it is relevant to determine vicarious liability. Foreseeability and control determine when people become responsible for harms caused by their agents, children, animals, and dangerous objects. To determine someone's liability for her use of a simple, predictable robot, we should examine foreseeability. However, to determine someone's liability for the actions of an advanced A.I. that has agency, the level of control over the A.I. will be most relevant. As emergence advances, such torts will affect the legal incentives faced by robot and A.I. designers only to the extent that they can exert control. Control, therefore, will determine the usefulness of legal incentives.

Because A.I. agents with agency are so far inexistent, A.I. law remains a system for allocating rights and duties among people based on the foreseeability of robot actions alone. Agency may remain absent in A.I. for a long time, but probably not forever.[87] The law needs to address emergence as foreseeability (or unpredictability) now, and will need to address emergence as agency at some point in the future.

With this distinction in mind, Table 1, shown above, can be modified to portray both types of emergence.

---

ity and noting that "the larger question will turn out to be who is liable for the infringement that an emergent system occasions"); Balkin, *supra* note 15, at 51–55 ("[F]rom the standpoint of law—as opposed to the standpoint of engineering—the problem posed by emergence is the problem of assigning responsibility for the unpredictable behavior of robots and AI systems.").

84. *See* Daniel Dennett, *When HAL kills, who's to blame?*, *in* RETHINKING RESPONSIBILITY IN SCIENCE AND TECHNOLOGY (FIORELLA BATTAGLIA, NIKIL MUKERJI, AND JULIAN NIDA-RÜMELIN EDS.) 203 (2014); John Sullis, *When is a Robot a Moral Agent? in* MACHINE ETHICS 151–61 (Michael Anderson & Susan Leigh, eds. 2011) (arguing that advanced A.I. systems can be moral agents).

85. *See* Dennett, *supra* note 84 (explaining higher order intentionality and arguing that it is a precondition for moral responsibility).

86. Foreseeability as a limit to people's liability is a central concept of tort law, with its most prominent exposition in Judge Cardozo's opinion in *Palsgraf*. Palsgraf v. Long Island Railroad 162 N.E. 99, 101 (N.Y. 1928) (limiting the scope of responsibility for negligent acts to the set of foreseeable results).

87. *See* Hallevy, *supra* note 10 (generally arguing across multiple works that responsibility for A.I. agents should already be present).

EMBODIMENT	EMERGENCE (AGENCY)	EMERGENCE (UNPREDICTABILITY)	SOCIAL VALENCE	ANALOG
Very High	Very High	Very High	Very High	Adult Human
Very High	High	High	Very High	Child
High	Low	Low	High	Domesticated Animal
High	Low	Very High	Low	Wild Animal
Very Low	Very Low	Low	Very Low	Corporation
Low	Very Low	Very Low	Very Low	Tool

Table 2: A.I. Analogs, Disaggregating Emergence

As Table 2 shows, the first type of emergence (unpredictability) is necessary to allocate responsibility for robots' actions among people. Absent vicarious liability, its relevance is thus conditional to the absence of the second type of emergence (agency), which would trigger allocating responsibility to the robot itself. The background analogies are chiefly determined by emergence because it is the characteristic that has to do with the issues of control and foreseeability, which are at the core of the questions about personal responsibility that the distinction between the different analogies seeks to answer.

In sum, this proposal shows some symmetry in how to treat A.I. agents. Emergence should be used to determine whether to treat a robot as a moral agent, and social valence should be used to determine whether to treat her as a moral patient. An A.I. agent that is both a moral agent and a moral patient would be on the "human" end of the scale portrayed above.[88] Currently, no A.I. agent has these degrees of autonomy and consciousness. To imagine a world where new beings do have such moral agency and patiency, and to draw lessons from it, we must turn to science fiction.[89]

V. APPLYING THE FRAMEWORK TO DOCTRINAL A.I. PROBLEMS

A. *Future Cases: Science Fiction and the Intuitiveness of the Framework*

We can now evaluate how the proposed framework comports with our intuition when applied to cases from science fiction, and how it helps us answer more difficult questions when applied to contemporary challenges. Note that, in each area of potential application, the purpose

88. *See supra* Part III, Section B.

89. *See* Balkin, *supra* note 17 (using a story by Isaac Asimov, Karel Capek's play RUR, and the legend of the Golem of Prague to explain the challenges of algorithms and the humans behind them). *See also infra* Part V, Section A.

of the framework is to provide a consistent method for drawing analogies between new technologies and existing subjects that the law already understands. Advocates and judges might propose more specific analogies for particular legal problems. And, of course, policymakers and regulators might choose to add new law.

Since the types of A.I. that (will) have rights and obligations are emergent technologies that do not yet exist—and will not exist for some time—the best examples of A.I. and of society's potential reactions to them come from works of fiction.[90] We cannot look even to proposed laws for insight because, as the European Parliament's proposal illustrated, "jurists on the whole possess poor, increasingly outdated views about robots and hence will not be well positioned to address the novel challenges they continue to pose."[91]

To help anticipate the legal challenges that more advanced technologies will present, we can evaluate science fiction works in which A.I. agents are integrated in society, examine how we believe they should be treated, and compare whether this matches the treatment the proposed framework would prescribe. Similar exercises have been conducted using works of literature;[92] here, I use movies. I evaluate two movies that have an A.I. agent as a main character: *Her*[93] and *Transformers*.[94]

In *Her*, the protagonist, Theodore, falls in love with Samantha, his computer operating system. Samantha has no body, but their relationship develops nonetheless as they both look for ways to overcome this impediment.[95] Samantha is independent, conscious, reflective, and sentient. We could define Samantha's characteristics as: very low embodiment, very high emergence, and very high social valence—under the meanings already described.

This example illustrates why embodiment does not help appropriately categorize an A.I. agent. Under the proposed rubric, we can assign Samantha to a category only according to her emergence and social valence, and so analogize her to an adult human. This is how Samantha is

---

90. Balkin, *supra* note 15, at 60 ("One reason why we may have strong notions of what robotics can do and the problems it can cause is that we have a rich literature of literary examples . . . ."); *see generally* Balkin, *supra* note 17 (using science fiction to explain the central characteristics and legal problems of robotics).

91. Calo, *supra* note 6, at 5 (arguing that American law still has enough experience with robots to predict the challenges that they will pose).

92. *See* Balkin, *supra* note 17 (using a story by Isaac Asimov, Karel Capek's play RUR, and the legend of the Golem of Prague to explain the challenges of algorithms and the humans behind them).

93. HER (Annapurna Pictures 2013).

94. TRANSFORMERS (Paramount Pictures 2007); *see also* TRANSFORMERS: REVENGE OF THE FALLEN (Paramount Pictures 2009); TRANSFORMERS: DARK OF THE MOON (Paramount Pictures 2011); TRANSFORMERS: AGE OF EXTINCTION (Paramount Pictures 2014).

95. For example, at one point in the movie Samantha suggests using a surrogate for them to be intimate.

treated during the movie and how the audience is invited to perceive her.[96] In the movie, Samantha is both a moral agent and a moral patient. In her relationship with Theodore, she is entitled to be treated with the same respect and consideration that other humans do and she is considered responsible for her reckless decisions, as Theodore getting upset by them shows. If Samantha existed, others would probably also deem her responsible for her actions.[97]

The autobots in *Transformers* present a similar story, even though they are embodied. In the movie, the human Sam Witwicky develops friendships with the autobots Bumblebee and Optimus Prime. Both autobots can make decisions (in fact, they often tell Sam what to do) and are fully sentient and conscious. Under the proposed framework, the autobots have very high embodiment, very high emergence (as agency), high social-valence-as-perception, and very high social-valence-as-sentience.[98]

Thus, Autobots, like Samantha, also trigger the analog of an adult. And this matches how the audience receives them: when minor characters in the movie treat autobots as cars (tools), not as humans, something feels wrong. Autobots, after all, can be main characters precisely because they feel human to the audience. If they existed, most of us would probably believe that they should be treated as such.

Other examples of how these characteristics determine the way in which people see an A.I. robot are Wall-E and Chappie in the eponymous movies,[99] whom people would probably treat as children, and who are treated as such by other characters in the films.[100] Both could be described as having very high embodiment, high emergence, and very high social valence, which would trigger that analogy in the framework proposed here. Other examples are Sentinels in *The Matrix*[101] and *X-Men: Days of the Future Past*,[102] and HAL in *2001: A Space Odyssey.*[103] Sentinels in these movies have very high emergence but low social valence and are seen as wild (and dangerous) animals. HAL, who much like

96. Of course, what is pertinent to the rights that we would find intuitive be subscribed to a character is not how the character was, as a matter of fact, treated during the movie. After all, a work of fiction could portray a world that is unjust or disagreeable. What is pertinent is how the work of fiction's portrayed reality triggers the audience's moral intuitions.

97. *See* Hallevy, *supra* note 10 (suggesting, across multiple works and with different examples, that advanced A.I. agents could be criminally liable for their offenses).

98. They are not treated as humans by all characters in the movie, which defines social valence in the traditional sense, and they are fully sentient and conscious, which defines social valence in the proposed sense.

99. WALL-E (Pixar Animation Studios 2008); CHAPPIE (Sony Pictures 2015).

100. In *Chappie*, one of the characters even says "he's a child!" CHAPPIE (Sony Pictures 2015).

101. THE MATRIX (Warner Bros. 1999).

102. X-MEN: DAYS OF THE FUTURE PAST (Twentieth Century Fox 2014).

103. 2001: A SPACE ODYSSEY (Metro-Goldwyn-Mayer 1968).

Samantha in *Her*, has very low embodiment, very high emergence, and high social valence, is portrayed as a disembodied adult.[104]

These examples illustrate that the framework above comports with our social norms and moral intuitions. We do not view these technologies in a binary, as either tool or full human adult. Each technology triggers in us a different analogy that determines how we see it and how we believe it should be treated. Not all A.I. robots are equal, and A.I. law should not treat them all the same.[105]

Having addressed these "extreme" fictional cases, we can turn our attention to more subtle ones.

B. Tort Law

The framework can be applied to doctrinal questions that A.I. poses for different areas of the law. As an illustration, in this section we turn to tort law, chiefly concerned with the question of liability. How and when the law should make people liable for robots' actions is a pressing question today, even if whether robots should have rights is not yet.[106]

Self-driving cars can show how the framework helps answer questions about new technologies using existing law.[107] Autonomous cars

---

104. *See* Dennett, *supra* note 84, at 203-14, 354-57 (analyzing HAL's moral responsibility in more detail). Also note that Dave, in particular, would disagree with the idea that disembodied HAL is unable to cause physical harm.

105. Some robots really are more equal than others.

106. *Id.* at 10 (explaining that robots present legal problems for humans and human relationships: "[t]he laws of robotics that we need in our age are laws that control and direct human beings who create, design and employ robots"). *See also id.* at 7 (demonstrating that robots present legal problems for relationships between humans); *id.* at 10 ("[W]hat we really need are laws of robotics designers and operators. The laws of robotics that we need in our [age] are laws that control and direct human being who create, design, and employ robots.").

107. Sven A. Beiker, *Legal Aspects of Autonomous Driving*, 52 STA. CLARA L. REV. 1145 (2012); Jack Boeglin, *The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation*, 17 YALE J.L. TECH. 171 (2015); Jean-François Bonnefon, Azim Shariff & Iyad Rahwan, *The Social Dilemma of Autonomous Vehicles*, 352 SCIENCE 1573 (2016); Frank Douma & Sarah Aue Palodichuk, *Criminal Liability Issues Created by Autonomous Vehicles*, 52 STA. CLARA L. REV. 1157 (2012); Sophia H. Duffy & Jamie Patrick Hopkins, *Sit, Stay, Drive: The Future of Autonomous Car Liability*, 16 SMU SCI. TECH. L. REV. 453 (2013); Jerone Dunbar & Juan E. Gilbert, *The Human Element in Autonomous Vehicles*, *in* ENGINEERING PSYCHOL. AND COGNITIVE ERGONOMICS: COGNITION AND DESIGN 339–62 (2017); Andrew P. Garza, *Look Ma, No Hands: Wrinkles and Wrecks in the Age of Autonomous Vehicles*, 46 N. ENGL. L. REV. 581 (2011); Mark Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, CAL. L. REV. (forthcoming); Dorothy J. Glancy, *Privacy in Autonomous Vehicles*, 52 STA. CLARA L. REV. 1171 (2012); Nathan A. Greenblatt, *Self-driving Cars and the Law*, 53 IEEE 46 (2016); Gary E. Marchant & Rachel A. Lindor, *The Coming Collision between Autonomous Vehicles and the Liability System*, 52 STA. CLARA L. REV. 1321 (2012); Maurice Schellekens, *Self-driving Cars and the Chilling Effect of Liability Law*, 31 COMPUT. L. SECUR. REV. 506 (2015).

have low emergence and low social valence. According to the framework, therefore, they can best be analogized to domesticated animals. This idea gives background for a desirable regulatory approach. It remains independent of specific regulatory modifications that regulators might choose to have—such as the possibility of a system shutdown. More importantly, it does not prevent specific analogies that we might establish for particular legal problems.

Imagine a self-driving car suffers a system failure and drives off a cliff.[108] The person inside the car survives with severe injuries and sues the car manufacturer. For simplicity, imagine the car's programmers are employees of the manufacturer and not independent contractors. Is the company liable for the car's unexpected behavior?

The domesticated animal analogue would lead us to conclude the company is strictly liable. As the manufacturer owns the driving program, it would be liable for its malfunctions—but not for the car itself, which is owned by the principal passenger. A dog owner is liable under strict liability for injuries caused when her dog bites a neighbor, even if the dog never bites and this behavior is highly unexpected.[109] The dog owner is the cheapest cost avoider; she can more easily prevent the injury than can the neighbor, setting aside minimal precautions the neighbor can take as part of his general duties of care.[110] The same result and justification makes sense for autonomous cars. The analogy successfully captures the features autonomous vehicles share with domesticated animals that matter most for tort law. When the vehicle is in self-driving mode, the manufacturer's programmer is the person who can both best control it and predict its failure at the lowest information cost.[111] Thus,

---

108. See *Tesla fatal crash: 'autopilot' mode sped up car before driver killed, report finds*, THE GUARDIAN, June 7, 2018, *https://www.theguardian.com/technology/2018/jun/07/tesla-fatal-crash-silicon-valley-autopilot-mode-report* [https://perma.cc/A7VQ-V72K] (reporting on a similar incident in which a Tesla car crashed against a highway divider).

109. Common law traditionally holds that owners of domestic animals such as dogs are strictly liable for any foreseeable harm their animal causes; typically, after "one free bite," the dog's "vicious propensities" are established and future harm becomes foreseeable. *See, e.g.*, Collier v. Zambito, 807 N.E.2d 254 (N.Y. App. Ct. 2004). Statutes in many states go further, making a pet owner strictly liable even for harms caused by a dog that did not show vicious propensities. *See, e.g.*, Pingaro v. Rossi 731 A.2d 523 (N.J. App. Ct. 1999) (interpreting "dog bite" statute).

110. *See generally*, Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972) (introducing the concept of the cheapest cost avoider).

111. There could be other regulatory alternatives building on this analogue. For instance, there could be an industrial policy that incentivizes the robot development industry, as there have been in other areas such as weapon technology for the military, and nuclear powers. A tax could be established with a general "social robotic insurance." In that case, the State would be subsidizing innovation by subsidizing insurance premiums. This would relieve pressure from an industry that is not yet mature.

applying liability to the manufacturer makes sense, both from a corrective justice and deterrence standpoint.[112]

A different analogue, such as that of a child, would have shifted the legal question from whether the elements of strict liability are present to whether there was negligent supervision. Here again, the "parent" would be the company that owns the software, and the principal passenger would be a licensee. However, this analog would have failed to capture the technology's low level of emergence: the driving software has no agency, so its programmers have a more direct relationship with its actions than do parents with those of their children. What we want from them is not ex-post active supervision, but rather ex-ante care in the design and programming. Consequently, this analogy would have failed to set liability in a way that it induces adequate levels of care.[113]

C. Copyright Law

Another contemporary regulatory debate this framework implicates is the question of how the law should allocate rights of authorship for works composed by a writing algorithm.[114] How do emergence and social valence affect copyright law? The easiest legal question arises if the A.I. agent is advanced enough to be sentient and conscious. In that case, its level of emergence and social valence lead to the analogy of a human, and the law should consider allocating a copyright to the A.I. agent, just as it allocates copyrights to humans. As part of that consideration, we must also ask the extent to which the goals of copyright law apply to such sentient and conscious A.I. agent.[115]

If the law seeks to encourage the production of creative works (ends-focused), and this (advanced and fictional) A.I. responds to these legal incentives,[116] then granting copyright to her for her works would

---

112. *See generally* JULES L. COLEMAN, RISKS AND WRONGS (2002) (explaining a corrective justice approach to tort law); GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS (1970) (explaining a deterrence approach to tort law).

113. One could ask if the case is different from one of product liability, which would analogize the self-driving car to a tool. One consequential difference is that in many common law jurisdictions, manufacturer liability requires fault, whereas the analogy to liability for animals shifts the burden of proof onto the manufacturer. So if strict liability is appropriate there is a policy reason to choose a different analogy than a tool. A second reason is that the analogy of a domesticated animal adds the element of foreseeability. This is as it should be because self-driving cars behave more unpredictably than do most products under strict liability. In the same way that the child analogy would overestimate its level of emergence, the tool analogy would underestimate it.

114. Bridy, *supra* note 13; Grimmelmann, *supra* note 13; Vertinsky & Rice, *supra* note 13.

115. This addresses Posner's objection to reasoning by analogy. See Posner, *supra* note 40, at 768-70.

116. This would also be the case if humans respond to these legal incentives by

be appropriate.[117] If the goal instead is to foster the creation of these works by humans (process-focused), then no copyright should be granted to the A.I. agent.[118] Having applied the legal category of a human to an A.I. agent, the decision to grant copyright then ultimately depends on whether the law values the result (more works created) or the process (more people creating works).[119] If the law values the result, the copyright should be allocated either to the A.I. agent or an entity responsible for it (as copyrights may now be allocated to people or, in employment relationships, their employers). If the law rather values the process, then, given the lack of a human creative process, the work should remain in the public domain.[120]

The most difficult legal question, and the most present one, is what to do with A.I. agents that do not resemble people as much. Picture a writing algorithm with a high level of emergence (unpredictability) but a low level of social valence, such as *Mathgen*, a software that "writes" (randomly generates) mathematics papers.[121] In a similar way to how, above, we viewed emergence (as foreseeability) in order to allocate responsibility for the actions of a self-driving car, here we must look at emergence to allocate what rights, if any, pertain to works created by a writing algorithm. A writing algorithm's emergence will determine the extent to which people who interact with it will respond to legal incentives to create content. Because *Mathgen*'s randomness produces a high level of unpredictability, the applicable legal category for its level of emergence would be that of a wild animal.[122]

Technology that lacks high emergence, and therefore does not generate highly unpredictable work products, would lead to a different

---

creating more writing algorithms, as we will see later in the section.

117. *See* Stanley M. Besen & Leo J. Raskind, *An Introduction to the Law and Economics of Intellectual Property*, 5 J. ECON. PERSP. 3, 11–18 (1991) (explaining how copyright law balances restrictions on access with the creation of incentives to develop new creative works).

118. *But see* Bridy, *supra* note 13, at 9-20 (arguing that all creativity is algorithmic and works created by autonomous computers are not relevantly different to those created by humans).

119. *See* Pamela Samuelson, *Privacy as Intellectual Property?*, 52 STAN. L. REV. 1125, 1134 (1999) ("Intellectual property law grants exclusive rights in information-based creations in order to promote development . . . of information and a creation of new property rights [for those purposes] seems almost inevitable.").

120. This is different, of course, from the question of whether the creative process of writing the algorithm's code warrants rights of authorship of the programmers over such code (not the works resulting from it).

121. MATHGEN, http://thatsmathematics.com/mathgen/ [https://perma.cc/3N2N-6F23] (last visited May 28, 2018).

122. Their similar levels of unpredictability is relevant both for allocating rights and allocating responsibility, as seen in the lase section. As discussed before, situations involving wild animals are treated by the common law as ultra-hazardous activities. Therefore, exposing a wild animal subjects the person who did so to strict liability for harm that the animal actually and proximately caused. *See* note 108.

analogy and result: such technology, like a word processor on a laptop, should be analogized to a tool. How such processors contrast with an emergent writing algorithm was evident in the example of *Tay* (although in this section we are concerned with how such emergence affects the allocation of rights and not responsibilities). In *Tay*'s example, we saw that programming *Tay* to type and tweet a message is different from merely typing and tweeting a message directly because, as *Tay* becomes emergent, it will tweet messages that its creators cannot predict. For the purposes of copyright law, the case of writing algorithms like *Mathgen* and *Tay* resembles one in which a wild animal wrote the piece more than one in which a person typed it on her computer.

With the new analogy in mind, we can ask again the extent to which the ends of copyright law are served by granting copyright to these A.I. agents.[123] A similar question was asked in *Naruto v. Slater*, a recent famous case regarding the copyright claim to pictures taken by a wild animal.[124] In the case, a black ape grabbed a camera owned by photographer David Slater and took a selfie, which Slater published and which then became viral. PETA brought an action on behalf of the animal, claiming that the animal was the author. The district court held that only humans can obtain copyright under statute, which the Ninth Circuit upheld.[125] What would be true of granting copyright to *Mathgen* was here found to be true of granting copyright to apes: granting such rights would neither lead to more creative works, nor encourage more people to create these works themselves—and so it would not serve the goals of copyright law.[126]

In the case of *Mathgen*, the works currently remain in the public domain.[127] This contrasts with *Naruto v. Slater*, in which the Ninth Circuit allocated copyright to the photographer, rather than keep the work in the public domain. One rationale for its ruling may be that copyright

---

123. The difference between using this analogy and simply calling the programmer the author of the work (which would analogize the software with a tool) is that we have, and should have, the same policy discussion that we do for animals; that is, answering whether the work should be copyrightable (and by whom) or else remain in the public domain.

124. Naruto v. Slater, No. 15-CV-04324-WHO, 2016 WL 362231 (N.D. Cal. Jan. 28, 2016) (holding that the wild animal lacked standing and rights of authorship should be retained by the photographer who owned the camera and enabled the animal to take the pictures).

125. Naruto v. Slater, 126 U.S.P.Q.2d 1464, 1471 (9th Cir. 2018) (holding that humans alone can obtain copyright under statute).

126. *See* Annemarie Bridy, *The Evolution of Authorship: Work Made by Code*, 39 COLUM. J.L. & ARTS 395, 401 ("[B]ecause U.S. copyright law is grounded in the protection of economic rather than moral rights, it's not inconsistent with first principles to recognize authorship in non-natural persons . . . [but] there may be policy reasons not to proliferate copyrights in this way.").

127. *See* Grimmelmann, *supra* note 13, at 403 ("I would like to talk about computer-authored works—I would like to, except that they don't exist. Copyright law doesn't recognize computer programs as authors, and it shouldn't.").

might incentivize artists to incur costs so that other animals will create similar works.[128] The same logic applies to algorithms like *Mathgen*: the person in a position to incur costs to foster creation is the programmer, and such a person—no less than the photographer in *Naruto v. Slater*—may deserve copyright protection. If one agrees with *Naruto v. Slater* that the photographs should not have remained in the public domain, then one should see favorably a change in doctrine to provide the programmers of writing algorithms with copyright over their works. If one disagrees with the ruling due to how indirectly those incentives might operate, then one should consider that both photographs taken by animals and works of writing algorithms should remain in the public domain.

D. Criminal Law

For another implication of the framework, regarding debates on criminal liability,[129] recall Apple's virtual assistant Siri.[130] As in the previous examples, the framework proposed allows us to treat Siri as something different than a human or a tool. Siri has low emergence, given that its programmers can predict most, albeit not all, of its outputs relatively well; Siri also has low social valence, given that it does not seem to develop relationships of trust.[131] This could make her fit either in the analog of a corporation or the analog of a domesticated animal.[132] While *Mathgen*'s example showed that an analogy must inform legal choices based on the policy goals involved, Siri's example shows that choosing the most appropriate analog itself may depend on the legal question.

In a recent case, an iPhone user who had recently committed murder asked Siri for advice on how to hide a body, and Siri responded with helpful advice that led to his temporary success in hiding the crime.[133] Would Siri be an accessory? To answer this question, we would need to

128. One could object that Slater's photograph was serendipitous, but serendipitous findings are also protected by intellectual property.

129. See Hallevy, *supra* note 10; Hallevy, *supra* note 10; HALLEVY, *supra* note 10; HALLEVY, *supra* note 10.

130. *Siri*, APPLE, https://www.apple.com/ios/siri/ [https://www.apple.com/ios/siri/] (last visited May 28, 2018).

131. This might change in the future as Apple continues to develop Siri. *See, e.g.*, David Pierce, *How Apple Finally Made Siri Sound More Human*, WIRED (Sept. 7, 2017), https://www.wired.com/story/how-apple-finally-made-siri-sound-more-human/ [https://perma.cc/4372-64KS].

132. Note that Siri's level of emergence is much lower than that of Samantha's in the fictional example above.

133. Independent Staff, *Florida Man Accused of Killing His Friend Asked Siri Where to Hide The Body, Court Hears*, THE INDEPENDENT (Aug. 13, 2014), https://perma.cc/V6HS-58BX; Jim Dalrymple II, *Man Accused Of Murder Had a Photo on his Phone of Asking Siri Where to Hide a Body*, BUZZFEED (Aug. 13, 2014,), https://perma.cc/PA5X-8V7C.

allocate responsibility for the outcome (hiding the body). We should then focus on Siri's level of emergence, rather than her social valence: the question does not concern whether the iPhone user could be deceived but rather whether anyone was in a position to predict the outcome, and what incentives should be set for people in such a situation going forward; this will determine if anyone should be seen as an accessory to the crime committed by the user.[134]

Similar to the self-driving car example, Siri's software is owned by Apple, and licensed to iPhone owners.[135] By programming Siri, Apple (including its employees) is capable of taking due care to prevent Siri from aiding iPhone users in murder. While it would not have been easy for Apple to anticipate that Siri would give such advice (and, surely, had the company anticipated the result it would have prevented it), Siri's level of autonomy is not developed enough to make her actions as unpredictable as those of a person.

Because, for this case, emergence is more relevant than social valence, and Siri does have some level of emergence, Siri should not fall in the legal category of a corporation or a tool (equivalent in terms of social valence). Instead, she should fall in the category of a domesticated animal (equivalent in terms of emergence). As with *Mathgen* and *Tai*, the relationship between the programmer and the advice Siri gives is more indirect than is the relationship when a person gives advice directly (as through a chat tool). Unlike *Mathgen* and *Tay*, however, Siri's output is not random, warranting a different analog (a domesticated animal rather than a wild animal).

Under the legal category of a domesticated animal, criminal liability for the company, or for its employees, would be highly unlikely unless there is evidence that the company intended Siri to help cover up murder in such a way.[136] In most jurisdictions pet owners are not criminally

134. A different case would be if Siri deceives or defrauds one of its users. Because, for this question, the person's vulnerability towards Siri is the most important consideration, social valence becomes relevant. To determine one's level of protection against Siri's deceit, the most appropriate analog would again be a domesticated animal. For consumer protection purposes, we should think of a software like Siri as a trained monkey that will do as its programmers say. In this hypothetical example, Apple can anticipate with relative certainty what Siri will do, and can use her ability to have people relate to her at some level to exploit them commercially. Siri could, for example, pretend to care only about a user and then point out that he has been rather sedentary lately, so he should purchase a sponsored fitness app. By analogy, Apple should be held liable for Siri's deceptive behavior, if there were any, as well as for any fraud that Siri committed.

135. *Software License Agreements*, APPLE, https://www.apple.com/legal/sla/ [https://perma.cc/8BUF-BVFN] (last visited May 28, 2018).

136. This operates under the assumption that Siri gave advice. Another interpretation could say that Siri did not engage in criminal behavior at all (no *actus reus*) and did not give advice, but merely passed on information found on the Internet, which the user could have easily found. If a murderer searches "how do I hide a body" on Google, Google would not be criminally liable if it directs her to a website

liable for the actions of their pets unless the pets have been declared high-risk by local law.[137] In tort law, however, this analogy would lead to finding Apple responsible under strict liability.[138] This coincides with the policy aims involved (deterrence and reparation). The supervisory negligence-based responsibility that other analogs would trigger would be inadequate in this case: analogizing to a child would unhelpfully incentivize ex-post supervision, but analogizing to a domesticated pet appropriately incentivizes ex-ante levels of care in design which can more reliably help avoid the harm.

The lack of criminal liability for the company under this analogy leads to the question of whether it would make sense at all to find an A.I. agent, analogized to a human, criminally responsible for an outcome. A central question in criminal law is whether one can either directly or indirectly attribute *mens rea*.[139] A characteristic of criminal law is that it sanctions morally blameworthy behavior for the objectives of deterrence and rehabilitation. From society's perspective, Siri did not engage in morally blameworthy behavior because it has no moral agency (it could not tell right from wrong), which would render such analogy inadequate.

More generally: what would it mean to hold a robot criminally or civilly liable? Would it mean to take its money, lock it up, or deactivate it?[140] Presumably, A.I. agents would also not care about money or time. Fiction aside, it is not even obvious that A.I. agents would have self-preservation motives. Animals are subject to natural selection pressures, but because robots are designed and not naturally selected, they would not necessarily have motives like those of humans or animals, thereby rendering usual forms of punishment and deterrence useless. This line of reasoning coincides with the prior consideration about Siri's level of unpredictability leading to the analog of an animal, not a human. The legal categories of an adult or a child would imply a level of emergence in the sense of autonomy that the technology lacks.

---

which gives out the same information. To curb this question, imagine that, instead, a user asks Siri for advice on how to commit suicide and she gives encouraging advice on how to do so. In this hypothetical, Siri would have committed a crime in some jurisdictions (and not merely been an accessory to one). The considerations above would remain applicable.

137. For example, in New York, dog bites lead to criminal liability for their owners only if the dog had been previously declared to be dangerous, the owner's own negligence led to the bite, and the bite produced a serious injury (death or serious disfigurement). N.Y. AGRICULTURE AND MARKETS LAW §123 (Consol. 2017).

138. *See supra* note 108.

139. MODEL PENAL CODE § 2.01-2.02 (Am. Law Inst., Official Draft 1962).

140. Depending on the implications of the new E.U. Parliament initiative of declaring robots "electronic persons," they might not even be subject to deactivation in European Convention on Human Rights states.

VI. Conclusion

Analogies are the best available method to resolve the regulatory uncertainty surrounding new technologies, particularly technologies having some but not all human characteristics. The proposed framework helps draw salient analogies between new technologies and existing subjects the law already treats. These analogies enable the law to avoid rehashing doctrinal and regulatory debates for each technology, so they allow the law to cope with yet unforeseen advances.

Currently, emergence is essential to define the level of foreseeability and people's liability for A.I.'s actions and social valence is essential to define the protections needed against deception in consumer law. When A.I. technology advances and A.I. robots become more like people, emergence will be relevant to determine moral agency and social valence will be relevant to determine moral patience. In both cases, the most applicable characteristic and analog will be determined by the legal question at hand and by the purposes of the law in the regulated relationship.

For non-specifically regulated purposes, A.I. robots that have very high emergence and very high social valence should be analogized to adults, and those with very low emergence and very low social valence, to common tools. The distinction, however, should not be dichotomous, as many legal categories exist between the two poles. A.I. robots with high emergence and very high social valence should be analogized to children for liability purposes. Those with low emergence and high social valence are most analogous to domesticated animals. Those with low emergence and low social valence should be treated similarly to wild animals. And those with low emergence and very low social valence can be analogized to corporations.

This set of legal categories captures the disruptive elements of A.I. for the law. As such, it begins to address some of the complex doctrinal questions involving one of the most disruptive technologies of our time. Areas of law, such as tort law, may not need to change substantively to accommodate A.I., so long as their existing legal categories are applied intelligently.