

# You Might Be a Robot<sup>1</sup>

Mark A. Lemley<sup>2</sup> & Bryan Casey<sup>3</sup>

## Abstract

*As robots and artificial intelligence (AI) increase their influence over society, policymakers are increasingly regulating them. But to regulate these technologies, we first need to know what they are. And here we come to a problem. No one has been able to offer a decent definition of robots and AI—not even experts. What’s more, technological advances make it harder and harder each day to tell people from robots and robots from “dumb” machines. We’ve already seen disastrous legal definitions written with one target in mind inadvertently affecting others. In fact, if you’re reading this you’re (probably) not a robot, but certain laws might already treat you as one.*

*Definitional challenges like these aren’t exclusive to robots and AI. But today, all signs indicate we’re approaching an inflection point. Whether it’s citywide bans of “robot sex brothels” or nationwide efforts to crack down on “ticket scalping bots,” we’re witnessing an explosion of interest in regulating robots, human enhancement technologies, and all things in between. And that, in turn, means that typological quandaries once confined to philosophy seminars can no longer be dismissed as academic. Want, for example, to crack down on foreign “influence campaigns” by regulating social media bots? Be careful not to define “bot” too broadly (like the California legislature recently did), or the supercomputer nestled in your pocket might just make you one. Want, instead, to promote traffic safety by regulating drivers? Be careful not to presume that only humans can drive (as our Federal Motor Vehicle Safety Standards do), or you may soon exclude the best drivers on the road*

*In this Article, we suggest that the problem isn’t simply that we haven’t hit upon the right definition. Instead, there may not be a “right” definition for the multifaceted, rapidly evolving technologies we call robots or AI. As we’ll demonstrate, even the most thoughtful of definitions risk being overbroad, underinclusive, or simply irrelevant in short order. Rather than trying in vain to find the perfect definition, we instead argue that policymakers should do as the great computer scientist, Alan Turing, did when confronted with the challenge of defining robots: embrace their ineffable nature. We offer several strategies to do so. First, whenever possible, laws should regulate behavior, not things (or as we put it, regulate verbs, not nouns). Second, where we must distinguish robots from other entities, the law should apply what we call Turing’s Razor, identifying robots on a case-by-case basis. Third, we offer six functional criteria for making these types of “I know it when I see it” determinations and argue that courts are generally better positioned than legislators to apply such standards. Finally, we argue that if we must have*

---

<sup>1</sup> © 2019 Mark A. Lemley and Bryan Casey.

<sup>2</sup> William H. Neukom Professor, Stanford Law School; partner, Durie Tangri LLP.

<sup>3</sup> Fellow, Center for Automotive Research at Stanford (CARS); Lecturer, Stanford Law School.

Thanks to Ryan Calo and Rose Hagan for comments on an earlier draft.

*definitions rather than apply standards, they should be as short-term and contingent as possible. That, in turn, suggests regulators—not legislators—should play the defining role.*

## Introduction

*“If it looks like a duck, and quacks like a duck, we have at least to consider the possibility that we have a small aquatic bird of the family Anatidae on our hands.”*

—DOUGLAS ADAMS<sup>4</sup>

*“If it looks like a duck and quacks like a duck but it needs batteries, you probably have the wrong abstraction.”*

—DERICK BAILEY<sup>5</sup>

In the heat of the 2018 midterms, robots seemed poised to intervene in a second consecutive election cycle. Noticing an odd pattern of communications from Twitter accounts supporting Ted Cruz, an enterprising journalist tweeted:

At 2 in the morning, hundreds of Ted Cruz supporters all woke up to issue the same exact statement on Twitter. Nothing suspicious about this. Nope. These are all just proud, patriotic American citizens, all deep in the heart of Texas.<sup>6</sup>

Within hours, users had retweeted the journalist’s statement more than 30,000 times.<sup>7</sup> And its subtext—for anyone living under a rock since the 2016 election—was clear. The tweet implied that Cruz’s curiously duplicative supporters were not flesh-and-blood, but bots.

Closer inspection, however, showed something unexpected. The duplicates in question originated from an identical tweet posted by Cruz’s campaign earlier that day.<sup>8</sup> Crucially, Cruz’s original tweet incorporated a relatively new feature: an embedded “button” allowing users to

---

<sup>4</sup> DOUGLAS ADAMS, DIRK GENTLY’S HOLISTIC DETECTIVE AGENCY 216 (1987).

<sup>5</sup> Derick Bailey, *SOLID Development Principles – In Motivational Pictures*, LOSTECHIES (Feb. 11, 2009), <https://lostechies.com/derickbailey/2009/02/11/solid-development-principles-in-motivational-pictures/>.

<sup>6</sup> Lamar White, Jr. (@LamarWhiteJr), TWITTER (Sept. 13, 2018 12:20 AM) <https://twitter.com/LamarWhiteJr/status/1040138113279045632>.

<sup>7</sup> See *id.* (topping 30,000 retweets by September 14, 2018).

<sup>8</sup> Team Cruz (@TeamTedCruz), TWITTER (Sept. 12 2018 9:52 AM), <https://twitter.com/teamtedcruz/status/1039919700241797120?lang=en>. Team Cruz’s original tweet read: “I choose @tedcruz for #TXSen because unlike @betoorourke, I think refusing to stand for the national anthem is disrespectful to those who have sacrificed so much for America.” *Id.* The tweet includes a “Tweet #IStand” button below a campaign video. *Id.*

instantly reproduce it.<sup>9</sup> With a click, an automated script would construct a replica—one that perfectly mirrored the original’s language but did so under the user’s own Twitter moniker. And the script, not the user, controlled when the tweet was sent.

With this discovery, reality again turned upside down. It seemed the accounts under scrutiny were “organic” after all. They had only resembled bots because of the automatic—dare we say robotic—replication made possible by the new feature.

When the dust from the tweet storm finally settled, a cautionary tale unfolded. Real tweets, believed to be fake, produced *genuinely*<sup>10</sup> fake news, believed to be real. The serpent of fakery, it seemed, had eaten its tail. At the center of this odd story was the increasingly unshakable suspicion that virtually everywhere we look humanity is besieged by undercover bots. Indeed, the classic idiom “on the internet, nobody knows you’re a dog”<sup>11</sup> might today be replaced with “no matter where you are, nobody knows you’re a robot.”<sup>12</sup> But as the Cruz campaign’s example shows, an even better formulation may be “nobody knows whether you’re a robot or not.”

Mere weeks after Cruz’s Twitter kerfuffle, this thorny challenge of sorting the “bots” from the “nots” went from an internet oddity to a legal reality. California enacted a statute making it illegal for an online “bot” to interact with consumers without first disclosing its non-human status.<sup>13</sup> The law’s definition of “bot,” however, leaves much to be desired. Among other ambiguities, it bases its definition on the extent to which “the actions or posts of [an automated] account are not the result of a person,”<sup>14</sup> with “person” defined to include corporations as well as “natural” people.<sup>15</sup> Truthfully, it’s hard to imagine any online activity—no matter how automated—that is “not the result of a (real or corporate) person” at the end of the day.<sup>16</sup>

---

<sup>9</sup> See Team Cruz, *supra* note 8.

<sup>10</sup> Sadly, we have to add the prefix “genuine” because even this phrase, in an Orwellian twist, no longer refers to news that is actually fake.

<sup>11</sup> Glenn Fleishman, *Cartoon Captures Spirit of the Internet*, N.Y. TIMES (Dec. 14, 2000), <https://www.nytimes.com/2000/12/14/technology/cartoon-captures-spirit-of-the-internet.html> (describing the New Yorker comic turned global catchphrase). *see also* *On the Internet, Nobody Knows You're a Dog Images*, KNOW YOUR MEME, <https://knowyourmeme.com/memes/on-the-internet-nobody-knows-youre-a-dog/photos>.

<sup>12</sup> See, e.g., David Kushner, *On the Internet, Nobody Knows You're a Bot*, WIRED (Sept. 1, 2005), <https://www.wired.com/2005/09/pokerbots/>.

<sup>13</sup> Cal. SB-1001 (available at [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201720180SB1001](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001)).

<sup>14</sup> *Id.* § 19740(a).

<sup>15</sup> *Id.* § 19740(d) (defining “person” as “a natural person, corporation, limited liability company, partnership, joint venture, association, estate, trust, government, governmental subdivision or agency, or other legal entity or any combination thereof”).

<sup>16</sup> See *id.*

But it's not just chatbots that present these definitional challenges. As robots and artificial intelligence (AI)<sup>17</sup> come to play greater roles in all areas of life—from driving, to weather forecasting,<sup>18</sup> to policing<sup>19</sup>—analogous issues have begun cropping up across a staggeringly diverse array of contexts.<sup>20</sup> In recent years, we've seen Google ask that their robots be regulated as if they were humans.<sup>21</sup> We've seen “pseudo-AI” companies ask that their human workers be regulated as bots.<sup>22</sup> We've seen countries grant robots legal rights (Saudi Arabia, for example, granted citizenship to a robot in 2017).<sup>23</sup> What's more, we've not only seen bots pretending to be human—the concern that prompted the California law—but an increasing number of humans pretending to be bots. One delightful example comes from Ford. In 2017, the automaker resorted to dressing its human drivers as car seats to run “driverless” vehicle experiments, thanks to state laws which forbade operating a car without a human driver at the wheel.<sup>24</sup> But beyond this somewhat cartoonish example lie many troubling ones. In fact, a host of emerging technologies like “DeepFakes,” “Lyrebird,” and “Duplex” make it easier to realistically pretend to be something you're not, without having to resort to dressing as a car seat.<sup>25</sup>

---

<sup>17</sup> We include both “traditional” robots and AI within the scope of this paper. On the problem with defining robots (and therefore the scope of what we cover) for purposes of this paper arguing that you can't define robots, see *infra* Part I.A.

<sup>18</sup> Bennat Berger, *The Predictive Powers of AI Could Make Human Forecasters Obsolete*, VENTURE BEAT (Feb. 28, 2018), <https://venturebeat.com/2018/02/28/the-predictive-powers-of-ai-could-make-human-forecasters-obsolete/>.

<sup>19</sup> See Andrew Selbst, *Disparate Impact in Big Data Policing*, 52 Ga. L. Rev. 109 (2017).

<sup>20</sup> See Part I.C and Part I.D *infra*.

<sup>21</sup> Jerry L. Mashaw & David L. Harfst, *From Command and Control to Collaboration and Deference: The Transformation of Auto Safety Regulation*, 34 YALE J. ON REG. 167, 268 (2017) (discussing Google's request for regulatory clarity regarding its “Self-Driving System”).

<sup>22</sup> See, e.g., *id.*; Richard Wray, *SpinVox Answers BBC Allegations Over Use of Humans Rather Than Machines*, THE GUARDIAN (Jul. 23, 2009), <https://www.theguardian.com/business/2009/jul/23/spinvox-answer-back>; Ellen Huet, *The Humans Hiding Behind the Chatbots*, BLOOMBERG (Apr. 18, 2016), <https://www.bloomberg.com/news/articles/2016-04-18/the-humans-hiding-behind-the-chatbots>.

<sup>23</sup> See Zara Stone, *Everything You Need to Know About Sophia, The World's First Robot Citizen*, FORBES (Nov. 7, 2017), <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#8bf42f046fa1>. Now if only Saudi Arabia would treat women as citizens . . .

<sup>24</sup> See Aarian Marshall, *That Guy Dressed Up as a Car Seat to Solve a Robocar Riddle*, WIRED (Aug. 8, 2017), <https://www.wired.com/story/virginia-self-driving-car-seat-disguise-van/>; Darrell Etherington, *Ford Disguised a Man as a Car Seat to Research Self-Driving*, TECH CRUNCH (Sept. 17, 2017), <https://techcrunch.com/2017/09/13/ford-disguised-a-man-as-a-car-seat-to-research-autonomous-driving/> (noting that Ford resorted to the costume because “you actually still do need to have someone behind the wheel in real-world testing”).

<sup>25</sup> See *infra* notes \_\_\_ – \_\_\_ and accompanying text.

The blurring of these lines doesn't stop there, however. For the foreseeable future, many entities that we think of as "robots" – and a surprising number we think of as "people" – are better understood as hybrids falling somewhere in between. Most self-driving cars, for example, aren't fully self-driving. They operate autonomously in some circumstances. But humans are there to take over when the unexpected happens.<sup>26</sup> Meanwhile, many human drivers aren't fully autonomous either. They rely on electronic maps to update traffic conditions in real time, on calendar apps to tell them when to leave, and on embedded technologies like adaptive cruise control or lane detection to help them get from Point A to Point B safely. And cars are just one example. Though we rarely think about it, we've actually outsourced significant parts of our memory and information processes to devices like search engines, cell phones, and smart watches. True, so far it's pretty easy to tell whether something is a human or a robot. But even today we have considerable trouble distinguishing a robot from other machines we perceive as lacking intelligence. Is a Tesla, for example, a robot (as opposed to a mere "car") because of its "Autopilot" capabilities? Is a drone with automatic stabilizing technologies a robot, even if it's being flown remotely by a human? How about an airliner whose autopilot system can override human efforts to course correct, as recently occurred with Lion Air?<sup>27</sup>

While many—including ourselves<sup>28</sup>—have written of the policy challenges posed by these emerging technologies, our focus is different.<sup>29</sup> We ask not "What should be done?" but "What should it be done to?" The law will regulate robots, human enhancement technologies, and many things in between. Indeed, it already does. But the blurring of the lines between machines, robots, and humans means that regulations specifically targeting robots need to be pretty clear about exactly who or what they're attempting to regulate. So too, for that matter, do regulations targeting humans but not robots.

Simply defining "robot" may seem like an obvious place to start. But as California's misbegotten "bot" legislation and Ford's costumed car seat indicate, crafting a one-size-fits-all definition can be surprisingly hard. Indeed, our central claim is that it can't be done, at least not well. The overlap between people, algorithms, computers, robots, and ordinary machines is

---

<sup>26</sup> But maybe not for long. Waymo now has permission to operate cars on all California roads with no one behind the wheel. <https://www.sfchronicle.com/business/article/Waymo-gets-green-light-for-robot-cars-in-13349173.php>.

<sup>27</sup> See *Lion Air Crash: Investigators Say Plane was 'Not Airworthy'*, BBC NEWS (Nov. 28, 2018), <https://www.bbc.com/news/world-asia-46121127>.

<sup>28</sup> See, e.g., Mark A. Lemley & Bryan Casey, *Remedies for Robots*, \_\_ U. Chi. L. Rev. \_\_ (forthcoming 2019); Bryan Casey, *Amoral Machines, Or: How Roboticians Can Learn to Stop Worrying and Love the Law*, 111 Nw. U. L. Rev. 1 (2017).

<sup>29</sup> See, e.g., Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Cal. L. Rev. (forthcoming 2019); Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671 (2016); Stuart Minor Benjamin, Debate: *Algorithms and Speech*, 161 U. PA. L. REV. 1445 (2013); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87 (2014); Toni M. Massaro & Helen Norton, *Siri-ously? Free Speech Rights and Artificial Intelligence*, 110 Nw. U. L. Rev. 1169 (2016); Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 Calif. L. Rev. 513, 514–15 (2015).

sufficiently great that there is no good legal definition of a robot. As the great computer scientist Alan Turing observed almost a century ago, there's something exceptional about robots and AI that make them exceptionally difficult to define.<sup>30</sup> And, in the end, it might be impossible to come up with a satisfying definition that regulates only the robots or humans we *really* want to. This is particularly true because the nature of robots is changing fast, and legal definitions set with today's technology in mind will rapidly become obsolete.

If we need to regulate robots but can't explicitly define them, what do we do? One possible answer is to simply do nothing. But that seems implausible. Robots and AI increasingly affect our lives, and courts, legislatures, and regulators are already facing that fact.<sup>31</sup> A second possibility is to live with imperfect definitions. That will mean confronting both over- and under-regulation, to say nothing of the obsolescence problem and the challenges of gameable definitions.<sup>32</sup>

In this Article, we argue that a better approach is to embrace the ineffable nature of robots and adapt our legal tools accordingly. We may not be able to successfully define robots *ex ante*. But as with obscenity, quite often we will know them when we see them. In other words, a common law, case-by-case approach may provide a promising means of successfully navigating the definitional issues presented by robots—one that builds and adapts its definitions inductively over time rather than trying to legislate it.

Inductive definition has significant implications for how we regulate. First, we should avoid attempts to explicitly define robots in statutes and regulations whenever possible. Society is better served by regulating acts rather than entities. Some behavior may be more common among robots than humans. But it is the behavior and its consequences that we will normally care about, not who (or what) engaged in it. Put another way, given the definitional challenges, the law is better off regulating verbs, not nouns. Second, when we do need to tailor our laws to specific entities, courts are better than legislatures at these sorts of accretive, bottom-up approaches. As such, we should rely on the common law to the extent we can, rather than rushing in with new regulations. Third, if the common law proves insufficient, we should prefer regulatory rulemaking to legislation. Regulation can more easily incorporate evidence and diverse perspectives, and it's also easier to change when we (inevitably) screw it up. Finally, if we do need legislation specific to bots, it should be tailored as narrowly as possible and should include safeguards that allow us to revisit definitions as the technology evolves.

In Part I, we discuss the origins and growth of robots, the blurring of lines between machine and human behavior, and the human impacts that robots are beginning to produce. In Part II, we discuss efforts to define robots and AI in legislation, regulation, and academic discourse, and argue that those efforts are doomed to fail. Finally, in Part III, we offer suggestions for how to regulate robotic behavior even when we don't really know what a robot is.

---

<sup>30</sup> See Part I.A *infra*.

<sup>31</sup> See Part I *infra*.

<sup>32</sup> See Part II.A *infra*.

## I. Opening Pandora's Bots

Today, we're witnessing a "Cambrian explosion"<sup>33</sup> of robotics technologies, but also an explosion of "cyborgs" and "spoofers." And as might be predicted, the proliferation of these technologies has also given rise to previously unimaginable social, economic, and political impacts. This Part maps the origins and growth of robots, the blurring of lines between machine and human behavior, and the human impacts we're beginning to see.

### A. On the Definition of Robots in Papers Arguing That You Can't Define Robots

Before beginning, however, a caveat. For many readers (especially those with technical backgrounds) our every use of the word "robot" (and especially our inclusion of "disembodied"<sup>34</sup> AI and algorithms in the scope of this piece) will invoke Inigo Montoya's infamous rejoinder to Vizzini: "You keep using that word. I do not think it means what you think it means."<sup>35</sup>

We know. Indeed, that's exactly the point. For better or worse, we live in a world where the term is used to describe all manner of entities—from superhumanly-intelligent software systems to simple pneumatic machines.<sup>36</sup> We have no intention of shouting down those who may use the word improperly. Rather, we hope to meet them on their level, wherever it may be. And that means beginning with a definitional dragnet wide enough to catch all comers.

Why adopt this potentially unwieldy approach? Because, unfortunately, it will be the rare committee of roboticists that is tasked with drafting the next piece of legislation or administrative interpretation governing the technology. Tomorrow's archetypical robot regulator might be someone who prefers to use the word "droid" in a statute because that's what she heard in Star Wars; or someone who gets squeamish about chatbots because of the Terminator series' dystopian "Skynet" takeover; or even someone who steadfastly believes "the Internet is, in fact, a series of tubes."<sup>37</sup>

Our criteria, in other words, is subjective, not objective. We use "robot" in this paper to describe anything that is *or could conceivably be* perceived as one. We don't care whether reasonable roboticists might disagree. We care whether untrained, unfamiliar, or unconcerned policymakers might. That means we discuss technologies that fall squarely within the bullseye of

---

<sup>33</sup> See Gill A. Pratt, *Is a Cambrian Explosion Coming for Robotics?*, IEEE SPECTRUM (Aug. 31, 2015), <https://spectrum.ieee.org/automaton/robotics/robotics-hardware/is-a-cambrian-explosion-coming-for-robotics> (arguing that such an explosion is fast approaching, if it hasn't arrived already).

<sup>34</sup> See Part II.A *infra*.

<sup>35</sup> See Nobody 115 & Brad Kim, *You Keep Using That Word, I Do Not Think It Means What You Think It Means*, KNOW YOUR MEME, <https://knowyourmeme.com/memes/you-keep-using-that-word-i-do-not-think-it-means-what-you-think-it-means> (describing the villain's misuse of the word "inconceivable").

<sup>36</sup> See Part II.A *infra* and accompanying notes (explaining why linguistic limits and human psychology conspire to make us especially likely to use the word too broadly).

<sup>37</sup> Sarah Kliff, *The Internet is, in Fact, a Series of Tubes*, WASH. POST (Sept. 20, 2011), [https://www.washingtonpost.com/blogs/wonkblog/post/the-internet-is-in-fact-a-series-of-tubes/2011/09/20/gIQALZwfik\\_blog.html](https://www.washingtonpost.com/blogs/wonkblog/post/the-internet-is-in-fact-a-series-of-tubes/2011/09/20/gIQALZwfik_blog.html) (describing Senator Ted Stevens's widely ridiculed metaphor).

some expert definitions.<sup>38</sup> But it also means we include some unlikely candidates. Yet as we will see, even simplistic software can produce deceptively complex entities—both in terms of how we perceive them and how we define them.<sup>39</sup>

Having said that, a few preliminary distinctions are worth establishing at the outset. Generally speaking, people talk about robots when things have physical form and about AIs when they are referring to software. AIs can (and increasingly will) inhabit robots; self-driving cars are both robots and AIs. Hardware devices can seem like robots even though they're actually controlled by human operators (*e.g.* remote-controlled drones). And not all software is intelligent; the most powerful AI systems tend to (though do not necessarily) incorporate machine learning and the ability to modify their own decision-making algorithms. Whether embodied or disembodied, it is the intelligent characteristics of complex software and robots that are most likely to trigger regulations, and so we will focus on both hardware and software AIs.

With that caveat, we can now rewind the clock. For we begin our discussion at, well, the beginning—with the very first utterance of the word that will preoccupy us for the remainder of this piece.

## B. On the Origins of Robots

“Robot” first appeared on the linguistic stage in Karel Čapek’s 1920 play, *Rossum’s Universal Robots*.<sup>40</sup> The Czech playwright used “robot” in place of “automaton.”<sup>41</sup> Though rarely heard today, the term automaton referred to the types of ironclad “mechanical men” featured in early 20<sup>th</sup> century works of fiction like the *Wizard of Oz*.<sup>42</sup> Čapek derived “robot” from the old Slavonic “rabota,” meaning “servitude,” “forced labor,” or “drudgery”<sup>43</sup>—fitting, given that his fictional inventors billed their bots as “mass produced workers who lack[ed] nothing but a soul.”<sup>44</sup> Two decades later, though, the sci-fi writer Isaac Asimov would liberate the term “robot” from Čapek’s negative connotation. Asimov envisioned robot-filled futures that bordered on

---

<sup>38</sup> All of which are, themselves, are contested (as Part II.A shows).

<sup>39</sup> See, *e.g.*, Part II.B (discussing the human tendency to anthropomorphize even simple programs designed to merely echo our own words).

<sup>40</sup> Christopher Intagliata, *The Origin Of The Word ‘Robot’*, SCIENCE FRIDAY: SCIENCE DICTION (Apr. 22, 2011), <https://www.sciencefriday.com/segments/the-origin-of-the-word-robot/>. The play was also called “R.U.R.” *Id.*

<sup>41</sup> Dennis Jerz, *R.U.R. (Rossum’s Universal Robots)*, D.G. JERZ BLOG (Jun. 13, 1999), <https://web.archive.org/web/20070826040529/http://jerz.setonhill.edu/resources/RUR/index.html> (noting that the widespread success of “immediately popularized the word, supplanting the earlier ‘automaton’”).

<sup>42</sup> See *id.*

<sup>43</sup> See Intagliata *supra* note 40.

<sup>44</sup> See *id.* (quotations removed).

utopian.<sup>45</sup> His novels sealed the word's place in everyday speech and inspired a rising generation to enter the field of robotics.<sup>46</sup>

The sci-fi world may have originated the term robot. But it was the computer scientist Alan Turing who laid the theoretical foundation for actually building one—not in his oft-cited work involving “intelligent machines,”<sup>47</sup> but in a lesser-known piece called *On Computable Numbers* he penned while still a graduate student.<sup>48</sup> In an era when the word “computer” referred exclusively to people who crunched numbers for a living,<sup>49</sup> Turing would lay the groundwork for the profession's obsolescence.<sup>50</sup>

Turing's 1936 piece proposed a theoretical machine capable of accomplishing any task put before it.<sup>51</sup> The only catch? The machine first required programming, meaning that the underlying process for solving the problem had to be reducible to basic, machine-readable instructions. At a time before transistors, much less microprocessors, this meant relying on the “atomic [machine] operations of scanning, marking, erasing, [and] moving to left and right.”<sup>52</sup> Yet, archaic as it may sound today, Turing's concept worked. He showed that if a given task “could be described by simple instructions encoded on a paper tape,” it could be computed by machines.<sup>53</sup> One such machine might sum two input numbers; another might scan an input for a specific symbol; and so forth.

For Turing, though, this was just the beginning. Having established the viability of these “Turing Machines,”<sup>54</sup> he then showed how a higher-order system could sit atop multiple

---

<sup>45</sup> John Markoff, *Technology; A Celebration of Isaac Asimov*, N.Y. TIMES (Apr. 12, 1992), <https://www.nytimes.com/1992/04/12/business/technology-a-celebration-of-isaac-asimov.html>.

<sup>46</sup> See *id.* (quoting the computer science pioneer and founder of MIT's Artificial Intelligence Laboratory, Marvin Minsky, who wrote that after “‘Runaround’ appeared in . . . March 1942 . . . I never stopped thinking about how minds might work).

<sup>47</sup> Here, we refer to Turing's work *Computing Machinery and Intelligence* which famously spawned the “Turing Test.” Alan Turing, *Computing Machinery and Intelligence*, 49 MIND 433 (1950) [hereinafter “*Computing Machinery and Intelligence*”].

<sup>48</sup> Alan Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, s2-42 PROCEEDINGS OF THE LONDON MATHEMATICAL SOCIETY 230 (1937) [hereinafter “*On Computable Numbers*”].

<sup>49</sup> See, e.g., Brynn Holland, *Human Computers: The Women of NASA*, <https://www.history.com/news/human-computers-women-at-nasa>.

<sup>50</sup> See Watson *supra* note **Error! Bookmark not defined.** (asserting that, with *On Computable Numbers*, Turing had “invented the computer”). Charles Babbage and Ada Lovelace actually invented the first computer nearly 200 years ago but lacked the machines to build it. <https://www.brainpickings.org/2015/06/15/the-thrilling-adventures-of-lovelace-and-babbage-sydney-padua/>.

<sup>51</sup> See *On Computable Numbers supra* note \_\_\_ (describing the concept of “Turing Machines”).

<sup>52</sup> See Watson *supra* note **Error! Bookmark not defined..**

<sup>53</sup> See *id.*

<sup>54</sup> See *On Computable Numbers supra* note \_\_\_.

machines as a kind of scaffolding.<sup>55</sup> This “Universal Turing Machine” (UTM) could simulate the inputs and outputs of any single machine—meaning that complex tasks requiring more than one computation could be combined and automated.<sup>56</sup>

With this insight came the possibility of what we now call “software.”<sup>57</sup> And with software came the possibility—if only a distant one—of building machines as complex as Čapek’s and Asimov’s robots. Indeed, Turing saw no reason why machines with sufficient processing power couldn’t one day rival the thinking prowess of even the most talented “human computers” of his time.<sup>58</sup>

This possibility, in turn, raised an intriguing question for the young scientist. Suppose a machine did achieve such a feat one day. How, then, could we tell whether the machine was *truly* thinking, as opposed to blindly executing instructions?

Turing posed this question in a piece that would become one of the most celebrated works in computer science. But surprisingly, he actually found this challenge more daunting than the first. Indeed, the same intellect that was uncowed by the prospect of building a “thinking machine” would simply throw in the towel when asked to explicitly define one.

Ever fond of reducing complex tasks into simpler subparts, Turing launched his inquiry by noting that an attempt to define thinking machines “should begin with definitions of the meaning of the terms ‘machine’ and ‘think.’”<sup>59</sup> And this, in his view, was impossible. Never one to quit at impossible, Turing would turn the question on its head. Turing wrote, “Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.” He framed this new question as a test he called the “imitation game.”<sup>60</sup> The test delegated the task of distinguishing between humans and machine “imitators” to individual judges on a case-by-case basis. Judges were pit against two veiled actors. And after

---

<sup>55</sup> See *id.*

<sup>56</sup> See *id.*

<sup>57</sup> Joel Achenbach, *What ‘The Imitation Game’ Didn’t Tell You About Turing’s Greatest Triumph*, *Wash. Post* (Feb. 20, 2015), [https://www.washingtonpost.com/national/health-science/what-imitation-game-didnt-tell-you-about-alan-turings-greatest-triumph/2015/02/20/ffd210b6-b606-11e4-9423-f3d0a1ec335c\\_story.html?utm\\_term=.83b4bb640349](https://www.washingtonpost.com/national/health-science/what-imitation-game-didnt-tell-you-about-alan-turings-greatest-triumph/2015/02/20/ffd210b6-b606-11e4-9423-f3d0a1ec335c_story.html?utm_term=.83b4bb640349). (arguing that Turing had essentially invented software).

In an interview with the *Washington Post*, physicist Freeman Dyson remarked:

[Turing] invented the idea of software, essentially. It’s software that’s really the important invention. We had computers before. They were mechanical devices. What we never had before was software. That’s the essential discontinuity: That a machine would actually decide what to do by itself.

<sup>58</sup> See *id.* (describing Turing’s prediction for future computers).

<sup>59</sup> See *Computing Machinery and Intelligence supra* note \_\_\_\_.

<sup>60</sup> See *id.*

a probing conversation, they would attempt to determine which was human and which was machine.<sup>61</sup>

Are you a robot or not? For Turing, creating an explicit definition to distinguish between the two *ex ante* was a non-starter. The better approach, in his view, would mirror one famously adopted some fifteen years later by Supreme Court Justice Potter Stewart: “I’ll know it when I see it.”<sup>62</sup>

### C. Gosh, You’ve Got Some Really Nice Toys Here

In the decades after his untimely death, Turing’s insights in *On Computable Numbers* would transform civilization. His contribution gave rise to a software sector that traded tape for transistors and, from there, “ate the world.”<sup>63</sup> As early as the 1960s, software could be found in industries ranging from finance to national defense.<sup>64</sup> Eventually, it became the driving force behind even more transformative breakthroughs, such as personal computing and the Internet.

As for machines exhibiting “thinking” or “intelligence,” though, the results were less rosy. To be sure, Turing’s “Imitation Game” grew in cultural cachet—appearing in novels, scholarship, and films. But even into the 1990s, its practical relevance remained limited. Few, if any, machines exhibited sophistication enough to be mistaken as “intelligent.”<sup>65</sup> Some flirted with the possibility of fooling “Turing Test” judges. But their interactions constituted something closer to parlor tricks than true smarts.<sup>66</sup>

That wasn’t for lack of trying, though. Beginning in the late 1950s, the field of inquiry known as “artificial intelligence” (AI) exploded with interest. A generation inspired by Asimov descended upon college and industry campuses, intent on building machines as intelligent as his

---

<sup>61</sup> See *id.*

<sup>62</sup> *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964) (Stewart, J., concurring) (“I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it.”).

<sup>63</sup> Marc Andreessen, *Why Software is Eating the World*, WALL ST. J. (Aug. 20, 2011), <https://www.wsj.com/articles/SB10001424053111903480904576512250915629460>.

<sup>64</sup> See MICHAEL MAHONEY, *FROM 0 TO 1: AN AUTHORITATIVE HISTORY OF MODERN COMPUTING* (N. Y.: Oxford U.P., 2002 ed.).

<sup>65</sup> JOSEPH WEIZENBAUM, *ELIZA—A COMPUTER PROGRAM FOR THE STUDY OF NATURAL LANGUAGE COMMUNICATION BETWEEN MAN AND MACHINE* (1976) (describing Eliza, a program that exhibited intelligence by formulaically echoing speech back to subjects).

<sup>66</sup> See Luciano Floridi, et al., *Turing's Imitation Game: Still an Impossible Challenge for All Machines and Some Judges—An Evaluation of the 2008 Loebner Contest*, 19 *Minds & Machines*, 145 (2009).

fictional robots.<sup>67</sup> But the field's results proved anticlimactic, leading to a decades-long period of stagnation known as the AI Winter.<sup>68</sup>

During this fruitless period, Turing's test for "thinking machines" loomed large over the field. Some viewed it as a barometer of progress.<sup>69</sup> The inventor, Hugh Loebner, even established a \$100,000 prize for passing it in an attempt to spur progress.<sup>70</sup> But others regarded the Turing Test at best as imperfect and at worst as a distraction.<sup>71</sup> And these critics certainly had justification for doing so. Indeed, Loebner's prize would eventually be claimed by crude devices wielding not intelligence but deception—leading one AI pioneer, Marvin Minsky, to remark in 1995: "I do hope . . . that Mr. Loebner will indeed revoke his stupid prize, save himself some money, and spare us the horror of this obnoxious and unproductive annual publicity campaign."<sup>72</sup> Computer scientists, frustrated with their inability to design intelligent computers, began to resist the idea of human-like intelligence in devices altogether.

Just two years after Minsky's comment, however, the levees holding back true machine intelligence showed their first cracks. In 1997, IBM's Deep Blue chess engine defeated the reigning world champion, Garry Kasparov.<sup>73</sup> IBM's breakthrough was followed by a series of impressive, albeit incremental, accomplishments.

Then the levees burst. Within a five-year span, machines would defeat world "Jeopardy"<sup>74</sup> champions, knock humans off the all-time-high scoreboards of classic and even modern video

---

<sup>67</sup> See James Hendler, *Avoiding Another AI Winter*, IEEE Intelligent Systems (2008), <https://www.computer.org/csdl/mags/ex/2008/02/mex2008020002.pdf>.

<sup>68</sup> *See id.*

<sup>69</sup> See Floridi et al., *supra* note \_\_.

<sup>70</sup> *Id.*

<sup>71</sup> PETER STONE ET AL., STANFORD UNIV., ARTIFICIAL INTELLIGENCE AND LIFE IN 2030: REPORT OF THE 2015-2016 STUD. PANEL 50, 51 (2016), [https://ai100.stanford.edu/sites/default/files/ai\\_100\\_report\\_0831fnl.pdf](https://ai100.stanford.edu/sites/default/files/ai_100_report_0831fnl.pdf). NAT'L SCI. & TECH. COUNCIL, EXEC. OFFICE OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE 6, 25 (2016). "The philosopher Bertrand Russell ridiculed [tests that required the viewer to disprove a computer's humanity, likening it to asking a sceptic to disprove there is a china t]eapot revolving around the sun while insisting the teapot is too small to be revealed." cite

<sup>72</sup> See PETER STONE ET AL., *supra* note \_\_.

<sup>73</sup> See James O'Malley, *The 10 Most Important Breakthroughs in Artificial Intelligence*, TECHRADAR (Jan. 10, 2018), <https://www.techradar.com/news/the-10-most-important-breakthroughs-in-artificial-intelligence>.

<sup>74</sup> John Markoff, *Computer Wins on 'Jeopardy!': Trivial, It's Not*, N.Y. Times (Jun. 16, 2011), <https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html?pagewanted=all&r=0>.

games,<sup>75</sup> and outperform humans at complex image recognition tasks.<sup>76</sup> This series of breakthroughs culminated in DeepMind’s earth-shaking victory over the “Go” grandmaster Lee Sedol—“the robotics community’s own ‘Kasparov moment.’”<sup>77</sup> (And a feat that, beforehand, was routinely described as “decades” in the future.)<sup>78</sup>

Nowadays, figuring out whether a veiled chess master, “9p dan Go player,”<sup>79</sup> poker champion,<sup>80</sup> or even video game pro<sup>81</sup> is a machine is often easy. Pit it against whoever (or whatever) is regarded as the world’s best opponent. If it wins, it’s almost certainly not human.

Why did these breakthroughs arrive only in the last decade, half a century after “Universal Turing Machines” (aka software programs) had established themselves as a global force? Recent advances in the subfield of artificial intelligence known as “machine learning” certainly account for some of the story.<sup>82</sup> But there’s actually no single variable that explains the timing. A confluence of factors appears to be responsible, ranging from cheap and compact computing power, to the democratization of scientific knowledge through networking technology, to modern optics and sensor advances.<sup>83</sup> Combined with other modern advances, these systems have produced something Turing could only have imagined in his Princeton University dorm room: machines with intelligence truly worthy of comparison to Čapek’s robots.

#### D. Blurred Lines

With great intelligence comes great possibility. And intelligent machines are no exception.<sup>84</sup> Today, we’re in the midst of a “Cambrian explosion” of machine learning software,

---

<sup>75</sup> See Sage Lazzaro, *Watch Google’s DeepMind AI Computer Dominate Old Atari Games*, THE OBSERVER (Jan. 26, 2015), <https://observer.com/2015/01/watch-googles-deepmind-ai-computer-dominate-old-atari-games/>; <https://kotaku.com/googles-deepmind-ai-just-beat-two-pros-at-starcraft-ii-1832034863>.

<sup>76</sup> See James O’Malley, *supra* note \_\_. John Markoff, *How Many Computers to Identify a Cat? 16,000*, N.Y. TIMES (Jun. 25, 2012), <https://www.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html>.

<sup>77</sup> See Bryan Casey, *Amoral Machines, Or: How Roboticians Can Learn to Stop Worrying and Love the Law*, 111 NW. U. L. REV. (2017). The phrase refers to IBM Deep Blue’s defeat of the world chess champion Garry Kasparov in 1997.

<sup>78</sup> Tom Simonite, *Google DeepMind AI Declares Galactic War on StarCraft*, WIRED (AUG. 9, 2017), <https://www.wired.com/story/googles-ai-declares-galactic-war-on-starcraft-/>.

<sup>79</sup> The highest professional Go ranking.

<sup>80</sup> See Noam Brown & Tuomas Sandholm, *Libratus: The Superhuman AI for No-Limit Poker* (2017).

<sup>81</sup> At least, for some video games. See Simonite *supra* note \_\_ (describing games that robots have mastered).

<sup>82</sup> For a detailed discussion of “machine learning,” see, e.g., Lemley & Casey *supra* note \_\_.

<sup>83</sup> See Lemley & Casey *supra* note \_\_ at 12–19.

<sup>84</sup> See *id.* The social effects of these possibilities are disputed. Compare Harley Shaiken, *A Robot Is After Your Job; New Technology Isn’t a Panacea*, N.Y. TIMES, Sept. 3, 1980, at A19 with Louis Anslow,

large-scale “good old-fashioned AI” systems,<sup>85</sup> advanced hardware componentry, and “shared reality” apps, among other technologies.<sup>86</sup> Our tendency to anthropomorphize, well, *all of them* has led many to lump these diverse technological categories together as merely “robots.”<sup>87</sup> This Article adopts that convention. But it’s nevertheless worth delineating a few of the key emerging categories.

First, there are the usual “robot” suspects: the factory bots, assembly arms, advanced hardware applications, and military drones we’ve come to know in decades past. Yet even these have evolved dramatically. Many low-tech systems of yesteryear have sprung eyes, ears, noses, and advanced interactive capabilities.<sup>88</sup> These capabilities, in turn, have enabled far more dynamic human-machine interactions—leading to the coinage of “cobots”<sup>89</sup> for those working side by side with humans and, for those of a less mobile variety, “smart” or “IoT” (short for “Internet of Things”) devices.<sup>90</sup> Meanwhile, on the “disembodied” side of the bot spectrum, we still see the scraping, crawling, trawling, and indexing bots of the 1990s dutifully going about the drudgery of cataloging the Internet (albeit, again, often with dramatically increased powers).<sup>91</sup> And so, too, do we see the “daemons,” “scripts,” “sybils,” “botnets,” and “malware” variously described as “bots” or “robots” in years past.<sup>92</sup>

But beyond these longer-lived categories is an awe-inspiring influx of innovation. In the last decade, robots have invaded our roads, skies, offices, farms, hospitals, bedrooms, courtrooms, and digital worlds.<sup>93</sup> By one recent estimate, 80% of enterprises have already

---

*Robots Have Been About to Take All the Jobs for More than 200 Years*, TIMELINE (May 16, 2016), <https://timeline.com/robots-have-been-about-to-take-all-the-jobs-for-more-than-200-years-5c9c08a2f41d>.

<sup>85</sup> See Will Knight, *An Old-Fashioned AI Has Won a Starcraft Shootout*, MIT TECH. REV.: THE DOWNLOAD (Nov. 16, 2018), <https://www.technologyreview.com/the-download/612438/an-old-fashioned-ai-has-won-a-starcraft-shootout/> (describing GOFAI and how it remains capable of beating state-of-the-art machine learning models).

<sup>86</sup> See Pratt, *supra* note \_\_.

<sup>87</sup> See Part II.A *infra* for additional commentary on this.

<sup>88</sup> See Yen Duong, *Researchers Hope a 'Robo-Nose' Could Give K-9 Officers A Break*, DUKE TODAY (Nov. 19, 2018), <https://today.duke.edu/2018/11/researchers-hope-robo-nose-could-give-k-9-officers-break>.

<sup>89</sup> See Bernard Marr, *The Future Of Work: Are You Ready For Smart Cobots?*, FORBES (Aug. 29, 2018), <https://www.forbes.com/sites/bernardmarr/2018/08/29/the-future-of-work-are-you-ready-for-smart-cobots/>.

<sup>90</sup> Robert Gorwa & Douglas Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*, ICA 2018, Prague, 3 (forthcoming 2019), <https://arxiv.org/pdf/1801.06863.pdf>.

<sup>91</sup> See Gorwa & Guilbeault, *supra* note \_\_ (describing these various categories).

<sup>92</sup> See *id.*

<sup>93</sup> For a more detailed discussion of these changes, see, e.g., Lemley & Casey, *supra* note \_\_; Calo, *supra* note \_\_.

invested in robotics, and 63% have invested in machine learning systems specifically.<sup>94</sup> This proliferation has been enabled not just by software breakthroughs but also by distributed computing, powerful new remote operating capabilities, battery improvements, and sensory hardware advances.<sup>95</sup> Today, robots using these modern technologies can rival doctors at diagnosing ailments,<sup>96</sup> advance scientific research,<sup>97</sup> spawn uncannily human digital personas,<sup>98</sup> and even serve as witnesses to murder.<sup>99</sup> They deliver us groceries from the sidewalks and, soon, will deliver burgers (or even human organs<sup>100</sup>) from the skies.<sup>101</sup> And their ever-improving ability to think, sense, and act intelligently is causing them to blur the boundaries of what, exactly, it means to be human.

But this convergence isn't just about the steady march of machines toward more human-like capabilities. It's coming from both ends. Alongside robotics advances have come a host of other technologies better described, roughly, as "human enhancement" or "biohacking."<sup>102</sup> These boundary-blurring technologies include systems such as Elon Musk's "Neuralink," MIT's "Mind-Reading" headsets, and even chip implants that can unlock doors or generate passwords.<sup>103</sup> The end goal of these "cyborg" or "cybot" applications? To turn human bodies into computers. And that's without even considering the human-enhancing computers already so

---

<sup>94</sup> Louis Columbus, *80% Of Enterprises Are Investing In AI Today*, FORBES (Oct. 16, 2017), <https://www.forbes.com/sites/louiscolumnbus/2017/10/16/80-of-enterprises-are-investing-in-ai-today/#21e4a7a54d8e> (this study uses the narrow definition of AI).

<sup>95</sup> See *infra* note \_\_\_ – \_\_\_ and accompanying text.

<sup>96</sup> *AI vs. Doctors*, IEEE SPECTRUM, <https://spectrum.ieee.org/static/ai-vs-doctors>.

<sup>97</sup> Kaveh Wendell, *The Big Picture: Even Scientists Are Being Automated*, AXIOS (Oct. 20, 2018), <https://www.axios.com/automating-science-2e43b5c8-cb4a-44cb-bfce-f26272b1bd4c.html>.

<sup>98</sup> See Will Knight, *Meet the Fake Celebrities Dreamed Up By AI*, MIT TECH. REV.: THE DOWNLOAD (Oct. 31, 2017, 5:20 PM), <https://www.technologyreview.com/the-download/609290/meet-the-fake-celebritiesdreamed-up-by-ai/>. <https://www.theguardian.com/world/2018/nov/09/worlds-first-ai-news-anchor-unveiled-in-china>.

<sup>99</sup> See Louis Casiano, *Amazon's 'Alexa' May Have Witnessed New Hampshire Double Slayings, Cops Say*, FOX NEWS (Nov. 15, 2018), <https://www.foxnews.com/tech/amazon-ordered-to-turn-over-alexa-recordings-in-double-murder>.

<sup>100</sup> See Michelle Hampson, *Maryland Test Confirms Drones Can Safely Deliver Human Organs*, IEEE SPECTRUM (Nov. 19, 2018), <https://spectrum.ieee.org/the-human-os/robotics/drones/test-run-confirms-that-drones-can-safely-deliver-organs-for-transplant>.

<sup>101</sup> See Nick Bastone, *Uber May Start Delivering Burgers by Drones as Soon as 2021 Because Its CEO Says 'We Need Flying Burgers'*, BUS. INSIDER (Oct. 22, 2018), <https://www.businessinsider.de/uber-planning-drone-food-delivery-2021-2018-10?r=US&IR=T>.

<sup>102</sup> Stefan Nicola, *Biohackers Are Implanting Everything From Magnets to Sex Toys*, BLOOMBERG (Oct. 19, 2018), <https://www.bloomberg.com/news/articles/2018-10-19/biohackers-are-implanting-everything-from-magnets-to-sex-toys>.

<sup>103</sup> See *id.* (noting "Even neuroscientists who work in the field, who know full well how difficult it is to build working brain gear that passes muster with medical regulators, feel a sense of potential.").

integrated into our lives that we scarcely see them as such: the cell phones, laptops, and smart watches that put the world's collective knowledge at our fingertips, as well as the digital assistants, lane-keeping systems, and electronic atlases many now depend on to navigate the world.

There is yet another set of entities—ranging from the malicious to the mere troll—that we might call “spoofers.” In this category fall the robots and humans in the business of deception. Some are humans pretending to be robots. They range from the costumed car seat we saw in the introduction<sup>104</sup> to Russia's “most modern robot”—whose advanced capabilities turned many heads in 2018 before it turned out to be a guy in a costume.<sup>105</sup> But robots also pretend to be human. That's been true for years, as frustrated video gamers confronted with bots pretending to be people can attest.<sup>106</sup> Thanks to the democratization of computer science know-how, state-sponsored “influence campaigns,” and a profound cheapening of computing power, this category of spoofer is more prolific than ever. Even actors of limited means can now use bots to generate thousands of public comments,<sup>107</sup> put celebrity faces on pornographic videos,<sup>108</sup> inflate friend and follower lists,<sup>109</sup> or maliciously influence the electoral process.<sup>110</sup>

This is just a sampling of the new robots (and quasi-robots). And the list grows longer by the day. By 2021, for example, it's estimated there will be almost as many personal-assistant

---

<sup>104</sup> See *supra* notes \_\_\_ – \_\_\_.

<sup>105</sup> Andrew Griffin, 'Russia's Most Modern Robot' Revealed to be Just a Person in a Suit, INDEPENDENT (Dec. 12, 2018), <https://www.independent.co.uk/life-style/gadgets-and-tech/news/russia-robot-person-in-suit-fake-hoax-most-modern-advanced-a8680271.html>. The OG of spoofers was the Mechanical Turk. This term, now appropriated by Amazon for online piecework, originally referred to a fake chess playing machine from a century ago that claimed to be automated but in fact concealed the fact that a human inside of it was doing all of the work, See Simon Schaffer "Enlightened Automata", in Clark et al. (Eds),, *The Sciences in Enlightened Europe*, Chicago and London, The University of Chicago Press (1999), pp. 126-165.

<sup>106</sup> See, e.g., *MDY Indus. v. Blizzard Ent.*, 629 F.3d 928 (9<sup>th</sup> Cir. 2010) (permitting World of Warcraft to ban accounts that used bot technology to auto-level).

<sup>107</sup> Issie Lapowsky, *How Bots Broke The FCC's Public Comment System*, WIRED (Nov. 28, 2017), <https://www.wired.com/story/bots-broke-fcc-public-comment-system/>.

<sup>108</sup> Samantha Cole, *We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now*, VICE: MOTHERBOARD (Jan. 28, 2018, 1:13 PM), [https://motherboard.vice.com/en\\_us/article/bjye8a/reddit-fakeporn-app-daisy-ridley](https://motherboard.vice.com/en_us/article/bjye8a/reddit-fakeporn-app-daisy-ridley). See also BuzzFeed (@BuzzFeed), *You Won't Believe What Obama Says In This Video!*, TWITTER (Apr. 17, 2018, 8:00 AM), <https://twitter.com/BuzzFeed/status/986257991799222272>.; All Things Considered: *Technologies to Create Fake Audio and Video Are Quickly Evolving*, (Tim Mak, Nat'l Pub. Radio broadcast Apr. 2, 2018), <https://www.npr.org/2018/04/02/598916380/technologies-to-create-fake-audioand-video-are-quickly-evolving>.

<sup>109</sup> See <https://ag.ny.gov/press-release/attorney-general-james-announces-groundbreaking-settlement-sellers-fake-followers-and> (New York attorney general crackdown on companies that sell fake social media followers).

<sup>110</sup> Chris Baraniuk, *How Twitter Bots Help Fuel Political Feuds*, SCI. AM. (Mar. 27, 2018), <https://www.scientificamerican.com/article/how-twitter-bots-help-fuel-political-feuds/>.

bots on the planet as people. Indeed, everywhere we look, innovators seem bent on making the judge's role in a Turing test ever more difficult. And though the future of these technologies remains unclear, one thing is certain. Our machines will only continue to become more human, and we humans will continue to become more machine.

### E. Tech's Next Killer App

As we've now seen, robots, cyborgs, and spoofers are tremendously varied. But their effects on society are even more so. As Andrew Ng recently stated, "Almost anything you can do with less than a second of mental thought, we can probably now automate."<sup>111</sup> And while this is no doubt an overstatement, it encapsulates just how pervasive the technology's influence has become. Robots of vast sophistication now power Google's search engine, Microsoft's server farms, Wall Street's financial institutions, UPS's logistics chains, and Amazon's warehouses.<sup>112</sup> "Global business value derived from [robotics] is projected to total \$1.2 trillion in 2018, an increase of 70 percent from 2017."<sup>113</sup> And it's "forecast to reach \$3.9 trillion in 2022."<sup>114</sup>

But alongside the benefits ushered in by robots has come an increasing recognition of the technology's "dark side."<sup>115</sup> 2018 was a watershed year. It saw prominent lawmakers and critics accuse Google of "creating an automated advertising system so vast and subtle that hardly anyone noticed when Russian saboteurs co-opted it in the last election;"<sup>116</sup> Facebook of using robots to silo society into "ideological echo chambers;"<sup>117</sup> and Amazon of misusing its facial recognition technology to promote "authoritarian surveillance."<sup>118</sup> The year also saw the violent

---

<sup>111</sup> Kaveh Wendell, *Small, Narrow — and Revolutionary AI*, AXIOS (Oct. 17, 2018), <https://www.axios.com/narrow-ai-small-data-145b688f-489c-4c9c-bc29-837143d1f3c2.html>.

<sup>112</sup> Elizabeth Woyke, *How UPS Uses AI to Outsmart Bad Weather*, MIT TECH. REV. (Nov. 21, 2018), <https://www.technologyreview.com/s/612445/how-ups-uses-ai-to-outsmart-bad-weather/>.

<sup>113</sup> Gartner Newsroom, *Gartner Says Global Artificial Intelligence Business Value to Reach \$1.2 Trillion in 2018*, GARTNER (Apr. 25, 2018), <https://www.gartner.com/newsroom/id/3872933>.

<sup>114</sup> *See id.*

<sup>115</sup> *See* Michael Birnbaum, *Europe's Antitrust Cop, Margrethe Vestager, Has Facebook and Google in Her Crosshairs*, Wash. Post (May 12, 2018), [https://www.washingtonpost.com/world/europe/europes-antitrust-cop-margrethe-vestager-has-facebook-and-google-in-her-crosshairs/2018/05/10/519eb1a0-47cd-11e8-8082-105a446d19b8\\_story.html?noredirect=on&utm\\_term=.274713b7daad](https://www.washingtonpost.com/world/europe/europes-antitrust-cop-margrethe-vestager-has-facebook-and-google-in-her-crosshairs/2018/05/10/519eb1a0-47cd-11e8-8082-105a446d19b8_story.html?noredirect=on&utm_term=.274713b7daad) (quoting Margrethe Vestager, European Commissioner for Competition).

<sup>116</sup> *See* Charles Duhigg, *The Case Against Google*, N.Y. TIMES (Feb. 20, 2018), <https://www.nytimes.com/2018/02/20/magazine/the-case-against-google.html>.

<sup>117</sup> *See id.*

<sup>118</sup> *Emails Show How Amazon is Selling Facial Recognition System to Law Enforcement*, ACLU OF NOR. CAL. (May 22, 2018), <https://www.aclunc.org/news/emails-show-how-amazon-selling-facial-recognition-system-law-enforcement>.

potential of robots made plain. In February, a driverless vehicle deployed by Uber struck and killed a pedestrian without so much as braking.<sup>119</sup> Months later, a counter-robot defense system thwarted an assassination attempt on Venezuelan President Maduro by two remote controlled DJI drones armed with explosives.<sup>120</sup>

Gone, it seems, are the days when tech's most consequential "crashes" were limited to websites. And after decades of enjoying public and political adoration "as a dazzling national asset," society's attitude toward the tech sector and its bots has taken an abrupt turn.<sup>121</sup> Governments, industries, and ordinary citizens are increasingly discovering that the same technologies capable of promoting human flourishing are also capable of undermining it.<sup>122</sup> Indeed, through robotics, it seems we've opened Pandora's box. And we are increasingly keen to regulate what we've found inside.

## II. These Aren't the Droids You're Looking For

The rise of robots represents a technological paradigm shift—one with profound social, economic, and legal ramifications. We don't intend to discuss all of the ways policymakers should respond to these challenges. There's a growing body of scholarship doing precisely that—our own included.<sup>123</sup> Our focus here is different. It's directed not toward robot "regulability,"<sup>124</sup> *per se*, but what might be properly understood as its prerequisite: what we call "definability."

To regulate robots directly, we need to first establish what one is. But as we will see, that can be easier said than done. In this Part, we show why robots pose formidable challenges for those attempting to regulate the technology. Then, we overview the broader constraints posed by technologies that not only lack easy definition but also blur the lines drawn by our existing policy frameworks. Ultimately, we conclude that efforts to create explicit, *ex ante* definitions of the technology are doomed to failure. That leads us, in Part III, to consider the alternatives.

### A. "I Am Not A Robot"

---

<sup>119</sup> See Aarian Marshall, *Uber's Self-Driving Car Just Killed Somebody. Now What?*, WIRED (Mar. 19, 2018), <https://www.wired.com/story/uber-self-driving-car-crash-arizona-pedestrian/>.

<sup>120</sup> Christoph Koettl and Barbara Marcolini, *A Closer Look at the Drone Attack on Maduro in Venezuela*, N.Y. TIMES (Aug. 10, 2018), <https://www.nytimes.com/2018/08/10/world/americas/venezuela-video-analysis.html>.

<sup>121</sup> See Duhigg, *supra* note \_\_.

<sup>122</sup> See *id.* (noting that these "controversies point[s] to [a] growing anxiety that a small number of technology companies are now such powerful entities that they can destroy . . . social norms with just a few lines of computer code").

<sup>123</sup> See Lemley & Casey *supra* note \_\_; sources cited note \_\_.

<sup>124</sup> See Lessig, *supra* note \_\_.

As Part I makes clear, we've opened Pandora's bots. Robots, cyborgs, and spoofers of all shapes and applications are affecting the world—sometimes on a global scale. Given this reality, it is only natural for policymakers to intervene.

Today, our public and scholarly forums bristle with calls for regulation.<sup>125</sup> We've seen the CEOs of Facebook and Google hauled before Congress to explain the role of robots in destabilizing U.S. elections,<sup>126</sup> and FCC officials issue an extensive report examining whether bots mounted “coordinated attacks” on its public comment process involving “net neutrality.”<sup>127</sup> The pressure is strong to regulate these new technologies.

To properly reach the targets of regulation, there must be mechanisms for distinguishing between the entities it should and shouldn't cover. But as we show in the following sections, legislators and regulators are likely to encounter significant difficulties properly defining (and therefore properly regulating) robots.

### 1. *Definability and Impossibility*

Since the earliest writings on robots, “there has been substantial confusion as to exactly what [one] is and what exactly [one] does.”<sup>128</sup> Virtually all of us have encountered the word, or one its close cousins. But most who use it would be hard-pressed to offer a definition. Even among experts, disagreement reigns.<sup>129</sup> What explains this “semantic conundrum?”<sup>130</sup>

---

<sup>125</sup> Commentators include pundits, politicians, media outlets, technical experts, military leaders, and laypeople alike. See, e.g., Madeline Lamo & Ryan Calo, *Regulating Robot Speech*, UCLA L. REV. (forthcoming 2019) (noting that bot concerns “have led to calls for the government to step in”). “These and other concerns have led to calls for the government to step in. Commentators, including the head of the Allen Institute for Artificial Intelligence writing in the New York Times, publicly urge regulators to adopt a requirement that all bots identify themselves as non-human. Some regulators are heeding that call.” Oren Etzioni, *How to Regulate Artificial Intelligence*, NEW YORK TIMES, <https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html> (Sept. 1, 2017). 6 Cal. SB-1001 (available at [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201720180SB1001](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001)). 7 S. 3127 (available at <https://www.congress.gov/bill/115th-congress/senatebill/3127/text>).

<sup>126</sup> See Kevin Roose & Sheera Frenkel, *Mark Zuckerberg's Reckoning: 'This Is a Major Trust Issue'*, N.Y. TIMES (Mar. 21, 2018), <https://www.nytimes.com/2018/03/21/technology/mark-zuckerberg-q-and-a.html>; Jillian D'Onfro, *Google's Sundar Pichai Was Grilled on Privacy, Data Collection, and China During Congressional Hearing*, CNBC (Dec. 11, 2018), <https://www.cnbc.com/2018/12/11/google-ceo-sundar-pichai-testifies-before-congress-on-bias-privacy.html>.

<sup>127</sup> See Lapowsky, *supra* note \_\_.

<sup>128</sup> See Gorwa & Guilbeault, *supra* note \_\_ at 1.

<sup>129</sup> See Matt Simon, *What is a Robot?*, WIRED (Aug. 24, 2017), <https://www.wired.com/story/what-is-a-robot/> (describing range of disagreement).

<sup>130</sup> See *id.*

Some of it owes to the limits of language. After all, many of us use words like “drone,” “chatbot,” or “autonomous vehicle” with apparent precision—often unaware that these coarse categories actually describe a maddeningly complex (and rapidly evolving) amalgamation of human and machine components. Perhaps most commonly, our difficulty categorizing robots stems from our tendency to anthropomorphize them. For better or worse, we love to see ourselves in objects.<sup>131</sup> And if, in the words of Arthur C. Clarke, “any sufficiently advanced technology is indistinguishable from magic,”<sup>132</sup> then it’s fair to say that any sufficiently *intelligent* technology will inevitably be seen as a robot.<sup>133</sup> This explains the ceaseless headlines of robots “stealing our jobs,” accusations of robots acting “racist” or “sexist,” and a recent candlelight vigil UC Berkeley students held for a delivery bot that caught fire on campus.<sup>134</sup> It also explains why even after showing us that a chatbot “therapist” is simply echoing our own words back to us, we still can’t help but carry on extensive conversations with it.<sup>135</sup>

These quirks of language and psychology account not only for why our modern usage of “robot” is so multifaceted, but also why the term is so culturally potent. As Kate Darling recently remarked, “[The word] generates a lot of attention and fascination and sometimes fear.”<sup>136</sup> As she notes, “[I]t’s much sexier to call something a robot than call something a dishwasher.”<sup>137</sup> It’s this emotional—even “magical”—connection we have to robots that can lead to serious

---

<sup>131</sup> See Calo, *supra* note \_\_ at 545–548 (describing our tendency to anthropomorphize robots).

<sup>132</sup> ARTHUR C. CLARKE, *PROFILES OF THE FUTURE: AN INQUIRY INTO THE LIMITS OF THE POSSIBLE* (Revised Ed., Harper & Row, 1973) at 21.

<sup>133</sup> There is, however, a separate strain of human psychology that mentally categorizes robots as incapable of nuanced or complex behavior. As Ryan Calo puts it, “[J]udges seem to aggressively understand a robot to be something without discretion, a machine that is programmed and does exactly what it’s told. . . . Courts don’t have their minds around the differences between people and robots.” Jason Koebler, *Legal Analysis Finds Judges Have No Idea What Robots Are*, Vice News: Motherboard (Feb. 27, 2016), [https://motherboard.vice.com/en\\_us/article/nz7nk7/artificial-intelligence-and-the-law](https://motherboard.vice.com/en_us/article/nz7nk7/artificial-intelligence-and-the-law). Calo’s work pulls together remarkable stories of robots in American law, both as subjects and as cautionary tales. Ryan Calo, *Robots in American Law*, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2737598&download=yes](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2737598&download=yes) (2016). He “concludes 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.” *Id.*

<sup>134</sup> See Mark Prigg, *Artificial Intelligence Is 'Shockingly' Racist and Sexist - But Only Because They Are Being Fed The Wrong Data, MIT Study Finds*, DAILYMAIL (Nov. 12, 2018), <https://www.dailymail.co.uk/sciencetech/article-6399459/Artificial-intelligence-racist-sexist-fed-wrong-data.html>; Shaiken, *supra* note \_\_; The Download Editors, *A Food Delivery Robot Burst Into Flames—And Now People Have Made a Candlelight Vigil For It*, MIT TECH. REV (Dec. 17, 2018), <https://www.technologyreview.com/the-download/612618/a-food-delivery-robot-burst-into-flames-and-now-people-have-made-a-candlelit/>.

<sup>135</sup> See WEIZENBAUM, *supra* note \_\_.

<sup>136</sup> See Simon, *supra* note \_\_.

<sup>137</sup> See *id.*

definitional headaches. We almost can't help but use the term too loosely. And that tendency,<sup>138</sup> in turn, causes us to describe technologies of vastly different forms and functions under the catchall phrase "robot."

This messy typological reality hasn't stopped some from trying to define the term. Expert communities and dictionaries have taken a shot at it.<sup>139</sup> But even relatively recent attempts can appear outdated in mere months.<sup>140</sup> The Merriam-Webster Dictionary—a frequent resource of the Supreme Court—is worth a look, if only to see how far afield we can go. It offers three definitions:

1. a machine that resembles a living creature in being capable of moving independently (as by walking or rolling on wheels) and performing complex actions (such as grasping and moving objects).
- 2a. a device that automatically performs complicated, often repetitive tasks (as in an industrial assembly line).
- 2b. a mechanism guided by automatic controls.
3. a person who resembles a machine in seeming to function automatically . . .<sup>141</sup>

None of those comes close to capturing what might interest regulators in robots. In fact, they don't even come close to distinguishing robots from other machines. True, they make no pretense of being anything but generalist definitions, though the Supreme Court's penchant for looking up words in a statute in generalist dictionaries should give us significant pause when invoking this excuse.<sup>142</sup>

One might hold out hope that experts could do better. But even here, there's little cause for optimism. Indeed, experts themselves are the first to admit there's "no [] good universal

---

<sup>138</sup> As well as other causal factors that are too numerous to address in this Article.

<sup>139</sup> See *infra* note \_\_\_ and accompanying text.

<sup>140</sup> See Lemley & Casey, *supra* note \_\_\_ (discussing the rapidly changing technological landscape); Neil Richards & William Smart, *How Should the Law Think About Robots?*, WEROBOT (May 10, 2013) ("offer[ing] a definition of robots as non-biological autonomous agents that we think captures the essence of the regulatory and technological challenges that robots present"); "However, it appears that these stakeholders often continue to talk past each other, largely due to a lack of basic conceptual clarity."

<sup>141</sup> "Robot." *Merriam-Webster.com. Merriam-Webster*, n.d. Web. 1 Jan. 2018. Ironically, older versions of the dictionary seemingly do a better job at defining "robot"—albeit, still a very imperfect one. See 40 Cust. Ct. 610, 610 (1958) (citing the 1958 Webster's New International Dictionary definition of robot as: "Any automatic apparatus or device that performs functions ordinarily ascribed to human beings, or operates with what appears to be almost human intelligence").

<sup>142</sup> See, e.g., *Octane Fitness v. Icon Health & Fitness*, 572 U.S. \_\_\_ (2014) (resolving dispute over the definition of an "exceptional case" warranting attorneys' fees by looking up the term "exceptional" in the dictionary).

definition” of robot.<sup>143</sup> As Matt Simon sheepishly observed, “[A]sk three different roboticists to define a robot and you’ll get three different answers.”<sup>144</sup> They’ll range from short and sweet, but altogether vague, descriptions (such as “a physically embodied artificial intelligent agent”);<sup>145</sup> to labyrinthine definitions involving multiple categories, clarifications, and case studies (such as the one currently offered by the International Federation of Robotics);<sup>146</sup> to caveat-filled definitions that lean heavily on words like “usually,” “normally,” or “generally” in order to maintain relevance.<sup>147</sup>

To survey the range of definitions is to realize just how difficult this typological challenge is. In fact, history teaches that our efforts to define “robot” follow a familiar trajectory. Definitions appear. They date. They disappear. Some small number survive this churn. But their longevity is born of tradeoffs. They endure by relying on vagaries or generalities that render them under- or over-determined and, therefore, unilluminating. In other words, they don’t capture what it is that makes us want to regulate a particular technology in a particular way.

---

<sup>143</sup> See Simon, *supra* note \_\_ (quoting Kate Darling); Calo, *supra* note \_\_ (noting “[f]ew complex technologies have a single, stable, uncontested definition [and] [r]obots are no exception”).

<sup>144</sup> See Simon, *supra* note \_\_.

<sup>145</sup> See, e.g., Mel Seigel, *What is the Definition of a Robot?*, SERIOUS SCIENCE (Jul. 3, 2015), <http://serious-science.org/what-is-the-definition-of-a-robot-3587> (noting that many define “a robot is a machine that senses, thinks, and acts,” but personally preferring a definition that includes the additional element of communication); Simon, *supra* note \_\_ (quoting Anca Dragan’s definition as: “physically embodied artificially intelligent agent that can take actions that have effects on the physical world”); *id.* (quoting Hanumant Singh’s definition as: “a system that exhibits ‘complex’ behavior and includes sensing and actuation”).

<sup>146</sup> See, e.g., H. James Wilson, *What Is a Robot, Anyway?*, HARV. BUS. REV. (Apr. 15, 2015), <https://hbr.org/2015/04/what-is-a-robot-anyway> (quoting the IFR’s partial definition as: “A reprogrammable, multifunctional manipulator designed to move material, parts, tools, or specialized devices through various programmed motions for the performance of a variety of tasks”).

<sup>147</sup> See, e.g., Simon, *supra* note \_\_ (quoting Kate Darling’s definition as: “a physical machine that’s usually programmable by a computer that can execute tasks autonomously or automatically by itself”); “Robot.” *Merriam-Webster.com. Merriam-Webster*, n.d. Web. 1 Jan. 2018. A.K. GUPTA & S.K. ARORA, *INDUSTRIAL AUTOMATION AND ROBOTICS* (Laxmi Pub. LTD ed. 2007) at 267 (quoting the Webster dictionary definition as: “An automatic device that performs functions normally ascribed to humans or a machine in the form of a human”).

Other definitions: "a reprogrammable manipulator device" British Department of Industry "Robotics is that field concerned with the intelligent connection of perception to action." Mike Brady. Bots are have been generally defined as “automated agents that function on an online platform” (Franklin and Graesser 1996) I'd say, yes, a robot is a physically embodied artificial intelligent agent."

To date, there remains no consensus definition of “robot,” much less its common technological constituents such as “artificial intelligence,” “automating software,”<sup>148</sup> or “sensory perception.” No wonder, perhaps, that Turing viewed the undertaking as impossible.<sup>149</sup>

## 2. “How can it not know what it is?”<sup>150</sup>

Some of the least problematic efforts to sort the bots from the nots have come from the computer science field. And given their origin, it’s perhaps unsurprising that they take a Turing-esque approach.<sup>151</sup> Tools such as IMAGINATION and ASIRRA use “challenge-response” tests to separate humans from machines without resorting to explicit definition.<sup>152</sup> Like the Turing Test, these “exclusion protocols” play their gatekeeping functions by presenting users with challenges that—at least in theory—only humans can accomplish.<sup>153</sup>

By now, we’ve all encountered them. CAPTCHA’s tedious “I am not a robot” test is one especially prevalent example. But protocols like these aren’t there simply to annoy us. Rather, “[i]f you got rid of them, all hell would break loose.”<sup>154</sup> Indeed, the first ever “Robot Exclusion Protocol” (REP) emerged out of necessity.<sup>155</sup> This protocol—made famous by the “robots.txt” embedding—helped (and continues to help) organizations stop poorly executed web crawlers from accessing private information or inadvertently crashing their servers.<sup>156</sup> Since the REP’s introduction, these techniques have expanded beyond the web and into the realms of telecommunication, finance, and media sharing.<sup>157</sup> Today, “bot blockers” help mitigate fraud,

---

<sup>148</sup> See Calo, *supra* note \_\_ (noting “[f]ew complex technologies have a single, stable, uncontested definition [and] [r]obots are no exception”); Simon, *supra* note \_\_ (discussing other contested definitions in the field); Johann Schumann & Willem Visser, *Autonomy Software: V&V Challenges and Characteristics*, IEEE (2006), <https://ti.arc.nasa.gov/m/profile/schumann/PDF/SV2006.pdf> (describing a survey of experts that showed no consensus definition of the term “autonomy”).

<sup>149</sup> See *supra* note \_\_ – \_\_ and accompanying text.

<sup>150</sup> Ridley Scott, *Blade Runner* (1982).

<sup>151</sup> In fact, some of them reference Turing directly. The acronym CAPTCHA, for example, stands for “completely automated public Turing test to tell computers and humans apart.”

<sup>152</sup> Jeremy Elson et al., *Asirra: A CAPTCHA that Exploits Interest-Aligned Manual Image Categorization*, ACM CCS 2007 (2007), <https://www.microsoft.com/en-us/research/wp-content/uploads/2007/10/CCS2007.pdf>.

<sup>153</sup> See *id.*

<sup>154</sup> Erica Naone, *The CAPTCHA Arms Race*, MIT TECH. REV. (FEB. 24, 2009), (<https://www.technologyreview.com/s/412207/the-captcha-arms-race/>).

<sup>155</sup> See *id.*

<sup>156</sup> See *id.*

<sup>157</sup> See Shuman Ghosemajumder, *The Imitation Game: The New Frontline of Security*, INFOQ QCON (Dec. 4, 2015), <https://www.infoq.com/presentations/ai-security> (describing evolutions and uses in industry).

prevent system overloads, thwart bots from tampering with online auctions, and even stop robocalls – all by distinguishing (or purporting to distinguish) between robots and human beings.<sup>158</sup>

But as a host of evidence shows (not to mention the dozens of robocalls you still likely get every year), these tests aren't foolproof.<sup>159</sup> And we're not just talking about Matt Unsworth's stylus-wielding robot that got around the CAPTCHA technology by physically clicking the box.<sup>160</sup> Rather, numerous findings show that it's possible to subvert exclusion protocols by a variety of means.<sup>161</sup> Some efforts rely on robots themselves to "break" screening mechanisms.<sup>162</sup> And where robots fail, attackers can always opt for a "mechanical turk" approach—whereby "spammers simply enlist[] networks of humans to attack CAPTCHAs."<sup>163</sup> The going black market rate for solving them is just pennies per solution.<sup>164</sup>

True, defenders can always respond to sophisticated attackers by upping the ante, *i.e.* making the test harder to pass. And we've seen exactly that. CAPTCHA tests evolved from transcribed snippets of malformed text to trickier image- or object-recognition tasks.<sup>165</sup> These adaptations worked well for a time. For it was not so long ago that only a handful of experts and companies held hopes of automating complex image-recognition tasks. But nowadays, off-the-

---

<sup>158</sup> See *id.*; Bursztein et al., *How Good are Humans at Solving CAPTCHAs? A Large Scale Evaluation* (2010), [https://web.stanford.edu/~jurafsky/bursztein\\_2010\\_captcha.pdf](https://web.stanford.edu/~jurafsky/bursztein_2010_captcha.pdf).

<sup>159</sup> It was only four years ago that *Wired* trumpeted the end of CAPTCHAs, because the "I am not a robot" check box was going to make it possible to definitively identify robots based on how they moved the cursor. Andy Greenberg, *Google Can Now Tell You're Not a Robot With Just One Click*, *WIRED* (Dec. 3, 2014), [https://www.wired.com/2014/12/google-one-click-recaptcha/?fbclid=IwAR2mD5qe\\_abJfDxNGqcVNUvTTo3s6yain8DMYwhcRujBube8LIXzQ7AP48o](https://www.wired.com/2014/12/google-one-click-recaptcha/?fbclid=IwAR2mD5qe_abJfDxNGqcVNUvTTo3s6yain8DMYwhcRujBube8LIXzQ7AP48o).

<sup>160</sup> See Rob Price, *This Robot Can Trick Computers into Thinking It's a Human*, *BUS. INSIDER* (Jan. 27, 2017), <https://www.businessinsider.com/robot-tricks-captcha-into-thinking-its-a-human-by-stylus-video-2017-1>.

<sup>161</sup> See Ghosemajumder, *supra* note \_\_ (describing various means of circumventing exclusion protocols).

"Software that could reverse CAPTCHA with some accuracy by analysing patterns in the generating engine started being developed soon after the creation of CAPTCHA. In 2013, researchers at Vicarious announced that they had developed a system to solve CAPTCHA challenges from Google, Yahoo!, and PayPal up to 90% of the time. In 2014, Google engineers demonstrated a system that could defeat CAPTCHA challenges with 99.8% accuracy. In 2015, Shuman Ghosemajumder, former click fraud czar of Google, stated that there were cybercriminal sites that would defeat CAPTCHA challenges for a fee, to enable various forms of fraud." *Id.* at \_\_.

<sup>162</sup> See *id.*

<sup>163</sup> See Naone, *supra* note \_\_.

<sup>164</sup> See *id.*

<sup>165</sup> See Ghosemajumder, *supra* note \_\_ (describing this evolution).

shelf tools enable developers of remarkably ordinary abilities to do so.<sup>166</sup> And they do so even as it gets harder and harder for real human beings to distinguish the letters from the background.<sup>167</sup>

This trend points to a growing challenge for exclusion protocols. The concept relies on a state of the world whose days appear numbered. For CAPTCHA to work, there must be problems that are easy for humans but hard for machines. But unfortunately, such problems aren't in endless supply. And as the technological state-of-the-art advances, it seems this gap between human and machine is shrinking. We are rapidly approaching a convergence point where robots are at least as good at image recognition as humans. And that means, each day, there are fewer ways to raise the difficulty of tests like CAPTCHA without also making them more difficult for humans. Already, many exclusion protocols are overly exclusive—eliminating machines *and* humans in droves. As the security expert Matt Blaze recently lamented:

[I was] [l]ocked out of my retirement account because I apparently lack sufficient proficiency in identifying "mountains or hills" and "signs" [required by modern CAPTCHAs]. I guess I should feel happy that they have protections in place to ensure that my money can only be stolen by someone with really good image classification skills.<sup>168</sup>

It seems people, just like robots, “can’t rely on their memories.”<sup>169</sup> And even if we could find a magic bullet for separating humans from bots, that wouldn’t stop attackers from recruiting armies of gig workers to help their robots circumvent protocols.

This reality suggests a cautionary tale. New robot detection capabilities seem to give way, inexorably, to new forms of avoidance.<sup>170</sup> This is true of exclusion protocols like CAPTCHA. But it’s also true of a host of other robot-generated “mimicry” or “deception” technologies, such as the face-swapping “Deepfakes,”<sup>171</sup> the voice-mimicking “Lyrebird,”<sup>172</sup> and the intonation-replicating “Duplex.”<sup>173</sup> No matter the specific use case, these technologies seem to share this

---

<sup>166</sup> See *id.*

<sup>167</sup> See *id.*

<sup>168</sup> Matt Blaze (@mattblaze), TWITTER (Dec. 10, 2018 10:47 AM), <https://twitter.com/mattblaze/status/1072201083613777920>.

<sup>169</sup> Ridley Scott, *Blade Runner* (1982).

<sup>170</sup> See, e.g., Rick Gladstone, *Photos Trusted but Verified*, N.Y. TIMES: LENS (May 7, 2014), <https://lens.blogs.nytimes.com/2014/05/07/photos-trusted-but-verified/> (describing the website IZITRU.COM, <https://www.izitru.com>, which is spearheaded by Dartmouth’s Dr. Hany Farid. It allows users to upload photos to determine if they are fakes. The service is aimed at “legions of citizen journalists who want to dispel doubts that what they are posting is real”).

<sup>171</sup> See Sarah Scoles, *These New Tricks Can Outsmart Deepfake Videos—for Now*, WIRED (Oct. 17, 2018), <https://www.wired.com/story/these-new-tricks-can-outsmart-deepfake-videosfor-now/>.

<sup>172</sup> See Bahar Gholipour, *New AI Tech Can Mimic Any Voice*, SCI. AM. (May 2, 2017), <https://www.scientificamerican.com/article/new-ai-tech-can-mimic-any-voice/>.

<sup>173</sup> See Drew Harwell, *A Google Program Can Pass as a Human on the Phone. Should It Be Required to Tell People It’s a Machine?*, WASH. POST, <https://www.washingtonpost.com/news/the->

common theme. Gatekeepers set detection boundaries. Innovative bots break them. Rinse; repeat.<sup>174</sup> And so an arms race is born.<sup>175</sup> Indeed, our very attempts to explicitly define “robot,” it seems, can be incentive enough for innovators to redefine it.<sup>176</sup>

That’s not all bad. Anyone looking forward to riding in self-driving cars should be quite happy that robots are getting better at recognizing whether an image contains a stop sign. The innovation arms race is bringing us new and better technology, and that’s undoubtedly positive. But those same technological trends will make it harder to distinguish humans from robots, blurring the lines of perception, if not reality.

### 3. *Between a Bot and a Hard Place*

It’s not just robots that are changing in ways that make them hard to distinguish from people. So is human behavior. Texting, “liking,” “emoting,” “GIFing,” “video linking,” “sharing,” “autocorrecting,” “smart completing,” “upvoting,” “following,” and “retweeting”—all of these are recent cultural innovations, even on a time scale measured in “robot years.”<sup>177</sup> Today, more and more human interactions occur through the mediation of a computer. And in a strange twist, these new forms of communication have actually made us easier to imitate.<sup>178</sup> After all, it’s precisely the complex grammar, syntax, and intonation of natural conversation that machines have such difficulty navigating.<sup>179</sup> As Charles Seife notes, “Dispense with them, and we’re speaking a much more machine-friendly language.”<sup>180</sup>

Modern bots can now simply “like,” “share,” “follow,” or “retweet” to amplify an online message or account. And that means they’re less likely to give the game away by inadvertently using poor grammar or issuing off-color remarks.<sup>181</sup> When celebrities (or even ordinary folks) use

---

[switch/wp/2018/05/08/a-google-program-can-pass-as-a-human-on-the-phone-should-it-be-required-to-tell-people-its-a-machine/?noredirect=on&utm\\_term=.6035140c540c](https://www.nytimes.com/2018/05/08/a-google-program-can-pass-as-a-human-on-the-phone-should-it-be-required-to-tell-people-its-a-machine/?noredirect=on&utm_term=.6035140c540c) (May 8, 2018).

<sup>174</sup> Fascinatingly, some of the technologies themselves follow a similar pattern. Cade Metz & Keith Collins, *How an A.I. ‘Cat and Mouse Game’ Generates Believable Fake Photos*, N.Y. TIMES (Jan. 2, 2018), <https://www.nytimes.com/interactive/2018/01/02/technology/ai-generatedphotos.html>; Stan Horaczek, *Spot Faked Photos Using Digital Forensic Techniques*, POPULAR SCIENCE (July 21, 2017), <https://www.popsci.com/use-photo-forensics-to-spot-faked-images>.

<sup>175</sup> See Scoles, *supra* note \_\_.

<sup>176</sup> See *id.*

<sup>177</sup> See Marty Swant, *We Spoke With Sophia the Robot to Find Out If She’s More Hype or Glimpse of the Future*, AD WEEK (Jan. 18, 2018), <https://www.adweek.com/digital/how-sophia-the-robot-is-used-to-market-the-future-of-artificial-intelligence/> (using the term “robot years” as a riff off of “dog years”).

<sup>178</sup> See Charles Seife, *Humans Are Getting More Botlike on Twitter*, SLATE (Nov. 02, 2018), <https://slate.com/technology/2018/11/cesar-sayoc-twitter-humans-behaving-bots.html>.

<sup>179</sup> See Seife, *supra* note \_\_.

<sup>180</sup> See *id.*

<sup>181</sup> See *id.*

“social growth” services such as Instagress, Archie, or Boostio to recruit an army of robo-followers, few of us are any the wiser.<sup>182</sup> Similarly, we may not think twice about the veracity of floods of emotive reactions to headline news—at least until we discover that Facebook removed “754 million fake accounts in the [one] quarter alone.”<sup>183</sup>

The popularity of these new forms of interaction owes, in part, to human nature. Our curiosity propels us toward novel ways of interacting. And beyond novelty, there’s also plain old laziness. Companies like Instagram and Twitter live and die by their ability to render communication “frictionless.”<sup>184</sup> So why not make voicing support for our favorite politicians as easy as the “click of a button” (as we saw in the introduction)?<sup>185</sup>

But these changes are sometimes driven by more surreptitious motives. Consider, for example, Facebook’s “like” button. Its original introduction was an act of aesthetic genius. It cleaned up user interfaces sometimes littered with hundreds of comments indicating generic support for a post or remark.<sup>186</sup> But Facebook’s subsequent expansion of the button into six expressive emojis, while responsive to user demand,<sup>187</sup> was perhaps less about user experience and more about its balance sheet.<sup>188</sup> The company’s introduction of the “love,” “haha,” “yay,” “wow,” “sad,” and “angry” emojis “was a subtle move . . . to help [its] advertisers.” Allowing users to signal more specific reactions to posts or events, in turn, allowed Facebook’s advertising robots to target potential customers with much greater precision.<sup>189</sup> Or even to treat robots as customers; one court has held that visits by search engines and other “artificial intelligence” agents, as well as human viewers, count as advertising “impressions.”<sup>190</sup>

And Facebook is far from the only company nudging us toward more machine-friendly interactions for economic gain.<sup>191</sup> Gmail’s “smart complete” feature, for example, provides similarly lucrative information. The emailing service now suggests short, chatty replies to message. And when we defer to the robot’s advice, we provide it with valuable insights. As a

---

<sup>182</sup> See Josh Constine, *Instagram kills off fake followers, threatens accounts that keep using apps to get them*, TECH CRUNCH (Nov. 19, 2018), <https://techcrunch.com/2018/11/19/instagram-fake-followers/>; See Chengcheng Shao et al., *The Spread of Low-Credibility Content by Social Bots*, 9 NATURE COMMS. 4787 (2018).

<sup>183</sup> See Constine, *supra* note \_\_.

<sup>184</sup> See *id.*

<sup>185</sup> See *supra* notes \_\_ – \_\_ and accompanying text.

<sup>186</sup> Which explains why it has taken over the Internet.

<sup>187</sup> See Navin Prakash, *Each Reaction on Facebook Makes it Richer*, Towards Data Science (Jun. 18, 2018), <https://towardsdatascience.com/each-reaction-on-facebook-makes-it-richer-b5ea07f1bbc7>.

<sup>188</sup> See *id.*

<sup>189</sup> See *id.*

<sup>190</sup> *Go2Net, Inc. v. C I Host, Inc.*, 115 Wash. App. 73, 86 (2003).

<sup>191</sup> See Derek Thompson, *I Used Gmail Auto-Complete, and Now I Know I’m Worthless*, THE ATLANTIC (Dec. 14, 2018), <https://www.theatlantic.com/ideas/archive/2018/12/gmails-auto-complete-feature-makes-us-feel-like-robots/578155/>.

recent headline put it, “Google’s auto-complete makes me feel like a robot.”<sup>192</sup> But the more accurate line might read, “Google’s auto-complete lets robots feel more like us.”

The upshot of this brave new world of human-robot interaction? It’s not so much that Turing’s “I’ll know it when I see it” approach no longer works. It’s that “seeing”—at least through the kind of abbreviated interactions Turing originally envisaged—is no longer enough. In a world where robot interactions are getting more human-like and human ones are getting more—well—robotic, there are fewer dead giveaways. And this convergence can put those trying to distinguish machines from humans between a rock and a hard place. “Seeing” a bot now requires something more than a passing glance. It’s simply no longer the case that we can trust someone (or something) who claims “I am not a robot.”<sup>193</sup>

## **B. Actually, You Might Be a Robot**

Above, we saw that defining and distinguishing robots is hard, and grows harder by the day. One possible answer to this challenge is to avoid regulating robots, algorithmic decision making, and AI altogether. But that seems implausible. Each day, more statutes and regulations explicitly govern these technologies or end up inadvertently governing them thanks to laws on the books that never anticipated their emergence.<sup>194</sup>

If our goal is to take a nuanced approach to regulating robots, policymakers must understand three ways robot definitions can fail. Definitions can fail: (1) when drafting laws from scratch to cover robots; (2) when robots interact with existing laws drafted with either people or traditional machines in mind; and (3) when robots defy the definitional bounds of existing regulatory bodies. Want, for example, to crack down on “influence campaigns” deceptively deploying social media robots? Be careful not to define “robot” too broadly. For the supercomputer nestled in your pocket, your clothing, or even under your skin might just make you one. Want to promote traffic safety by regulating drivers? Be careful not to presume that only humans are capable of driving. Failing to anticipate that robots might soon be the best drivers on the road could stall the introduction of life-saving technologies.<sup>195</sup> Want, instead, to appoint regulators to a body tasked with overseeing healthcare professionals? Be careful not to tap only those with expertise in the human side of health, because, thanks to machine diagnostic tools that rival the abilities of doctors, you might soon be regulating robots.

In the following sections we explore each of these problems.

### *1. Failing to Define Robot Carefully*

---

<sup>192</sup> *See id.*

<sup>193</sup> This references the prompt that usually accompanies CAPTCHAs. *See supra* notes \_\_\_ – \_\_\_.

<sup>194</sup> *See infra* notes \_\_\_-\_\_\_ and accompanying text.

<sup>195</sup> As we’ve seen with driverless vehicles and Federal Motor Vehicle Safety Standards. *See infra* notes \_\_\_ – \_\_\_ and accompanying text.

The most straightforward definitional problems come from efforts to define which robots are the subject of regulation. As this section explores, even our most careful efforts to craft explicit definitions of robots can have unintended consequences.

One consequence is confusion. “When words are used sloppily, concepts become fuzzy, thinking is muddled, communication is ambiguous, and decisions and actions are suboptimal, to say the least.”<sup>196</sup> The California Bot Disclosure Act’s definition, for instance, is internally inconsistent.<sup>197</sup> It’s simply not possible to read it in a rational way. We’ve seen similar problems in other efforts to define new technology early in the process. The Computer Fraud and Abuse Act (CFAA),<sup>198</sup> for instance, attempts to define computer hacking, but has made a hash of it.<sup>199</sup> One attempt at legislative reform, and numerous court interpretations, haven’t been able to fix it in over thirty years.<sup>200</sup> As such, we’re stuck with a law we just can’t really understand.<sup>201</sup> Hopefully that won’t be true of California’s new bot bill. But it might be.

A second problem with new definitions is that they can be overbroad, entangling actors that are, in the famed meme, “not the droids you’re looking for.”<sup>202</sup> Once again, the CFAA provides a good example. That Act defined a “federal interest computer” in a way that initially seemed limited, but which rapidly grew to encompass any computer or device connected to the Internet, which is today essentially every computer and phone—and a surprisingly large number of refrigerators.<sup>203</sup> Maybe we need a law regulating authorized access to refrigerators,<sup>204</sup> but it is unlikely the CFAA is the right one, precisely because it covers a number of devices no one thought in 1986 would come within the statute.

We’ve already begun to see overbreadth in robot statutes and regulations. The California Bot Disclosure Act is overbroad in some respects; we can imagine any number of actions or posts not generated by robots that are nonetheless “not the result of a person.”<sup>205</sup> And we’ve seen similar problems with local and municipal level efforts. The city of Houston, for example, was recently swept up in a moral panic over “robot sex brothels”—leading to a hasty ordinance that banned far more than just bots (more on this in Part III.B).<sup>206</sup> Meanwhile, the New York City

---

<sup>196</sup> See S. Kaplan, *Bayes is for Eagles*, 39 IEEE TRANSACTIONS ON RELIABILITY 130, 131 (1990).

<sup>197</sup> See *supra* notes \_\_\_ – \_\_\_ and accompanying text.

<sup>198</sup> See Kerr, *Norms of Computer Trespass*, *supra* note \_\_\_; Mayer, *supra* note \_\_\_; Orin Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, 94 MINN. L. REV. 1561, 1563-65 (2010).

<sup>199</sup> See generally Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, *supra* note \_\_\_.

<sup>200</sup> See *id.*; Kerr, *Norms of Computer Trespass*, *supra* note \_\_\_; Mayer, *supra* note \_\_\_.

<sup>201</sup> That sometimes rises to the level of a constitutional problem. See generally Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, *supra* note \_\_\_.

<sup>202</sup> Star Wars: A New Hope (1977). Or, for the purists among us, just “Star Wars.”

<sup>203</sup> See Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, *supra* note \_\_\_.

<sup>204</sup> A large number of diet books suggest we might.

<sup>205</sup> See Cal. SB-1001, *supra* note \_\_\_.

<sup>206</sup> See *infra* notes \_\_\_ – \_\_\_ and accompanying text.

Council recently penned a bill aimed at facial recognition whose wording is so broad it likely extends to any image capturing system.<sup>207</sup> We've also seen federal legislation proposed or enacted that relies on extremely broad definitions of robots or artificial intelligence. Examples include the AI JOBS Act (H.R. 4829), which calls for a Department of Labor report on the impact of AI on the workforce by defining the term in a way that essentially includes any software;<sup>208</sup> and the Bot Disclosure and Accountability Act of 2018 (S.3127)<sup>209</sup> which has issues similar to the California bill.<sup>210</sup>

If robot definitions can be overbroad, they can also be underinclusive. There's no shortage of laws doomed to irrelevance because they tried to predict and regulate the way technology would develop.<sup>211</sup> A law that focused on a particular method of AI learning, for instance, could easily become irrelevant if the nature of the technology changes. Ironically, the California Bot

---

<sup>207</sup> See Int. No. 1170 (available at <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3704369&GUID=070402C0-43F0-47AE-AA6E-DEF06CDF702A&Options=Advanced&Search=>) (defining "biometric identifier information" as "a retina or iris scan, fingerprint, voiceprint, or scan of hand or face geometry, any of which is collected, retained, converted, stored or shared to identify an individual").

<sup>208</sup> See H.R. 4829 (available at: <https://www.congress.gov/bill/115th-congress/house-bill/4829/text>) (defining the term "artificial intelligence" as anything that can: "(A) think like humans (including cognitive architectures and neural networks); (B) act like humans (such as passing the Turing test using natural language processing, knowledge representation, automated reasoning, and learning); (C) think rationally (such as logic solvers, inference, and optimization); (D) act rationally (such as intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision-making, and acting); or (E) automate or replicate intelligent behavior").

<sup>209</sup> See S.B. 3127 (2017–18) (available at: <https://www.congress.gov/bill/115th-congress/senate-bill/3127/text>) (defining the term "bot" to include any "automated software program or process intended to impersonate or replicate human activity online"). The bill does a better job than most of focusing on functional characteristics, as opposed to focusing on entities themselves. But, as we'll see below, including other functional considerations in the definition (such as whether the reasons for bot "impersonation" are well intentioned) could strengthen it considerably. The current wording, for example, would prevent well-intentioned researchers from concealing the identity of bot accounts.

<sup>210</sup> Others include the National Security Commission Artificial Intelligence Act of 2018 (H.R. 5356), Fundamentally Understanding the Usability and Realistic Evolution of Artificial Intelligence Act of 2017 (FUTURE of AI Act) (H.R. 4625), and the AI in Government Act of 2018 (S.B. 3502).

<sup>211</sup> Semiconductor Chip Protection Act of 1984, 17 U.S.C. §§ 901–914 (Supp. IV 1986) (an entirely new IP right to protect semiconductor "mask works" that was only ever used twice, because the way we make semiconductors changed); R. Jason Richards, *The Utah Digital Signature Act As "Model" Legislation: A Critical Analysis*, 17 J. MARSHALL J. COMPUTER & INFO. L. 873 (1999) (discussing the weakness of the Utah Digital Signature Act of 1995 as "model" legislation for electronic signature technology); Jane Winn, *The Emperor's New Clothes: The Shocking Truth About Digital Signatures and Internet Commerce*, 37 IDAHO L. REV. 353-88 (2001) (explaining how laws written in the 1990s to regulate electronic commerce rapidly became obsolete because they didn't anticipate how people would behave); Calo, *supra* note \_\_ ("the Electronic Communications Privacy Act ("ECPA"), passed in 1986, interacts poorly with a post Internet environment in part because of ECPA's assumptions about how electronic communications would work.").

Disclosure Bill suffers from this shortcoming too. Like Schrödinger’s regulation, it’s simultaneously over- and under-inclusive. Because it applies only to posts that are not the result of a person, and defines person as “a natural person, corporation, limited liability company, partnership, joint venture, association, estate, trust, government, governmental subdivision or agency, or other legal entity or any combination thereof,”<sup>212</sup> it arguably doesn’t include anything that results from a computer owned by a corporation.<sup>213</sup>

Underinclusiveness is a particular risk because parties have an incentive to game the definition, designing their technology to take advantage of loopholes in statutes and regulations. There are any number of examples of this at work. Entire companies have been founded on regulatory arbitrage—Uber and Lyft, to name a few.<sup>214</sup> Sometimes the strategy works, as it did for those companies, and for car companies that avoided fuel-efficiency rules by redefining certain cars as light trucks<sup>215</sup> (and as it generally does for tax shelters).<sup>216</sup> In other cases, courts find ways to stop the arbitrage. But when they do, it’s generally by rewriting the statutes themselves or using malleable common law doctrines to, in effect, change the definition to include the gamer.<sup>217</sup> Sometimes the results fall somewhere in between. Internet language and content filters can block some content, but they are also pretty easy to get around, as any teenager can tell you.<sup>218</sup>

---

<sup>212</sup> See Cal. SB-1001, *supra* note \_\_ § 19740(d).

<sup>213</sup> It is an interesting philosophical question when the acts of a robot or AI are not “the result of a corporation.” A self-learning AI that “goes rogue” and does something unexpected might qualify, though even then it would likely have apparent authority to act on the corporation’s behalf. Microsoft’s Tay might qualify. But what about the Google map-generating AI that “cheated” to generate maps more quickly, <https://techcrunch.com/2018/12/31/this-clever-ai-hid-data-from-its-creators-to-cheat-at-its-appointed-task/>, or the drone that figured out a loophole in its instructions that caused it to start flying in the wrong direction. See Bryan Casey & Mark A. Lemley, *Remedies for Robots*, \_\_ U. Chi. L. Rev. \_\_ (forthcoming 2019).

<sup>214</sup> See, e.g., Bryan Casey, *Uber’s Dilemma: How the ADA May End the On-Demand Economy*, 12 U. MASS. L. REV. 126 (describing transportation network company efforts to avoid traditional taxi regulations by claiming status as a mere “platform”).

<sup>215</sup> Cites

<sup>216</sup> cites

<sup>217</sup> In both *Aereo* and *Grokster*, for instance, the Supreme Court created entirely new rules of copyright law because it didn’t like the fact that the defendants had complied with the letter of the old rules while arguably subverting their intent. For a discussion of these cases and other examples, see Mark A. Lemley & Mark P. McKenna, *Unfair Disruption* (working paper 2019); Dan L. Burk, *Perverse Innovation*, 58 WM. & MARY L. REV. 1 (2016).

<sup>218</sup> See Cory Doctorow, *Today, Europe Lost The Internet. Now, We Fight Back.*, EFF (Sept. 12, 2018), <https://www.eff.org/deeplinks/2018/09/today-europe-lost-internet-now-we-fight-back>. As the Electronic Frontier Foundation recently wrote:

Put it this way: if your occupation is figuring out how filters work and tinkering with getting around them, you can become skilled in the art. The filters used by the Chinese government to block images, for example, can be defeated by simple measures.

Finally, definitions tend to be static categories. Once the term robot is defined in law, we can change the definition only by passing a new one. And that's hard. Definitions written too early in the history of a technology may simply miss an important development altogether. As noted above, one possible consequence of this is irrelevance—the law, for example, may not apply to a whole swath of new technologies. But a potentially more serious risk is that the law may constrain the development of the technology itself by applying a definition written with one technology in mind to a changed world in which the line that once made sense no longer does. A bill written to regulate artificial intelligence in the 1970s would have missed the machine learning revolution altogether, for instance. It might have insisted on rules governing algorithms that assumed humans were always hard-coding them, which might have made modern training methods impossible.

California's bot law provides an interesting example of the problem of definition at work. After receiving word of Governor Brown's signature, the bill's sponsor Senator Bob Hertzberg—himself the victim of a prior bot "influence campaign"<sup>219</sup>—would voice his enthusiasm. His campaign's self-styled "Bot Hertzberg" account announced:

Thank you @JerryBrownGov for signing @SenateHertzberg's bill #SB1001 to require bots like me to be labeled! #BotHertzberg

But there was nothing automated about this post. It was written by the Senator himself (or, more likely, his staff). The Senator apparently missed the irony of broadcasting support for a law regulating bots through an account that, itself, claimed to be a bot but wasn't. In light of the tweet's obviously hand-crafted nature, the missive seemed guilty of the very kind of mischief his bill sought to prohibit.

## 2. *When Robots Encounter Human Laws*

If defining robot seems impossible, maybe we just shouldn't do it. But that comes with its own set of problems. Not defining robots doesn't mean we won't regulate them. Thanks to laws already on the books, your robots might (or might not) already be regulated. And those regulations may treat them as if they were people (or alternatively, as if they were toasters).<sup>220</sup> As applied to robots, regulations written with people (or toasters) in mind often make no sense. But without a specific exception for robots, they may be stuck complying with unnecessary, even silly, laws.

To start, human beings have biological limitations that robots don't. Wage and hour laws, child labor laws, and rules that limit how many hours doctors, pilots, or even truckdrivers can work all operate under the presumption that workers need sleep. Robots don't. But if our laws

---

Meanwhile, these filters that are bound to be thousands of times more effective than any copyright filter because they're doing a much more modest job with far more money and technical talent on hand. *Id.*

<sup>220</sup> And not in the Battlestar Galactica sense of the term.

<sup>220</sup> And not in the Battlestar Galactica sense of the term.

fail to acknowledge the distinction, a self-flying commercial airliner or a self-driving truck might have to be shut down for most of a day to comply with laws that never considered the possibility of a robot pilot or driver. Similarly, self-driving trucks with instantaneous response times don't need to leave as much space between themselves and the cars in front of them as humans do in order to prevent an accident. But the law requires them to hang back just the same. And the law requires various devices in a vehicle, like steering wheels and brake pedals, that are keyed to human anatomy. These assumptions baked into many of our laws can implicitly or explicitly impose all variety of strange obligations on robots, based on the false premise that they share the same relevant characteristics as human beings. Even if these ill-defined laws don't prevent robots from driving, they assume everyone's (or everything's) capabilities are equal without accounting for the new capabilities of robots.<sup>221</sup>

Robots face a similar problem when existing regulations put them in the category, not of humans, but of dumb machines. Take the robot many of us interact with most: the elevator.<sup>222</sup> Early elevators needed human operators to control the starting and stopping motion and to open the doors.<sup>223</sup> Some states and countries therefore required elevators to have human operators.<sup>224</sup> Today, that requirement is silly; the elevator is smart enough to get you where you want to go at the push of a button (or perhaps not even that; many modern elevators have no buttons at all and go to the destination floors identified in a fob or keycard without any passenger intervention). But there are still laws that require human elevator operators because they think of operating an elevator as something humans, not machines, do.<sup>225</sup> Those rules sound ridiculous to us today. But they certainly didn't seem so at the time. And it's precisely the kind of failure of imagination that we must guard against when regulating robots. After all, it's not so hard to imagine the intelligent entities of the future (human or machine) thinking much the same about the idea that complex ground and air transportation must be the exclusive province of people.

The problem here is, in some sense, the inverse of the one in the last section. There, we struggled to define robots in order to treat them differently than human beings or dumb machines. Here, we struggle with laws that aren't written with robots in mind but that nonetheless end up regulating what robots can do.

---

<sup>221</sup> See *infra* notes \_\_ – \_\_ and accompanying text.

<sup>222</sup> Don't think an elevator is a robot? Why not? It responds to simple requests by initiating an automated transportation process involving efficient routing algorithms.

Okay, to be fair, we don't really think of an elevator as a robot either, but it does show some of the problems with efforts to define the term. We – and likely you – just put it in a mental bucket, perhaps because it is familiar technology.

<sup>223</sup> James Shiffer, *The Last Elevator Operators in Minnesota, Maybe*, STAR TRIBUNE (May 2, 2014), <http://www.startribune.com/minnesota-s-last-elevator-operators/257249671/>.

<sup>224</sup> See *id.*

<sup>225</sup> See *id.*; Ernesto Londoño, *Rio de Janeiro Elevator Attendants 'Adore' Their Dying, Chatty Trade*, N.Y. TIMES (Nov. 25, 2018), <https://www.nytimes.com/2018/11/25/world/americas/brazil-rio-de-janeiro-elevator-attendants.html>.

When robots make these problems evident, regulators can respond in different ways—without necessarily rewriting the definitions on the books. Consider self-driving cars. Under general tort principles, the notion of “control” of the vehicle has typically been central to liability.<sup>226</sup> And in the vast majority of accidents, the human sitting in the driver’s seat is the presumptive “controller” for purposes of negligence liability.<sup>227</sup> But what are we to do with the drunk driver who fell asleep in his Tesla, which kept merrily (and safely) driving down the freeway until a cop noticed he was sleeping and tried to pull the car over?<sup>228</sup> The human behind the wheel is in trouble precisely because he *wasn’t* driving when he was supposed to, but it seems implausible to say he was actually in “control” of the car. Unfortunately, however, many assumptions made by state and federal traffic and safety laws seriously complicate this analysis. Numerous laws on the books, for example, explicitly define vehicle “drivers” and “operators” as human beings. Others simply leave the definition ambiguous, having been “written decades ago [by] authors [who] likely never considered the possibility that a car might not have a human driver at all.”<sup>229</sup>

In 2015, Google’s driverless vehicle outfit ran into a similar definitional roadblock when it tried to manufacture a visionary new car. As Mashaw and Harfst note, “many [NHTSA] standards require that a vehicle device or basic feature be located near ‘the driver’ or near ‘the driver’s seating position.’”<sup>230</sup> But Google’s automated vehicles were “entirely controlled by artificial intelligence, something Google called a ‘Self Driving System’ (SDS), such that no driver was needed, or indeed, wanted.” Worried about the costs associated with failed compliance—as well as the tort system’s reliance on vehicle code violations as evidence of negligence *per se*—the company wrote to the National Highway Traffic Safety Administration (NHTSA) expressing concerns that its driverless vehicles might not be able to certify compliance with Federal Motor Vehicle Safety Standards.<sup>231</sup>

NHTSA’s eventual response, however, was less than reassuring. The agency acknowledged the term’s ambiguity but lamented that “it can take substantial periods of time to develop rulemaking proposals and final rules, and that such proceedings are ill-suited as first-line

---

<sup>226</sup> See Bryan Casey, *Robot Ipsa Loquitur: What the First Robot Accused of Negligence Says About the Future of Tort* (working paper 2019) (discussing this phenomenon).

<sup>227</sup> See *id.*

<sup>228</sup> Timothy B. Lee, *It Took Seven Miles to Pull Over a Tesla with a Seemingly Asleep Driver*, ARS TECHNICA (Nov. 30, 2018), <https://arstechnica.com/tech-policy/2018/11/cops-pull-over-tesla-cruising-on-a-freeway-with-apparently-asleep-driver/>. It took seven miles for the cops to persuade the Tesla to pull over. *Id.* They did it by pulling in front of the car and slowing to a halt. *Id.* As an aside, this story indicates that one challenge self-driving cars face is how to understand and obey sirens and orders from ambulances and police.

<sup>229</sup> See *id.*; Mashaw & Harfst, *supra* note \_\_.

<sup>230</sup> See *id.*

<sup>231</sup> See *id.*

regulatory mechanisms to address rapidly-evolving vehicle technologies.”<sup>232</sup> As such, the agency recommended that Google instead petition for an outright exemption from the federal regulatory requirements so the company could continue to advance its deployments—which the company eventually succeeded in doing.

Despite the lackluster initial response, however, NHTSA eventually went on to produce guidance informing Google that its software system, in so many words, could in fact meet the definition of “driver” contemplated by the Federal Motor Vehicle Safety Standards. The agency noted:

If no human occupant of the vehicle can actually drive the vehicle, it is more reasonable to identify the driver as whatever (as opposed to whoever) *is* doing the driving. In this instance, an item of motor vehicle equipment, the SDS, is actually driving the vehicle.

This isn’t all that satisfying either, however. NHTSA seems to presuppose that a robot is in charge only if there isn’t a person available. But it still assumes a driver must be a person unless confronted with the impossibility of a human driver.

Numerous states have attempted to build upon NHTSA’s example by explicitly amending their vehicle codes to address the ambiguities identified by Google. But even these efforts have their quirks. California and Nevada, for example, now direct liability for traffic law violations toward the “operator” of the vehicle, defined as “the person who is seated in the driver’s seat, or if there is no person in the driver’s seat, causes the autonomous technology to engage.”<sup>233</sup> What happens, for example, when a passenger in a fully autonomous vehicle just wants to ride in the vehicle’s front-left seat (*i.e.* the driver’s seat)? Do they suddenly bear responsibility for an accident even when a passenger sitting right beside them doesn’t?

A different approach to revamping existing laws is simply to expand the universe of what is included. The FAA has taken this approach with unmanned aircraft. The blurring categories of aircraft emerging from the rapidly growing drone industry have forced the agency to rethink its old categories.<sup>234</sup> The title of Section 349 reads: “Exception for Limited Recreational Operations of Unmanned Aircraft.”<sup>235</sup> But, fearful of missing important new technologies, the FAA’s regulations now “refuse[] to set a threshold under which it bows out, insisting that everything not carrying people and capable of flight is an ‘unmanned aircraft.’”<sup>236</sup> Yet it’s not obvious that paper airplanes are something the FAA should worry about.

---

<sup>232</sup> *See id.* (quoting Letter from Paul A. Hemmersbaugh, Chief Counsel, NHTSA, to Chris Urmson, Director, Self-Driving Car Project, Google (Feb. 4, 2016), <http://isearch.nhtsa.gov/files/Google%20--%20compiled%20response%20to%2012%20Nov%20%2015%20interp%20request%20--%204%20Feb%2016%20final.htm>.) (quotations omitted).

<sup>233</sup> SB No. 1298, Chapter 570, SEC. 2. DIVISION 16.6. Autonomous Vehicles 38750.

<sup>234</sup> *See* David Schneider, *New FAA Rules for Drones Go Into Effect*, IEEE SPECTRUM (Oct. 10, 2018), <https://spectrum.ieee.org/automaton/robotics/drones/new-faa-rules-for-drones-go-into-effect>.

<sup>235</sup> *See id.*

<sup>236</sup> *See id.*

The placement of robots into either human or dumb machine regulatory boxes can also have more subtle effects, strengthening our subconscious views that devices which act human must have human motivations and limitations. In the wake of the fatal Uber accident in Arizona, for example, the police chief said that the vehicle couldn't possibly have seen the pedestrian because "she came out of the shadows."<sup>237</sup> He anthropomorphized a vehicle that doesn't see in the way same humans do. The vehicle's LIDAR, radar, and ultrasonic sensory systems can see objects in pitch blackness, so the comparison was irrelevant (more on this in Part III). But a legal system that assumes drivers are humans may naturally tend to impute human beliefs and limitations to robots, as the Arizona police chief did here.

### 3. *Siloed and Overlapping Definitions*

The definitional problem is further complicated by the expertise of the definer. We've seen that dictionaries, industry leaders, and legislators have already put together some pretty bad definitions of "robot." But surely expert agencies will do better, right? Some legislators hope so. The pending federal Bot Disclosure Bill sponsored by Senator Feinstein, for example, largely avoids defining bots and instead directs the "Federal Trade Commission to define that term 'broadly enough so that the definition is not limited to current technology.'"<sup>238</sup>

But deferring to "expert" agencies comes with its own set of tradeoffs. Such agencies are often "expert" only in their traditional domains, not in robotics. An agency charged with regulating healthcare professionals, for instance, may believe it only needs to appoint regulators with domain expertise over human healthcare providers. But thanks to rapid advances in machine diagnostic tools, these agencies might soon be regulating robots.<sup>239</sup> Without concerted efforts made to anticipate these trends, agencies may lack the technological expertise necessary to come up with good (or, maybe more accurately, less-bad) definitions.

The disruptive effects of robotics technologies may also push us to redefine regulatory institutions themselves. The technology's crosscutting nature increasingly means that single applications can implicate a staggering array of regulatory concerns. And "[a]s robots become more and more multi-purpose, it will be harder to imagine a priori how they will be used and, thus, harder to create comprehensive legislative and consumer protections for them."<sup>240</sup>

---

<sup>237</sup> See Timothy B. Lee, *Police Chief Said Uber Victim "Came From The Shadows"—Don't Believe It*, ARS TECHNICA (Mar. 23, 2018), <https://arstechnica.com/cars/2018/03/police-chief-said-uber-victim-came-from-the-shadows-dont-believe-it/>.

<sup>238</sup> See S.B. 3127, *supra* note \_\_.

<sup>239</sup> *AI vs. Doctors*, IEEE SPECTRUM, <https://spectrum.ieee.org/static/ai-vs-doctors> (surveilling the field of machine diagnostic tools); <https://yellrobot.com/china-surgeon-remote-surgery-5g-network/> (describing remote surgery performed by a doctor-robot combination); A. Michael Froomkin & Ian Kerr, *When AIs Outperform Doctors: The Dangers of a Tort-Induced Over-Reliance on Machine Learning and What (Not) to Do About It*, \_\_ ARIZ. L. REV. \_\_ (forthcoming 2019).

<sup>240</sup> Neil Richards & William Smart, *How Should the Law Think About Robots?*, WeRobot (2013) (manuscript at 12), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2263363](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2263363).

Autonomous vehicles provide an illustrative example. It was not so long ago that the targets of automotive regulation could be relied upon to keep to their proverbial lanes.<sup>241</sup> Drivers drove, wheels turned, manufacturers manufactured, taxis taxied, fleet dispatchers dispatched fleets, and so forth. But a new generation of “robotaxis” may implicate many of these regulatory concerns at once.<sup>242</sup> In a world with clear dividing lines between drivers, manufacturers, taxi providers, and delivery services, it may have made perfect sense to designate separate regulatory bodies for each. The Department of Motor Vehicles, for example, has historically been delegated responsibility for driver licensing. The National Highway Traffic and Safety Administration, meanwhile, has regulated manufacturing standards. And other agencies have regulated discrete transportation services like taxis and limousines. But in a world where these technologies confound traditional administrative mandates, all of that can change overnight. Emerging autonomous vehicle models will present regulatory concerns as widespread as congestion, road safety, wages, discrimination, grid usage, fuel efficiency, privacy, cybersecurity, unfair competition, and much more.

These crosscutting technological impacts may give rise to zero-sum battles over regulatory authority. There are, after all, lots of agencies out there. And robots that disrupt existing paradigms will likely cause significant hierarchical and jurisdictional conflict. A self-driving taxi service, for example, is potentially subject to federal regulation by NHTSA, by state departments of transportation, by local taxi agencies in every city, and by government bodies like airports that control access to important public lands, to name just a few. And to be successful, such services will need to operate in most if not all major markets, so there is a significant problem of horizontal as well as vertical overlap. If every city, state, and taxi commission gets to define self-driving cars in their own way, there’s the real risk of a patchwork of inconsistent definitions that would prove difficult, if not impossible, to navigate.

But there’s also reason to think that the crosscutting nature of robots could open new doors to regulatory cooperation. Indeed, effective regulation will depend on the industry- and jurisdiction-specific characteristics of robots. Self-driving cars, for example, present very different issues than medical robots. And medical robots deployed in rural hospitals will implicate distinct policy concerns that those in urban environments won’t. So while we may want industry-specific regulators to know something about the robots they’ll increasingly encounter, that doesn’t mean we’ll want a general-purpose regulator of robots to oversee them.<sup>243</sup> The right balance will combine sector- or jurisdiction-specific expertise with a deep understanding of the new issues robots will bring to that sector.<sup>244</sup> And that means that emerging robotics technologies could foster collaborations of a type the regulatory state has scarcely witnessed before—leading

---

<sup>241</sup> So to speak.

<sup>242</sup> See *infra* notes \_\_\_ – \_\_\_ and accompanying text.

<sup>243</sup> *But see* RYAN CALO, THE CASE FOR A FEDERAL ROBOTICS COMMISSION, BROOKINGS INST. CENT. FOR TECH. INNOVATION (2014); (outlining the merits of such an agency).

<sup>244</sup> The 2018 AI Now Report, for instance, calls for sector-specific, not general, regulation of robotics. MEREDITH WHITTAKER ET AL., AI NOW REPORT 2018, AI NOW (2018), [https://ainowinstitute.org/AI\\_Now\\_2018\\_Report.pdf](https://ainowinstitute.org/AI_Now_2018_Report.pdf).

regulators to branch out far from their traditional domains, work in partnership with private entities, or develop overlapping, complementary forms of oversight.<sup>245</sup>

It's also critical to note that not all robot regulation will come from the public sector. Some regulation may result from partnerships between public and private bodies. Further, companies, themselves, will also act as self-regulators—meaning they'll come up with their own policies and, therefore, definitions in order to decide how to deal with bots in their own businesses. Such policies for defining robots may, in turn, become *de facto* law as they speed ahead of state and federal legislators, administrators, and judicial bodies (who will have much on their plates). Twitter, Instagram, and Facebook for instance, are in the midst of bot purges.<sup>246</sup> To do so, they're effectively adopting a situational, implicit definition of bots.<sup>247</sup> These industry-government collaborations can produce good outcomes, as Kaminski argues, but they can also lead to capture and preclude effective regulation.<sup>248</sup>

The problem of defining robots, in short, is complicated by both the fact that lots of different agencies are likely to be offering definitions, and the fact that those agencies may have domain expertise but not expertise in robotics.

#### 4. *Regulating on Bot Time*

The difficulties associated with each failure mode above are further compounded by another problem: the pace of change in robotics and AI. As little as three decades ago, time could be said to be on the regulators' side. Technologies entered markets and regulators' lives, but at what looked (at least by modern standards) to be a measured pace. Then came what Larry Lessig called "internet time"—accelerating the pace of innovation to a degree that pushed policymakers to the breaking point. Now, it appears time is again shifting gears. As Moore's Law, Kyder's law,<sup>249</sup>

---

<sup>245</sup> Margot E. Kaminski, *Binary Governance: A Two-Part Approach to Accountable Algorithms*, 92 S. CAL. L. REV. \_\_ (forthcoming 2019).

<sup>246</sup> See *Constine*, *supra* note \_\_.

<sup>247</sup> See *id.*

<sup>248</sup> Among the considerable literature on agency capture, see, e.g., Stephen G. Breyer, *Regulation and Its Reform* (1994); 1A Areeda & Hovenkamp, *Antitrust Law* para. 241b2; James Buchanan & Gordon Tullock, *The Calculus of Consent* (1962); Daniel A. Farber & Philip P. Frickey, *Law and Public Choice, a Critical Introduction* (1991); George J. Stigler, *The Theory of Economic Regulation*, 2 BELL J. ECON. & MGMT. SCI. 3 (1971); Richard A. Posner, *Theories of Economic Regulation*, 5 BELL J. ECON. & MGMT. SCI. 335 (1974).

Microsoft, for instance, has called for government regulation of facial recognition AIs. [https://blogs.microsoft.com/on-the-issues/2018s/12/06/facial-recognition-its-time-for-action/?utm\\_source=t.co&utm\\_medium=referral](https://blogs.microsoft.com/on-the-issues/2018s/12/06/facial-recognition-its-time-for-action/?utm_source=t.co&utm_medium=referral). The AI Now Report also calls for regulation of facial recognition software. <https://medium.com/@AINowInstitute/after-a-year-of-tech-scandals-our-10-recommendations-for-ai-95b3b2c5e5>

<sup>249</sup> This law refers to the exponential growth of computer storage capacity (as opposed to Moore's law which described processing capacity).

and the only half-joking “Huang’s Law”<sup>250</sup> continue their inexorable march,<sup>251</sup> all signs suggest we’re moving from “internet time” to what might rightly be called “robot time” (or, perhaps, “AI time”).<sup>252</sup> Technologies of profound social, legal, and economic complexity are arriving at a speedier clip than ever. And that means definitions have a shorter shelf-life than ever and regulators have less time to get up to speed.

If current trends in judicial interpretation are any indication, we should be acutely concerned about this acceleration. Given the shift in statutory interpretation towards stilted, “originalist” and “plain meaning” readings of laws,<sup>253</sup> we seem especially apt to lock in definitions that ignore the context or intent of legislation.<sup>254</sup> If courts do no more than read the words for their supposed plain meaning at the time a statute was adopted, robot definitions are unlikely to be read in ways that are responsive to unforeseen changes in technology.

Absent a new approach, “definability” as we currently know it may not withstand this transition. As we move from “internet time” to a timescale exponentially faster, something will have to give. And one thing’s certain: it won’t be an attempt to regulate. Sure, the ideas that framed the first generation of cyberlaw centered on the Internet’s inability to be tamed.<sup>255</sup> And, today, we hear similar rumblings regarding robots. But the idealism of a regulation-free Internet didn’t last.<sup>256</sup> Regulation came—the good, the bad, and the ugly. It will come for robots too. And if we don’t develop more sophisticated approaches to defining technologies like driverless vehicles, autonomous drones, and deep fakes, we’ll be in deep trouble.

---

<sup>250</sup> See Tekla S. Perry, *Move Over Moore’s Law, Make Way for Huang’s Law*, IEEE SPECTRUM (Apr. 2, 2018), <https://spectrum.ieee.org/view-from-the-valley/computing/hardware/move-over-moores-law-make-way-for-huang-s-law> (this refers to the exponential growth of GPU processing power, a specific type of processing well-suited for modern AI approaches).

<sup>251</sup> Even if some have showed signs of recent slowing.

<sup>252</sup> See Lessig, *supra* note \_\_\_\_.

<sup>253</sup> See, e.g., WILLIAM ESKRIDGE ET AL., LEGISLATION AND STATUTORY INTERPRETATION (2d ed.).

<sup>254</sup> In *SAS Inst. v. Iancu*, for instance, the Court adopted a nonsensical reading of the America Invents Act based on its purported “plain meaning.” 584 U.S. \_\_\_\_ (2018).

<sup>255</sup> As Lessig wrote:

If there was a meme that ruled talk about cyberspace, it was that cyberspace was a place that could not be regulated. That it “cannot be governed”; that its “nature” is to resist regulation.

See Lessig, *supra* note \_\_\_\_ . A certain sitting president once famously likened it to “trying to nail Jello to the wall.” Bethany Allen-Ebrahimian, *The Man Who Nailed Jello to the Wall*, Foreign Pol. (Jun. 29, 2016), <https://foreignpolicy.com/2016/06/29/the-man-who-nailed-jello-to-the-wall-lu-wei-china-internet-czar-learns-how-to-tame-the-web/>. See also John Perry Barlow, *A Declaration of Independence for Cyberspace*, WIRED (1994); David Johnson & David Post, *Law and Borders: The Rise of Law in Cyberspace*, 46 STAN. L. REV. 1367 (1996).

<sup>256</sup> Lessig, *supra* note \_\_\_\_.

### III. Sorting the Bots from the Nots

Hopefully we've now persuaded you that defining robots is hard, and that so far we haven't done a very good job of it. By tradition, this is the part of a law review article where we are supposed to unveil our own, much better definition, which neatly solves all the problems we identified in the previous parts and ties everything up with a bow.

If that's what you're expecting, well, get used to disappointment.<sup>257</sup> We don't have a magic definition that will make everything fine. In fact, we don't think such a definition is possible. Instead, mitigation is our only hope. And in the subparts that follow, we offer strategies designed to accomplish precisely that.

#### A. Turing's Razor<sup>258</sup>

Our first suggestion is to avoid defining robots in categorical terms whenever possible. Policymakers should resist the idea of *ex ante*, explicit definition as much as possible. As noted above, that doesn't mean they shouldn't regulate. Maybe they should, maybe they shouldn't. But we know they will. Rather, when regulating, policymakers should first ask a threshold question: To define or not to define? Counterintuitively, not defining is sometimes the better course of action. Courts take this approach in the common law, building definitions inductively over time by deciding cases that present boundary issues.

We're not suggesting that the law should never define robots. Any regulation that governs robots will cover some devices and exclude others. What we suggest instead is what we might call "Turing's Razor": Whenever possible, establish whether a potential regulated entity is a robot without resorting to explicit, *ex ante* definitions. As Turing shows, purposeful omission can still allow for definition. Turing didn't define "thinking machines."<sup>259</sup> He offered a criterion for deciding whether something was an AI. For Turing, something is intelligent if it behaves in a way indistinguishable from the way intelligences behave.

Turing's Razor points us toward functional definitions of robots. But it has a second, larger implication: Whenever possible we should regulate verbs, not nouns. That is, what the law should generally target is the act that is being regulated as opposed to the actor. A good example of this approach is the Better Online Ticket Sales Act of 2016 (aka "BOTS Act"). The Act makes no attempt to define bot. Instead, it simply prohibits efforts to get around security protocols like CAPTCHA.<sup>260</sup> We don't actually need to decide whether you are a bot. As the BOTS Act demonstrates, we can achieve our goals by deciding whether someone (or something) is circumventing the protocol.

---

<sup>257</sup> Yep, *The Princess Bride* again.

<sup>258</sup> A razor, in philosophy, refers to a rule of thumb that allows us to avoid unnecessary actions or proverbially "shave off" unlikely explanations.

<sup>259</sup> See *supra* notes \_\_\_ – \_\_\_ and accompanying text.

<sup>260</sup> See Better Online Ticket Sales Act of 2016, 114th Cong., (2015-2016).

Similarly, it will often be true that we don't need rules that decide whether a car with certain autonomous features is or is not a robot. What we actually need are rules that regulate unsafe driving behavior, regardless of whether a human or a machine is engaging in it. That approach won't always work; we may need to decide whether it is safe to sit behind the wheel of a car while drunk, and that may well depend on how autonomously the car can act and where it is going. But even there, the underlying regulatory concern involves functional considerations, not the entity itself. By focusing on what we need to know in order to justify a particular regulation, we can regulate no more than we need to and avoid baking in poor definitions that are over- or underinclusive (or likely to rapidly become obsolete). If the question is "Can the car drive an impaired person home?" we may reach a different conclusion than if the question is "Does the car need a steering wheel?"

A focus on conduct, not status, is a good idea for other reasons. It may help us avoid discrimination against certain technologies or business models, and ultimately avoid discrimination against robots. It will allow us to accumulate knowledge and hone our definitions over time by giving us the flexibility to change course as the technology changes. And ultimately it may prevent unnecessary regulation by narrowing our legal rules to focus on identified problems rather than creating regulations that apply across the board to robots, whether we need them or not.

It may seem strange to suggest that we regulate robots without any definition of what a robot even is. But it's not unprecedented. Most famously, the Supreme Court applied this approach to "obscenity."<sup>261</sup> Justice Stewart didn't actually define the term in the course of regulating it. Instead, he offered a phrase that has since etched itself into the legal firmament: "I know it when I see it."<sup>262</sup> As it turns out, both robots and pornography share this central challenge of *Jacobellis*. They're both really, *really* tough to define.

True, that approach leaves us with some uncertainty. But any definition specifying what was obscene *ex ante* would almost certainly have been worse. The sorts of things at issue in the 1950s and 1960s look quite tame by modern standards. A statutory or precedential definition of obscenity set in that era would be far more restrictive than what seems appropriate today. We suspect the same will be true of robots.

Our first suggestion, therefore, is to follow Turing's Razor: don't say any more than you have to, if that.<sup>263</sup>

## **B. Defining Functionally**

Regulating verbs not nouns is easier said than done. Regulators following the Turing's Razor approach will want to identify the most salient functional features of robots. Here, we think at least six criteria merit consideration. And, conveniently for fans of mnemonic devices, they all

---

<sup>261</sup> *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964).

<sup>262</sup> *See id.* (Stewart, J., concurring) ("I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it.").

<sup>263</sup> Chili Palmer, in *Get Shorty*.

start with the letter “a.” They are: (1) agenda, (2) automaticity, (3) autonomy, (4) agency, (5) ability, and (6) anthropomorphization.

In offering these “six As,” however, we have no illusions as to their limitations. We make no pretense that they provide a definitive, one-size-fits-all methodology for categorizing the relevant regulatory criteria of robots. Indeed, part of our point is that there *is* no such methodology. Rather, our hope is that focusing regulators’ attention on each discrete consideration will encourage the right type of thinking—allowing them to anticipate and mitigate some of the definitional pitfalls we’ve seen in the past. The following sections explore each of these “six As,” offering real world examples of their successes and failures.

## 1. Agenda

We begin by considering “agenda”: that is, the motives held by those deploying robots, which presumably dictate the ends robots will serve. Paying attention to agenda is critical because not all robots seek to advance the same ends. Some, for example, are clearly malicious. They’re deployed to hack, deceive, fraudulently impersonate, or gain unfair competitive advantages.<sup>264</sup> But plenty of others operate in service of vital public interests, or simply make life better or safer for their users or customers. Failing to recognize that robots are “dual use” technologies that can enable positive or negative actors can result in legal rules with severe unintended consequences.

The Computer Fraud and Abuse Act (CFAA) is illustrative. The statute prohibits actors from “intentionally access[ing] a computer without authorization,” among other restrictions.<sup>265</sup> The CFAA was originally intended to target Cold War-era hackers.<sup>266</sup> But, unfortunately, its prohibitions make no distinctions between ill-intentioned hackers and individuals operating in the public interest. The statute’s failure to distinguish between good and bad actors has had a chilling effect on those deploying bots in the public interest. It has been invoked against “white hat” security researchers, journalists seeking to measure discrimination on Facebook, and academics trying to safeguard our voting system.<sup>267</sup> As these suits show, conduct that resembles hacking is sometimes necessary to advance a universally recognized good. With robots, too, our regulations need to be sensitive to the fact that failing to consider the question of agenda could stymie bad *and* good uses alike.<sup>268</sup>

---

<sup>264</sup> See *supra* notes \_\_\_ - \_\_\_.

<sup>265</sup> 18 U.S.C § 1030. See *supra* Part II.B.

<sup>266</sup> See Kerr, *Norms of Computer Trespass*, *supra* note \_\_\_.

<sup>267</sup> See, e.g., Jamie Williams, *D.C. Court: Accessing Public Information is Not a Computer Crime*, EFF (Apr. 12, 2018), <https://www.eff.org/deeplinks/2018/04/dc-court-accessing-public-information-not-computer-crime>.

<sup>268</sup> Madeline Lamo and Ryan Calo, for example, get at the notion of “agency” by distinguishing between commercial, corporate, and political bots in their recent discussion on robot free speech. Lamo & Calo, *supra* note \_\_\_ (discussing “the various ways that people can use bots for personal, corporate, and political ends, where questions of social impact are front and center.”).

By way of contrasting example, regulators should consider the Federal Trade Commission's (FTC) standard of unfair and deceptive business practices. Crucially, the standard is technologically agnostic. It focuses not on actors, *per se*, but on their motives.<sup>269</sup> So when it came to light that Ashley Madison—an app designed to facilitate extramarital affairs<sup>270</sup>—was inducing online purchases “by creating fake computer ‘hosts’ or ‘bots,’ which were programmed to generate and send messages to male members under the guise that they were real women,”<sup>271</sup> the FTC didn't get bogged down in technical esoterica.<sup>272</sup> Rather, the agency needed only to consider the agenda of those who had deployed the bots (in this instance, whether they had been used toward unfair and deceptive ends).

## 2. Automaticity

A threshold consideration facing any robot regulator is whether their rules should apply only to fully-automated systems or also extend to robots that require some level of human intervention. Ultimately, this is a question of “automaticity”: that is, the degree to which a robot accomplishes discrete tasks without humans intervening. As we've seen, many robotics technologies are not fully “autonomous” but are, in Jonathan Zittrain's pithy phrasing, “autonomish.”<sup>273</sup> They exist somewhere in the twilight between humans and fully-automated systems.<sup>274</sup> And sometimes, as Ford's example in the introduction shows, they're actually just people dressed as car seats.<sup>275</sup> Unfortunately, a “surprising” number of journalists, researchers, and industry experts fail to make these crucial distinctions between autonomous and autonomish bots.<sup>276</sup> And this, in turn, can lead to significant confusion among policymakers.

Automaticity won't matter for all robot regulations. We might, for example, not care that the driver of a vehicle is a robot if we're convinced it's safe. But there will likely be a host of regulations for which clarifying this threshold question is vital. A recent string of controversies

---

<sup>269</sup> See Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 812-14 (2015).

<sup>270</sup> The company's slogan is: "Life is Short. Have an Affair."

<sup>271</sup> Article: Artificial Intelligence Litigation: Can the Law Keep Pace with The Rise of the Machines?, Quinn, Emmanuel, Urquhart, & Sullivan LLP, <https://www.quinnemanuel.com/the-firm/publications/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>.

<sup>272</sup> *In re Ashley Madison Customer Data Sec. Breach Litig.*, 148 F. Supp. 3d 1378, 1380 (JPML 2015).

<sup>273</sup> Jonathan Zittrain (@zittrain), TWITTER (Mar. 19, 2018 3:34 PM) <https://twitter.com/zittrain/status/975870939811516416>.

<sup>274</sup> Or, as Part I.D shows, they may even be cyborgs.

<sup>275</sup> See Marshall, *That Guy Dressed Up as a Car Seat to Solve a Robocar Riddle*, *supra* note \_\_.

<sup>276</sup> See Gorwa & Guilbeault, *supra* note \_\_ (noting that a “surprising number of journalists and researchers describe human-controlled accounts as bots”).

involving “pseudo-AI”<sup>277</sup> companies is illustrative. Today, hundreds of companies claim to offer robots capable of making sense of users’ communications.<sup>278</sup> But as it turns out, many such companies conceal the fact that their so-called “bots” actually rely on humans to get the job done.<sup>279</sup> In some situations, this “Wizard of Oz design technique”<sup>280</sup> can be relatively innocuous. It may, for example, matter little to users that the chatbots they’re purportedly interacting with are, in fact, humans (as we recently saw with Facebook, X.ai, Clara, and a host of similar applications).<sup>281</sup> But for other business models, such as that of Edison Software, the question of human involvement can change the regulatory complexion entirely. Edison Software recently made headlines after offering its users the ability to automate intelligent email replies.<sup>282</sup> But a journalistic investigation revealed that its “AI technology” actually entailed engineers manually going “through the personal email messages of hundreds of users.”<sup>283</sup> To no one’s great surprise, the company conveniently failed to “mention that humans would view users’ emails in its privacy policy.”<sup>284</sup> And, far from unique, Edison is just one of dozens of others that have flouted fake automaticity as a way to dodge oversight concerns.<sup>285</sup>

Automaticity is fundamentally different from the question of whether we are, in fact, regulating a human being or a machine. If we don’t care—*i.e.* if we regulate behavior regardless of who or what performs it—it is less important. But if we are to treat automatic behavior more (or less) favorably than human actions, we ought to start by establishing whether the putative robot is really acting automatically.

### 3. Autonomy

Automaticity also affects how we think about a third critical consideration: “autonomy.” Purely automated systems that cannot deviate from the dictates of their designers may well be “automatic,” but they don’t have autonomy. By autonomy, we mean the extent to which an entity is empowered to make decisions. As technical advances increasingly allow robots to learn

---

<sup>277</sup> See Olivia Solon, *The Rise of 'Pseudo-AI': How Tech Firms Quietly Use Humans to do Bots' Work*, THE GUARDIAN (Jul. 6, 2018), <https://www.theguardian.com/technology/2018/jul/06/artificial-intelligence-ai-humans-bots-tech-companies>.

<sup>278</sup> See *id.*

<sup>279</sup> See *id.*

<sup>280</sup> See *id.* (quoting Alison Darcy, founder of Woebot).

<sup>281</sup> See *id.* (discussing Facebook’s reliance “on humans for its virtual assistant for Messenger, M”).

<sup>282</sup> See *id.*

<sup>283</sup> See *id.*

<sup>284</sup> See *id.*

<sup>285</sup> See, e.g., *id.*; Richard Wray, *SpinVox Answers BBC Allegations Over Use of Humans Rather Than Machines*, THE GUARDIAN (Jul. 23, 2009), <https://www.theguardian.com/business/2009/jul/23/spinvox-answer-back>; Ellen Huet, *The Humans Hiding Behind the Chatbots*, BLOOMBERG (Apr. 18, 2016), <https://www.bloomberg.com/news/articles/2016-04-18/the-humans-hiding-behind-the-chatbots>.

novel behaviors from unique environmental inputs, the clean dividing lines of automaticity and autonomy—even for fully-automated systems—are beginning to blur.<sup>286</sup> Modern machine learning systems, for instance, are given ultimate goals by a human controller, but can have considerable leeway in how they implement those goals (or in setting subsidiary goals that further the ultimate purpose). Here, “as legions of genies in bottles have taught us . . . [we need to] be careful what [we] wish for.”<sup>287</sup> For autonomous robots will pursue the ends they’re told to, even when doing so means subverting the ultimate purpose behind that goal. Today, examples of autonomous robots gone awry are a regular fixture of the news.<sup>288</sup> It is this very phenomenon that explains why drones powered by machine learning algorithms recently began flying in the wrong direction to take advantage of an inadvertent loophole in their programming, and also why a map-making AI decided it was easier to actually hide the fact that it was storing a full picture rather than abstracting elements of that picture into maps.<sup>289</sup>

#### 4. Agency

Autonomy implicates a fourth concern: “agency.” Those who regulate robots must confront the question of who or what is ultimately being held responsible for the robot’s conduct. Automatic systems that can only do what they’re told are relatively easy cases for assigning liability. But it is less obvious who should be responsible for the behavior of autonomous systems that depart from their designer’s intent in significant ways.<sup>290</sup> Consider the now notorious example of Microsoft’s Twitter chatbot, “Tay.” Unlike many other bots “designed to maintain a static internal state upon deployment, Tay’s system updated itself in real time by learning from interactions with users.”<sup>291</sup> Thus, when Microsoft first deployed Tay into the Twitterverse, it was a proverbial blank slate. But after a coordinated effort by trolls to tweet “misogynistic, racist, and Donald Trumpist remarks” at the bot, all that changed.<sup>292</sup> In mere hours, “Tay rapidly morphed

---

<sup>286</sup> See Ryan Abbott, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability* 86 Geo. Wash. L. Rev. 1 (2018) (arguing “[c]omputers are no longer just inert tools directed by individuals . . . [r]ather, in at least some instances, computers are given tasks to complete and determine for themselves how to complete those tasks”); Ryan Calo, *Robots in American Law* (2016), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2737598](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2737598) (noting “[i]f ever there were a line between human and robot spontaneity or skill, it is rapidly disappearing”).

<sup>287</sup> See Lemley & Casey, *supra* note \_\_.

<sup>288</sup> See *id.* (discussing examples of autonomous systems inadvertently betraying their programmer’s intent).

<sup>289</sup> See *id.*

<sup>290</sup> We discuss this issue in detail in a previous work. See Lemley & Casey, *supra* note \_\_ and sources cited therein.

<sup>291</sup> See *id.* at \_\_ (quoting James Vincent, *Twitter Taught Microsoft’s AI Chatbot to be a Racist Asshole in Less Than a Day*, THE VERGE (Mar. 24, 2016), <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>).

<sup>292</sup> See Vincent, *supra* note \_\_.

from a fun-loving bot . . . into an AI monster.”<sup>293</sup> But the question of who held responsibility for the bot’s dramatic transformation was less clear. Given that Tay’s seemingly rogue behavior resulted from the input of third parties, it’s at least plausible that the bot was no longer acting as an “agent” of Microsoft in a meaningful sense.<sup>294</sup>

Deciding the question of agency can be hard enough for systems that are fully automated. But things get even harder when thinking about autonomish systems. Sometimes autonomish systems can actually enhance the ability of humans to exert agency rather than taking control themselves. With unmanned military aircraft, for example, automated systems operating in the background can now handle more granular flight tasks, such as automatic stabilization.<sup>295</sup> And that, in turn, allows human pilots to dedicate even more attention to high-stakes questions, such as whether to fire at a target.<sup>296</sup>

But autonomish systems can also muddle questions of agency in complex robotics applications. Take autonomous vehicles. Automating the “dynamic driving task” can induce complacency in humans tasked with monitoring self-driving vehicles. And as we saw in Tempe, Arizona, this phenomenon can occasion horrific outcomes. In February of 2018, an autonomous vehicle deployed by Uber was travelling at about 35 mph when a woman stepped in front of it.<sup>297</sup> Neither the vehicle’s sensory system nor the “safety driver” on board spotted her in time to prevent the fatal collision.<sup>298</sup> And in its aftermath, the question of agency was a fraught one. Police initially blamed the pedestrian.<sup>299</sup> But the Arizonan government appeared to disagree, banning Uber’s autonomous vehicle program from the entire state.<sup>300</sup> Uber itself settled with the pedestrian’s estate.<sup>301</sup> And, all the while, the so-called “safety driver” at the heart of the event

---

<sup>293</sup> See Rachel Metz, *Microsoft’s Neo-Nazi Sexbot was a Great Lesson For Makers of AI Assistants*, MIT TECH. REV. (Mar. 27, 2018), <https://www.technologyreview.com/s/610634/microsofts-neo-nazi-sexbot-was-a-great-lesson-for-makers-of-ai-assistants/>.

<sup>294</sup> This question may, in turn, hinge on foreseeability. For example, if Microsoft were to make this same mistake twice, disclaimers that centered on agency would be less convincing.

<sup>295</sup> Cf Richard Water, *Artificial Intelligence: When Humans Coexist with Robots*, FIN. TIMES (Oct. 3 2018) (showing examples of “how AI can make humans far more effective without robbing them of control”).

<sup>296</sup> See *id.*

<sup>297</sup> See Daisuke Wakabayashi, *Emergency Braking Was Disabled When Self-Driving Uber Killed Woman*, Report Says, N.Y. TIMES (MAY 24, 2018), <https://www.nytimes.com/2018/05/24/technology/uber-autonomous-car-ntsb-investigation.html>.

<sup>298</sup> See *id.*

<sup>299</sup> See *id.*

<sup>300</sup> See Eric Newcomer, *Uber’s Autonomous Cars Banned in Arizona After Fatal Crash*, BLOOMBERG (Mar. 26, 2018), <https://www.bloomberg.com/news/articles/2018-03-27/uber-s-autonomous-cars-suspended-by-arizona-after-fatal-crash>.

<sup>301</sup> See Faiz Siddiqui, *Uber Reaches Settlement with Family of Victim Killed After Being Struck by One of Its Self-Driving Vehicles*, WASH. POST (Mar. 29, 2018), <https://www.washingtonpost.com/news/dr->

stood in the shadow of criminal liability—having been nominally responsible for preventing the accident despite being thrust into a position that study after study has revealed to be untenable: controlling an agent that is largely autonomous.

## 5. Ability

Where “agenda” asks what the desired end goal is, “ability” asks how a robot (or human) achieves that goal. Focusing on abilities, as opposed to who or what has those abilities, helps to avoid legal rules that create needless or inadvertent barriers to new technologies. Regulations out of Rwanda involving unmanned aerial vehicles (UAVs) provide a positive example. The country has drawn praise from the White House and two Secretaries of Transportation for its focus on performance-based legal rules for UAVs.<sup>302</sup> By establishing safety targets based on UAV abilities, the government allows companies to choose different “combination[s] of technologies and operational mitigations” they need to get the job done. In Lisa Ellman’s telling, “This gives the drone operators flexibility to build drones around performance [capabilities]” instead of “prescriptively requiring a certain type of technology.”<sup>303</sup> Ellman notes, for example, that, “Rather than the government saying ‘you have to use this kind of technology to stop your drone,’ they would say, ‘your drone needs to be able to stop in so many seconds.’”

The benefit of focusing on ability is not limited to drones.<sup>304</sup> Defining terms like “drivers” and “operators” by explicitly referencing the human sitting in the driver’s seat can inadvertently preclude robots. But there’s no reason why self-driving cars that are safer than human drivers should be banned simply because they’re not made of flesh and blood. Max Tegmark refers to

---

gridlock/wp/2018/03/29/uber-reaches-settlement-with-family-of-victim-killed-after-being-struck-by-one-of-its-self-driving-vehicles/?utm\_term=.a1926f1eda34.

<sup>302</sup> See Jake Bright & Samantha Stein, *African Experiments with Drone Technologies Could Leapfrog Decades of Infrastructure Neglect*, TECH CRUNCH (Sept. 16, 2018), <https://techcrunch.com/2018/09/16/african-experiments-with-drone-technologies-could-leapfrog-decades-of-infrastructure-neglect/>.

<sup>303</sup> See *id.*

<sup>304</sup> *Motherboard*, *supra* note \_\_: “The question of what a robot is and what that means has come up more often than you might think. In *Comptroller of the Treasury v. Family Entertainment Centers*, a special appeals court had to decide whether Chuck E. Cheese animatronic robots were considered “performers.” This distinction mattered because at the time, Maryland taxed food differently “where there is furnished a performance.” The court decided that “a pre-programmed robot can perform a menial task but, because a pre-programmed robot has no ‘skill’ and therefore leaves no room for spontaneous human flaw in an exhibition, it cannot ‘perform’ a piece of music ... just as a wind-up toy does not perform for purposes of [the law] neither does a pre-programmed mechanical robot.” For the purposes of that individual case, that seems like a fairly innocuous definition. But Calo notes that even at the time of this case, robots had begun performing basic autonomous tasks. “If ever there were a line between human and robot spontaneity or skill, it is rapidly disappearing,” he wrote.”

Calo, *supra* note \_\_, collects other good examples, including whether a submersible robot can claim discovery and ownership of maritime salvage.

this tendency to presume that intelligent entities “must be made of meat” as “carbon chauvinism.”<sup>305</sup> And it’s a pitfall that focusing on ability helps us to avoid.

But just because robots are capable of rivalling humans in increasingly diverse contexts doesn’t mean that all comparisons between human and bot abilities are equally valid. Rather, regulators should also be sensitive to instances where analogies break down. This could be as simple as acknowledging that one billion miles successfully driven by an autonomous vehicle in simulation are not necessarily comparable to one billion miles driven in the real world.<sup>306</sup> But it can also involve less direct analogies. Trucking laws can restrict the number of consecutive hours that commercial vehicles operate for.<sup>307</sup> The purpose of those laws is to reduce fatigue-related accidents caused by tired truckers. But because the laws don’t target the sleeping habits of the truckers directly, and instead promote healthy sleep through vehicle operation limits, they make little sense as applied to robots that don’t require sleep.

A focus on abilities is closely related to our suggestion in the last part that laws should regulate behavior, not status. But it can also be useful in defining robots for purposes of legislation. A trucking law that took into account whether a truck operator (either robot or human) required sleep—rather than assuming drivers couldn’t operate 24 hours a day—would focus on ability, not make assumptions about human abilities and limitations.

## 6. Anthropomorphization

So far, our functional criteria have largely focused on robotics technologies themselves. But when regulating robots as verbs, not nouns, it’s also important to also turn the tables by considering how humans will react to robots. As Ryan Calo and others have observed, how we regulate “automated systems— especially robots—[is] significantly shaped by the metaphors, stories, and names that we use to describe them.”<sup>308</sup> And our tendency to anthropomorphize them can lead to some bewildering outcomes.

Sometimes anthropomorphizing can cause us to describe robot abilities in ways that create more confusion than clarity. Researchers, journalists, and industry experts, for example, regularly “characterize modern image classifiers as ‘surpassing human abilities and effectively

---

<sup>305</sup> Eliza Strickland, *Interview: Max Tegmark on Superintelligent AI, Cosmic Apocalypse, and Life 3.0*, IEEE SPECTRUM (Sept. 14, 2017), <https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/interview-max-tegmark-on-superintelligent-ai-cosmic-apocalypse-and-life-3-0>.

<sup>306</sup> Cf. Fred Lambert, *Tesla’s Fleet Has Accumulated over 1.2 Billion Miles on Autopilot and Even More in ‘Shadow Mode’, Report Says*, ELECTREK (Jul. 17, 2018), [Error! Hyperlink reference not valid.https://electrek.co/2018/07/17/tesla-autopilot-miles-shadow-mode-report/](https://electrek.co/2018/07/17/tesla-autopilot-miles-shadow-mode-report/) (showing Tesla reports over 1 billion miles driven with Autopilot, but most of those are in better driving conditions like freeways).

<sup>307</sup> Nick Stockton, *Truckers Take on Trump Over Electronic Logging Device Rules*, WIRED (Apr. 3, 2018), <https://www.wired.com/story/trump-truckers-hours-service-electronic-logging-device/>. The same is true of doctors. See, e.g., James Hamblin, *No Doctor Should Work 30 Straight Hours Without Sleep*, THE ATLANTIC (Dec. 12, 2016), <https://www.theatlantic.com/health/archive/2016/12/no-doctor-should-work-30-straight-hours/510395/>.

<sup>308</sup> See Calo, *Robotics and the Lessons of Cyberlaw*, *supra* note \_\_\_\_.

proving that bigger data leads to better decisions,’ despite demonstrations that these networks rely on spurious correlations, e.g. misclassifying ‘Asians dressed in red’ as ping-pong balls.”<sup>309</sup> And even when claims are careful to provide caveats, loose comparisons to humans can nevertheless “portray a false sense of current abilities.”<sup>310</sup> A recent empirical result claiming to have achieved “dermatologist-level classification of skin cancer,” for example, underplays the fact “that classifiers and dermatologists perform fundamentally different tasks.”<sup>311</sup> Dermatologists in real clinical environments “encounter a wide variety of circumstances and must perform their jobs despite unpredictable changes.”

Robots in the medical realm are far from the only offenders. Our instinct to anthropomorphize also “plagues many [other] subfields of [AI].”<sup>312</sup> A growing literature on fairness in AI, for example, “often overloads terminology borrowed from complex legal doctrine, such as *disparate impact*, to name simple equations expressing particular notions of statistical parity.” In Zachary Lipton’s and Jacob Steinhardt’s telling, “This has resulted in a literature where ‘fairness,’ ‘opportunity,’ and ‘discrimination’ denote simple statistics of predictive models, confusing researchers who become oblivious to the difference, and policymakers who become misinformed about the ease of incorporating ethical desiderata into [AI].”<sup>313</sup>

Much of this is caused by our natural inclination to react to anything that appears to display complex behavior as if it were human. We anthropomorphize our pets and even our stuffed animals, so maybe it’s no surprise that we treat our Roombas as members of the family. That instinct will only deepen as robots start to look more human (or at least cuter) and mimic human behavior.<sup>314</sup> If we subconsciously expect a robot to act like a human being, we will be surprised and upset when it doesn’t, or when it makes mistakes that seem bizarre to us.

But the problems of anthropomorphizing don’t end there. Rather, as a recent uproar over a “robot sex brothel” in Houston shows, things can get even stranger. It all began when a Canadian company operating an “adult love dolls rent before you buy service” decided to open a

---

<sup>309</sup> Zachary C. Lipton & Jacob Steinhardt, *Troubling Trends in Machine Learning Scholarship* (forthcoming 2019).

<sup>310</sup> *Id.* This includes our tendency to set performance standards at the human level. It’s fleeting and unhelpful.

<sup>311</sup> See Lipton & Steinhardt *supra* note \_\_\_ at \_\_\_.

<sup>312</sup> See *id.* at \_\_\_.

<sup>313</sup> See *id.*

<sup>314</sup> Ryan Calo, *People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship*, 114 PENN ST. L. REV. 809 (2010) (citing studies showing that depicting a picture of eyes changes our assumptions about privacy); Kate Darling, “‘Who’s Johnny?’ Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy,” ROBOT ETHICS 2.0, eds. P. Lin, G. Bekey, K. Abney, R. Jenkins, Oxford University Press (2017); Kate Darling, *Extending Legal Protections to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects*, in: ROBOT LAW, eds. M. Froomkin, R. Calo, I. Kerr, Edward Elgar (2016).

storefront in Houston, Texas.<sup>315</sup> The company, Kinky S Dolls, “sells realistic-looking life-size dolls with basic artificial intelligence functions.”<sup>316</sup> And, among other offerings, it lets customers rent the dolls “by the half-hour or hour” at its place of business.<sup>317</sup>

When Kinky S Dolls began setting up shop in Houston, there didn’t “appear to be any local, state or federal laws specifically banning ‘robot brothels.’”<sup>318</sup> But after a petition titled “Keep Robot Brothels Out of Houston” amassed more than twelve thousand signatures (not to mention national publicity), policymakers set about changing that.<sup>319</sup> On October 3, 2018, the Houston City Council passed an ordinance effectively banning businesses like Kinky S Dolls.<sup>320</sup> But, instead of narrowly tailoring the ordinance by focusing on functional criteria, they simply “ban[ned] patrons from having sex with a device resembling a human at a business”<sup>321</sup>—a prohibition that, in theory, covers everything from blow up dolls to genital-shaped sex toys. And despite the troubling constitutional implications of its obviously overbroad wording, the ordinance passed unanimously.

The events in Houston point to a phenomenon that cries out for further attention. The policy response to the prospect of “robot sex” suggests that Masahiro Mori’s famous “uncanny valley” hypothesis<sup>322</sup> may extend as far as the regulatory realm. In fact, it seems the closer robots encroach on what we perceive as deeply human activities, the likelier they are to trigger a hasty regulatory backlash by horrified policymakers. As the Texas Governor recently noted, the city of Houston has “more brothels than Starbucks.”<sup>323</sup> But because Kinky S Dolls’s so-called brothel involved robots, the regulatory response was, well, uncanny.

We aren’t suggesting that people should stop anthropomorphizing (they won’t), or that the law should grant robots personhood because people treat them as human. But in deciding

---

<sup>315</sup> Tom Dart, ‘Keep *Robot Brothels Out Of Houston*’: Sex Doll Company Faces Pushback, THE GUARDIAN (Oct. 2, 2018), <https://www.theguardian.com/us-news/2018/oct/01/houston-robot-brothel-kinky-s-dolls-sex-trafficking>.

<sup>316</sup> See *id.*

<sup>317</sup> See *id.*

<sup>318</sup> See *id.*

<sup>319</sup> See *id.*

<sup>320</sup> Joel Shannon, *Proposed ‘Sex Robot Brothel’ Blocked by Houston Government: ‘We Are Not Sin City’*, USA TODAY (Oct. 4, 2018), <https://www.usatoday.com/story/news/nation-now/2018/10/03/sex-robot-brothel-blocked-houston-texas/1518298002/>.

<sup>321</sup> See *id.*

<sup>322</sup> Mori “hypothesized that a person’s response to a humanlike robot would abruptly shift from empathy to revulsion as it approached, but failed to attain, a lifelike appearance.” Masahiro Mori, *The Uncanny Valley: The Original Essay by Masahiro Mori*, IEEE Spectrum (Jun. 12, 2012), <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>.

<sup>323</sup> Alexandra Hart, *Does Houston Have More Brothels Than Starbucks?*, TEX. STAND. (Jul. 18, 2018), <https://www.texasstandard.org/stories/does-houston-have-more-brothels-than-starbucks/>.

how to regulate robots, one factor to consider is whether and to what extent people react as though the robot making decisions is human.

### C. Safeguarding Definitions

In the previous sections, we've urged regulators not to regulate robots directly, but instead to focus on behavior. When adopting this approach, we've also argued for an "I know it when I see it" definition that focuses on functional characteristics of interest rather than an all-or-nothing definition that tries to specify what a robot is.<sup>324</sup>

In this section, we assume legislators and regulators didn't listen to us. Perhaps the lure of a new law regulating robots was just too strong. Perhaps we're stuck with existing definitions and have built legal structures around them. Or maybe we do need to have some definition to work from; for instance, because the statute imposes criminal liability and the rule of lenity disfavors common law definitions of crimes.<sup>325</sup> If we are going to write definitions of robots into law, we offer a few ideas for how to mitigate the damage.

**Clouds, not Boxes.** First, if we can't use a functional or context-specific definition of robot, we should define using multiple categories rather than a single overarching definition. As Karni Chagal-Feferkorn explains, robots aren't a thing, they are a spectrum.<sup>326</sup> Past experience with putting things into boxes has led to efforts to position oneself just inside (or just outside) the box.<sup>327</sup> And the problem is particularly great when the contents of the boxes are constantly changing. The computer science community has developed the concept of "bounding boxes" to localize identified categories or concepts when exact definition is impossible.<sup>328</sup> We may not be able to perfectly identify what something is or where it is in advance. But we can identify a range of things it falls within or a zone in which it is located. The law needs "bounding buckets" to define robots—not an in-or-out definition, but one that identifies robots with various levels of confidence.

Setting boundary conditions allows definitions to evolve over time to meet changing circumstances. Precise definitions fix the technology and the legal rules in a particular moment, and that moment can quickly look outdated. We've seen examples of this with the Digital

---

<sup>324</sup> See *infra*, note \_\_.

<sup>325</sup> Intisar A. Rabb, *The Appellate Rule of Lenity*, 131 HARV. L. R. FORUM 179 (2018) (discussing the history of this rule of interpretation).

<sup>326</sup> Karni Chagal-Pfeferkorn, *Am I an Algorithm or a Product? When Products Liability Should Apply to Algorithmic Decision-Makers*, STAN. L. & POL'Y REV. (2019).

<sup>327</sup> Mark A. Lemley & Christopher R. Leslie, *Categorical Analysis in Antitrust Jurisprudence*, 93 IOWA L. REV. 1207 (2008). \_\_. See also *supra* notes \_\_-\_\_ and accompanying text (discussing examples in IP law).

<sup>328</sup> Prince Grover, *Evolution of Object Detection and Localization Algorithms*, TOWARDS DATA SCIENCE (Feb. 14, 2018), <https://towardsdatascience.com/evolution-of-object-detection-and-localization-algorithms-e241021d8bad> (discussing use of bounding boxes in image classification).

Millennium Copyright Act.<sup>329</sup> The Act cemented a 1998 view of the Internet that was obsolete by 1999—but one that, nevertheless, governs the law today.<sup>330</sup> By contrast, the concept of a “public accommodation” in the Americans with Disabilities Act was general enough that it now includes things like Netflix in the definition, even though Netflix didn’t exist when the law was written.<sup>331</sup> Our concept of a “public forum” may similarly be malleable enough to eventually apply to “spaces” online or in virtual reality even if they are privately owned. Applying legal regulation, not to existing devices that meet a particular statutory definition today but to devices now known or later conceived that have certain characteristics or perform certain duties, can help preserve a similar flexibility.

Defining by general concept rather than by strict boundaries also makes it harder to game the system. And, here, our bounding box analogy takes a very literal turn. Robots themselves have proven indispensable when it comes to catching other robots trying to game definitions. The Defense Department’s recent use of AI-powered forensics tools to spot “AI-made forgery” is one such example.<sup>332</sup> The “Deep Fakes” detecting technology uses “subtle cues in current [AI]-manipulated images and videos . . . to detect the presence of alterations” by bots.<sup>333</sup> “[S]trange head movements, odd eye color, and so on” serve as dead giveaways for bots trying to spot fakers. But even these tools have their limits.<sup>334</sup> Indeed, a researcher familiar with the project was also careful to note that these cat-and-mouse scenarios can quickly escalate into arms races—whereby new robots are “trained to outmaneuver forensics tools.”<sup>335</sup>

By now, there is a rich game theory literature exploring how to mitigate these types of arms races. But one approach we find particularly useful in the regulatory context involves setting multiple exclusionary criteria.<sup>336</sup> As we saw with Deep Fakes, relying on a small number of exclusionary criteria can allow adversaries to spot and circumvent them. But using many such criteria simultaneously helps to make this far more challenging. We think this approach is useful not just for bot detection systems, but also for regulators creating bounding buckets. Combining multiple functional considerations in our definitions not only helps to clarify one’s regulatory

---

<sup>329</sup> Pub. L. 105-304, 112 Stat. 2863 (1998) (codified as amended at 17 U.S.C. §§ 1201-05 (Supp. 1998), amended by 17 U.S.C.A. §§1201-04 (West Supp. 2000)).

<sup>330</sup> See *supra* notes \_\_-\_\_ and accompanying text.

<sup>331</sup> See Bryan Casey, *Title 2.0: Discrimination Law in a Data Driven Society*, *J. OF LAW & MOBILITY* (forthcoming 2019) (discussing this evolution of the term “public accommodation”).

<sup>332</sup> See Will Knight, *The Defense Department Has Produced the First Tools for Catching Deepfakes*, *MIT TECH. REV.* (Aug. 7, 2018), <https://www.technologyreview.com/s/611726/the-defense-department-has-produced-the-first-tools-for-catching-deepfakes/>.

<sup>333</sup> See *id.*

<sup>334</sup> See *id.*

<sup>335</sup> See *id.* (quoting Hany Faird, “a leading digital forensics expert at Dartmouth College”).

<sup>336</sup> See Steven Melendez, *Fake Documents Could Spur an AI Arms Race*, *FAST COMPANY* (Sept. 12, 2018), <https://www.fastcompany.com/90230758/fake-documents-could-spur-an-ai-arms-race> (showing how the use of multiple protocols serves as a better deterrent to attackers).

intent, but it also helps to build in redundancies that prevent companies from tweaking small characteristics to get into one bucket or out of another.

**Regulate, Don't Legislate.** A second way to limit the harm definitions can do is to prefer regulatory rulemaking to legislation. Regulators are more likely to consult experts in setting their definitions and more likely to take the time and effort to engage in notice-and-comment rulemaking,<sup>337</sup> thereby allowing problems with those definitions to surface.<sup>338</sup> Legislation that is passed at the last minute at the end of a legislative session, or that reflects a larger political compromise, is particularly likely to contain significant definitional flaws (or simply to sweep definitional problems under the rug).

Further, it's much easier to fix bad regulation than bad legislation. Regulators can change their definitions, sometimes without reopening the process of notice-and-comment. And if they're particularly worried about definitions becoming straightjackets for innovation, they can even opt for comprehensive guidance over formal rulemaking (as we've seen NHTSA do with national autonomous vehicle policy).<sup>339</sup> Regulators are also more likely to view regulation as an ongoing, iterative process than legislators, who may view their work on an issue as done once the bill becomes law. And this focus on the long game also makes regulators likelier to be on the lookout for legal rules that may inadvertently entangle new technological deployments (as we saw in Part II.B.2). What's more, regulators are less subject to partisan gridlock because they work for an executive of one party. And they work full time, while legislatures go in and out of session and sometimes sit only for a few months every two years.<sup>340</sup> So we're more likely to be stuck with a bad definition written into a statute than we are with regulations.

Finally, regulators are also uniquely situated to engage with industry. They're not only able to work alongside the private sector to describe and distinguish technologies, develop best-practices, and create industry-specific regulatory models, but they're also able to influence the private sectors' own efforts to self-regulate. If recent events are any indication, these efforts can be as influential as rulemaking itself. As Robert Gorwa and Douglas Guilbeault note, "[P]rivate (often rather opaque) policies can have serious political ramifications, potentially placing them more squarely within the remit of regulatory and legal authorities."<sup>341</sup> And if our goal is to take a nuanced approach to regulating robots, we'll want regulators in both the public *and* private sectors to increasingly coordinate.

---

<sup>337</sup> Administrative Procedures Act, 5 U.S.C. § 500 et seq.

<sup>338</sup> Kaminski, *supra* note \_\_.

<sup>339</sup> Mallory Locklear, *US Department of Transportation updates Autonomous Car Rules*, ENGADGET (Oct. 4, 2018), <https://www.engadget.com/2018/10/04/us-department-of-transportation-autonomous-car-rules/>. Iterative approaches like these have their own drawbacks.

<sup>340</sup> Kate Galbraith, *Biennial Blues?*, TEX. TRIBUNE (Dec. 31, 2010), <https://www.texastribune.org/2010/12/31/defying-national-trend-texas-clings-biennial-legis/> (noting "Texas is one of a dwindling number of states whose legislatures hold scheduled meetings only every two years").

<sup>341</sup> See Gorwa & Guilbeault, *supra* note \_\_ at \_\_.

To be clear, regulation is no panacea. Regulators, like legislators, are subject to capture.<sup>342</sup> Working closely with the affected industry can be a benefit, as Margot Kaminski suggests,<sup>343</sup> but it can also lead to relationships that are too cozy. There are other dangers with regulating robots in particular. A regulator is likely to regulate—that’s how they justify their existence, after all. And sometimes, perhaps most times, the right thing to do with new technology is to forebear from regulating.<sup>344</sup> So we’re not suggesting the creation of a new body to govern robots, as some have done.<sup>345</sup> But if we must regulate robot behavior, and if we must define robots to do so, writing those definitions into industry-specific regulations rather than either a general-purpose definition or statute is likely to minimize the harm we cause when we get that definition wrong, as we inevitably will.

**Risk-Mitigation Strategies.** Third, if we are to define robots, and especially if we’re going to do so in statutes, we should establish safeguards that permit us to review and revise the definitions in light of new information. Statutory definitions should come with sunset clauses<sup>346</sup> so that we don’t have to live with an obsolete definition for years while legislators are gridlocked or busy elsewhere. The application of those definitions should be subject to administrative and judicial review, not just for blind adherence to the words of the statute but to allow those who think they have wrongly been classified as robots (or, more rarely, wrongly classified as non-robots) to appeal that decision or seek an exception from the definition. One positive example here comes, ironically, from the Digital Millennium Copyright Act we criticized earlier. While that statute locked in obsolete definitions of Internet technology, missing peer-to-peer technology entirely even though it was less than a year away, it created a mechanism for Copyright Office regulators to consider and add exceptions to the anticircumvention provisions of the Act.<sup>347</sup> Regulators meet every three years and consider and rule on possible exceptions. That process has worked well for 20 years, and it has allowed us to permit now-common practices like unlocking a phone that weren’t contemplated by Congress but would otherwise have been prohibited by a strict application of the statute. We think that any legislative definition of robot should come with a similar process for making exceptions as technology evolves.

## Conclusion

---

<sup>342</sup> See *supra* note \_\_\_\_.

<sup>343</sup> See Kaminski, *supra* note \_\_\_\_.

<sup>344</sup> Anupam Chander, *How Law Made Silicon Valley*, 63 EMORY L.J. (2015); Josh Lerner & Greg Rafert, *Lost in the Clouds: The Impact of Copyright Scope on Investment in Cloud Computing Ventures*, National Bureau of Economic Research Working Paper 9 No. 21140.

<sup>345</sup> Calo, *supra* note \_\_\_\_.

<sup>346</sup> See Ittai Bar-Siman-Tov, *The Lives and Times of Temporary Legislation and Sunset Clauses*, AM. J. OF COMP. LAW. (2017) (discussing the role of sunset clauses and temporary legislation).

<sup>347</sup> See Pub. L. 105-304, *supra* note \_\_\_\_.

As the human impacts of robots and AI increase, so too will our efforts to regulate them. To regulate robots, we'll first need to establish what one is. As it turns out, this is not a straightforward task. As we've seen, many of our current attempts to define robots have failed miserably. Indeed, if you're reading this, you're (probably) not a robot, but the law might already treat you as one. The problem, however, isn't simply a matter of failing to hit on the right definition. Rather, for a variety of reasons, there may *be* no right definition of robots.

Instead of trying in vain to establish perfect definitions, there is a better alternative. We should do as Turing did when confronted with the challenge of defining robots: embrace their ineffable nature and adapt our regulatory tools accordingly. We may not be able to define robots *ex ante*. But as with obscenity, quite often we will know robots when we see them. In other words, a common law, case-by-case approach provides a promising means of successfully navigating the definitional issues presented by robots—one that builds and adapts its categories incrementally by focusing on the most salient functional characteristics. Want to regulate robots? Try instead regulating worrisome behavior regardless of who or what engages in it. Doing so avoids definitional traps and sharpens our regulatory focus—thereby making it less likely that the law will be easy to game and less likely that it will inadvertently interfere with innovation.