How should the law think about robots?

[Neil M. Richards](#)
Professor of Law
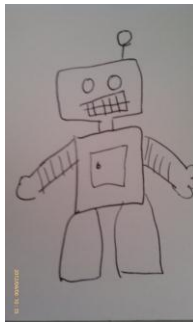Washington University in St. Louis

[William D. Smart](#)
Associate Professor of Computer Science
Director, [Media and Machines Laboratory](#)
Co-director, [Masters of Engineering in Robotics](#) program
Department of Computer Science and Engineering
Washington University in St. Louis

The robots are coming. We don't mean this in a scary, apocalyptic way, like in *The Terminator* or *Battlestar Galactica*, or in a tongue-in-cheek way, like in the Flight of the Conchords song "The Humans are Dead."[1] What we mean is this: Robots and robotic technologies are now mature enough to leave the research lab and come to the consumer market in large numbers. Some of them are already among us, like Rumbas, robotic carers in hospitals, drones used by military and law enforcement, and the prototype self-driving cars that have

---

[1] Flight of the Conchords, The Humans Are Dead ("It is the distant future, the year 2000. The world is very different ever since the robot uprising of the late 90s. There have been some major changes....All human life has been eradicated. Finally, robotic beings rule the world.").

started to appear on our roads.[2] These early technologies are just the start, and we might soon be witnessing a personal robotics revolution. These systems have the potential to revolutionize our daily lives and to transform our world in ways even more profound than broad access to the Internet and mobile phones have done over the past two decades. We need to be ready for them and, in particular, we need to think about them in the right way so that the lawmakers can craft better rules for them, and engineers can design them in ways that protect the values our society holds dear. But how should we do this?

This essay is an attempt to think through some of the conceptual issues surrounding law, robots and robotics, to sketch out some of their implications. It draws on our experience as a cyberlaw scholar and a roboticist to attempt an interdisciplinary first cut at some of the legal and technological issues we will face. Our paper is thus analogous to some of the first generation cyberlaw scholarship that sketched out many of the basics of the field, even before the field itself was a recognized one.[3] Our work (as well as the larger project of law

---

[2] Tom Vanderbilt, *Let the Robot Drive: The Autonomous Car of the Future is Here*, Wired Jan. 20, 2012..

[3] E.g., M. Ethan Katsh, *Software Worlds and the First Amendment: Virtual Doorkeepers in Cyberspace*, 1996 U. Chi. L. F. 335, 348–54 (1996); Joel R. Reidenberg, *Governing Networks and Rule-Making in Cyberspace*, 45 Emory L.J. 911, 917–20; David R. Johnson & David Post, *Law and Borders — The Rise of Law in Cyberspace*, 48 Stan . L. Rev. 1367, 1375 (1996); David Kushner, *The Communications Decency Act and the Indecent Indecency Spectacle*, 19 Hast. Comm. & Ent. L.J. 87, 131 (1996); David G. Post, *Anarchy, State, and the Internet: An Essay on Law-Making in Cyberspace*, 1995 J. Online L. art. 3, 12–17 (1995); Lawrence Lessig, *The Zones of Cyberspace*, 48 Stan. L. Rev. 1403, 1403 (1996); Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. Chi. L. Forum 207; Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 Harv. L. Rev. 501, 1999; Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1198–99 (1998); Lawrence Lessig, *Code and Other Laws of Cyberspace* (2000); , For an overview of more of this scholarship, along with a history of the development of the Internet, See Raphael Cohen-Almogor, *Internet History*, 2 Int'l J. Technoethics, Apr.-June 2011, available at http://www.hull.ac.uk/rca/docs/articles/internet-history.pdf.

and robotics that is just getting started) has one great advantage over that earlier cyberlaw work: it has the benefit of the cyberlaw project's wisdom of two decades of thoughtful analysis of the relationship between legal and technological norms in a period of revolutionary change. Cyberlaw can provide the blueprint for law and robotics, in both its successes and its challenges.

In this essay, we would like to advance four claims about the ways we, as scholars and as a society, should approach this problem. First, we offer a definition of robots as *non-biological autonomous agents* that we think captures the essence of the regulatory and technological challenges that robots present, and which could usefully be the basis of regulation. Second, we briefly explore the sometimes surprisingly advanced capabilities of robots today, and project what robots might be able to do over the next decade or so. Third, we argue that the nascent project of law and robotics should look to the experience of cyber-law, which has struggled instructively of the problems of new digital technologies for almost two decades. This experience has revealed one particularly important lesson: when thinking about new technologies in legal terms, the metaphors we use to understand them are crucially important. Lawyers are used to understanding legal subjects metaphorically, especially in developing areas of the law like new technologies. If we get the metaphors wrong for robots, the lessons of cyber-law reveal that it could have potentially disastrous consequences. Finally, we argue that one particularly seductive metaphor for robots should be rejected at all costs, the idea that robots are "just like people" and that there is a meaningful difference between humanoid and non-humanoid robots. We call this idea "the Android Fallacy."

1. What is a Robot?

Before we can think about these systems, we need to have a clear understanding of what we mean by "robot". The word itself comes from a Czech play from the 1920s, entitled

"R.U.R. (Rossum's Universal Robots)", by Karel Čapek.[4] In the play, the "robots" are artificial humans used as slave labor in a factory ("roboti" in Czech translates to "serf labor," with the associated connotations of servitude and drudgery). The term "roboticist", one who studies or creates robots, was coined by Isaac Asimov in 1941.[5] Even the etymology of the word suggests a device that is well-suited for work that is too dull, dirty, or dangerous for (real) humans.

So what is a robot? For the vast majority of the general public (and we include most legal scholars in this category), the answer to this question is inescapably informed by what they see in movies and, to a lesser extent, in literature and the popular science press. When asked what a robot is, they will generally give an example from a movie: Wall-E, R2-D2, and C-3PO are popular choices. Older respondents might also mention The Terminator or Johnny-5. Movie buffs will often mention Huey, Louie, and Dewie (from Silent Running), the false Maria (from Metropolis), the gunslinger (from Westworld), and an increasingly esoteric list of others. These are all clearly robots: they are all mechanisms, built from mechanical parts by humans (or other robots) to perform a specific dull, dirty, or dangerous job. They are all also anthropomorphic, or easy to anthropomorphize. R2-D2 is not human-like, but it is clear when "he" is "happy" or "irritated." Movie robots are plot devices, and work best when we can project human-like qualities (or the lack of them, in the case of *The Terminator*)[6] on them.

What about the less clear cases? HAL 9000 (from *2001: A Space Odyssey*) was an intelligent computer that controlled a large spaceship. In many senses, the ship was HAL's "body". Was HAL a robot? It could certainly move about and manipulate things in its world, two features that we expect of a

---

[4] "R.U.R. (Rossum's Universal Robots)", Karel Čapek, 1920. Translated by Paul Selver and Nigel Playfair, 2001. Dover Publications.

[5] "Liar!", Isaac Asimov. In Astounding Science Fiction, May 1941. Reprinted in "I, Robot" (1950).

[6] Of course, later Terminator movies explore the humanity of even these machines. The conclusion of *Terminator 2: Judgment Day* revolves around a copy of Arnold Schwarzenegger's now-obsolete model from the first film consciously sacrificing himself in order to save humanity.

robot. What about the replicants from *Blade Runner,* Cylons from *Battlestar Galactica* (2005), and Bishop the "synthetic person" from *Aliens*? They are human in appearance, but constructed from organic material, not metal. Do they meet our criteria for being a robot?

Even professional roboticists do not have a single clear definition. Arms that assemble cars, tele-operated submarines that explore the ocean depths, space probes hurtling through the void, remote-control cars augmented with little computers and sensors, and human-like androids all fall under the definition of "robot", depending on who you ask.

So how do we usefully define a "robot" for the purposes of this paper? In most of the examples above, the robots can move about their world and affect it, often by manipulating objects. They behave intelligently when interacting with the world. They are also constructed by humans. These traits are, to us, the hallmarks of a robot. We propose the following working definition:

> **A robot is a constructed system that displays both physical and mental agency, but is not alive in the biological sense.**

That is to say, a robot is something manufactured that moves about the world, seems to make rational decisions about what to do, and is a machine. It is important to note that the ascription of agency is subjective: the system must only *appear* to have agency to an external observer to meet our criteria.[7] In addition, our definition excludes wholly software-based artificial intelligences that exert no agency in the physical world.

Our definition intentionally leaves open the mechanism that causes the apparent agency. The system can be controlled

---

[7] This external ascription of agency is similar in spirit to the classic Turing test, where an external observer tries to identify a conversational partner, using what amounts to an Instant Messaging system, as either a human or a computer. A computer is said to have passed the Turing test if it causes the observer to reliably classify it as another human. While this is often seen as an intelligent test and a measure of artificial intelligence, it is more correctly thought of as a test of human-ness.

by clever computer software, or teleoperated by a remote human operator. While both of these systems are robots by our definition, the legislative implications for each of them are quite different, as we argue below.

2. What Can Robots Do?

Now that we have a definition of what robot is, we turn to what robots can do today. Since many of us are informed by movies, sound-bite media, and other unreliable sources, we are often poorly-informed about what state-of-the-art robots look like and what they can do right now. Robots have not yet reached the levels of capability the public associates with science fiction, but they are surprisingly close.

Until recently, the majority of "robots" in the world, over a million by some counts,[8] were the industrial automatons that assemble cars, move heavy parts, and otherwise make factory workers' jobs easier. These are, for the purposes of this paper and by our definition above not really robots; although they certainly have physical agency, they have no mental agency. Most of these systems perform set motions over and over, without regard for what is happening in the world. Spot-welding robots will continue to spot-weld even if there is no car chassis in front of them.

But "robots" within our definition do exist today. The most common robot in the world is now the iRobot Roomba (*see Fig. 1*), a small robot that can autonomously vacuum-clean your house. iRobot claims to have sold over 6 million of Roombas as of the end of 2010.[9] The number is certainly larger now, a year and a half later. These little critters *are* robots by our definition; they have both physical and mental agency. The computer algorithms that control them are simple, but they appear to make rational decisions as the scoot around the floor avoiding objects and entertaining your cat. The Roomba is fully autonomous, and needs no human assistance, despite operating in a cluttered real-world environment (your house); this is a more impressive

---

[8] International Federation of Robotics web site, http://www.ifr.org/industrial-robots/statistics/

[9] Transcript of iRobot Q4 2010 Earnings Call, http://seekingalpha.com/article/252090-irobot-ceo-discusses-q4-2010-results-earnings-call-transcript

achievement than one might think, especially given these inexpensive robots are available to consumers for only a few hundred dollars, depending on model.

Fig. 1: iRobot Roomba [iRobot Corporation].

Other more expensive robots are seeing heavy use in military settings all over the world. Cruise missiles, which meet our definition of robot,[10] have been used for many years by the United States military, and by other countries. More recently, remote-controlled drone aircraft, many of which we classify as robots, have seen heavy use in intelligence-gathering and offensive roles. Ground-based tele-operated robots, such as the Packbot (iRobot) (*see Fig. 2*) and the Talon (Foster-Miller) are becoming ubiquitous in modern military settings. These systems can replace human soldiers in dangerous situations: disabling a bomb, performing reconnaissance under fire, or leading the assault on a building. Based on extrapolations of earlier sales figures[11] for a single type of these ground robots, it is reasonable to estimate that

---

[10] The cruise missile clearly has physical agency, since it moves. It also has mental agency, to an external observer, since it can avoid terrain features while flying close to the ground. While this is done using a detailed map and a GPS device in modern missiles, the mechanism is unimportant to our definition. It is, of course, not capable of avoiding obstacles not in its map. However, it will rarely encounter such obstacles, to the external observer will not get to see this failure of apparent agency.

[11] "iRobot Delivers 3,000th Packbot", Robotics Trends, http://www.webcitation.org/5xPANQOLV

there are 10,000 such systems currently in use worldwide, in both military and civilian roles. These robots can drive around under remote control, often have an arm that can pick up and manipulate objects, and have a suite of sensors that relay data back to the operator. While they are completely controlled by a human operator, and currently have no autonomous capabilities, they often look intelligent to an external observer (who might be unaware that there is a human pulling the strings).

Fig. 2: iRobot Packbot [iRobot Corporation].

NASA has a long history of sending robots into space and to other worlds. The most successful recent examples are probably the Mars Exploration Rovers, Spirit and Opportunity (*see Fig. 3*). These were sent to Mars in 2003, and although no communication has been received from Spirit since March 2010, Opportunity is still operational after nine years on the surface. The rovers are *mixed initiative* or *shared autonomy* systems; they receive high-level instructions from human operators ("go over to that boulder"), but are responsible for their own low-level behavior (avoiding obstacles, for instance).

Fig. 3: A Mars Exploration Rover [NASA].

Finally, autonomous warehouse robots, designed by Kiva Systems, now help humans fulfill orders for several on-line retailers, including Zappos and Amazon.com (who recently acquired Kiva Systems for $775 million[12]). These robots bring whole racks of merchandise to a human who selects the appropriate items for a given order, and puts them in a shipping box. The robots are centrally-coordinated by the inventory system, and operate autonomously. The robots have no on-board sensors, and rely on wires embedded in the factory floor to determine their location. However, they certainly seem to have their own mental agency as they avoid each other and reconfigure the storage locations of items in the warehouse based on customer demand. (*See Fig. 4*).

*Fig. 4:The Kiva Automated warehouse system
[Kiva Systems].*

While not an exhaustive list of robots currently being used in the world, the above examples are representative of

---

[12] http://www.reuters.com/article/2012/03/19/us-amazoncom-idUSBRE82I11720120319

how robots are being used today. There are several common threads that run through each of these examples. In the cases where the robots are autonomous, the task is very well-constrained and the autonomy is at a relatively low level (avoiding things, as opposed to deciding who is an enemy on the battlefield). In several cases, there is no autonomy, and the robots are physical and perceptual extensions of remote operators. In the military setting, this is partly to the chain-of-command is respected and that responsibility for any given action rests with a human within this chain. However, it is also partly because, for many of the complex problems we encounter in the real world, we cannot yet build autonomous systems that can perform well. Finally, none of these systems interact directly with humans, other than their operator.[13] In fact, humans are often intentionally kept away from the robots, as has been the norm with industrial automation for over 60 years. The only exception in our list is the Roomba. However, when it interacts with you, it does so in the same way that it interacts with a table; the robot does not differentiate a human obstacle from a non-human one.

So much for robots that are actually in use. What can robots do in the research lab today? This is a more interesting list, since it suggests what robots will be doing in the real world in the coming years. A Google search for "cool robot video" uncovers over 31 million hits: robots dancing, climbing, swimming, jumping, folding towels, and fetching beer. Robots interacting with people, asking them questions and guiding them through shopping malls. Flying robot quadcopters[14] performing breathtaking acrobatics. Robots making cakes, cookies, pancakes, and full Bavarian breakfasts. Robots building maps and models of the world. The list seems almost endless. The vast majority of these videos come from research labs, either in academia or industry, and generally showcase some interesting new technical advance. For example, endowing a robot with the dexterity to control a tool with precision and to apply just the right amount of force is an important problem. Stirring a cake mixture with a wooden spoon until it is just the right consistency is a great test and

---

[13] At least they do not interact with humans who are likely to care about the legislative and consumer protections that apply to the robots.

[14] Small robot helicopters with 4 rotors, favored for their stability.

demonstration of this dexterity. Plus, it results in cakes that the graduate students can eat.

While it is impossible for us to briefly survey all of the current research going on worldwide, we can highlight some common themes that seem to be emerging. First, many research robots are now multi-purpose, rather than being designed for a single task in the way that many of the systems above are. For example, the PR2 robot from Willow Garage[15] appears in many videos, performing a variety of tasks. (*See Fig. 5*). Second, robots are starting to interact with people who do not know anything about robots. They no longer need to be accompanied by a graduate student escort, who has traditionally acted as a minder, interpreter, mechanic, and bodyguard. People with no experience of robots are now encountering and collaborating them them directly. While this is necessary if we are to fulfil the long-term potential of the technology, it also complicates matters hugely. Humans are unpredictable, easy to damage, and hard to please; a lot of research is currently aimed at allowing robots to deal with them gracefully and safely. Third, they are becoming more and more autonomous, as we solve the underlying technical challenges of perception and reasoning. Finally, there is an increasing focus on robots that work in the real world, not just in the lab. This requires us to deal with all of the uncertainty and unpredictability inherent in world in which we live.

Fig. 5: The PR2 research robot
[Willow Garage, Inc.].

---

[15] http://www.willowgarage.com/pr2/

These research robots are starting to make the transition into the real (or, at least, commercial) world. Google has a fleet of self-driving cars that have traveled more than 150,000 miles on the US road system without incident. (*See Fig. 6*). Robots used as therapeutic aids are available, and quickly becoming more widespread.[16] Robots are being evaluated as assistants in the homes of individuals with severe motor disabilities.[17] These trends will only accelerate in the coming years. More and more robots will enter our daily lives in the coming decade, and it is likely that many people will own a (useful) personal robot by 2022. This appearance of robots will drive a number of legislative challenges.

Fig. 6: A Google driverless car [Google].

As 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. In the extreme (and very far-future) case of a robot that can do everything a human can, there are few practical boundaries on what the robot can be used for. How does one legislate such a system? No other devices are like it, meaning we must come up with suitable analogies and metaphors, which, we claim, will be tricky.

As robots enter public life and our private homes,, the protections associated with them must be more comprehensive and robust than those currently in place for research robots. Most research robots come with many warnings and disclaimers, and rely on the users (who are trained professionals) not to do anything stupid. This is simply

---

[16] Paro theraputic robot, http://www.parorobots.com/

[17] Robots for Humanity Project,
http://www.willowgarage.com/blog/2011/07/13/robots-humanity

not practical for the general public, since they have to technical training, and cannot be relied on to exercise good judgement and caution.

As robots become more autonomous, the question of where liability rests when something goes wrong is complicated. Is it the manufacturer, the programmer, the user (who gave a bad instruction), or some combination of them all? The matter will be complicated in systems that are autonomous some of the time and tele-operated at other times, since this introduces a remote operator who might be controlling the robot in good faith, but with limited sensory information.

As robots enter the real world, our ability to predict what will happen decreases dramatically. Uneven floor surfaces, unexpected obstacles, small children, and a host of other factors make controlling the robot safely hard, and designing legislation that is comprehensive but does not overly constrain the use of the systems will be challenging.

3. Robo-Law and Cyber-Law

The sheer variety of applications that robots can and will be used for willthus put pressure on the legal system in a wide variety of substantive areas, including tort, contract consumer protection, privacy, and constitutional law, among others. Although robotic technologies will inevitably raise multiple novel legal questions, legal understandings of robots and how to think of them are in their infancy. There is currently very little scholarship on the intersection of law and robotics, though a few scholars have begun the task of beginning to think about the issues involving law and robotics in a systematic way.[18] Similarly, we are starting to see the first robot-specific laws being enacted, but such laws are currently rare enough to bring a sense of novelty. In June 2011, Nevada became the first state to pass a law regulating driverless

---

[18] E.g., Ian Kerr. *Bots, Babes, and Californication of Commerce*, 1 Univ. Ottawa L. & Tech. J. 285 (2004); M. Ryan Calo, *Open Robotics*, 70 Md. L. Rev. 571 (2011); M. Ryan Calo, *Robots and Privacy*, in Robot Ethics: The Ethical and Social Implications of Robotics, (Patrick Lin, George Bekey, and Keith Abney, eds., Cambridge: MIT Press, Forthcoming).

robotic cars.[19] The law granted rulemaking authority to the Nevada Department of Transportation to regulate the testing, safety, and ultimately the presence on its roads of "autonomous vehicles" using artificial intelligence and sensors such as GPS sensors and lasers. Under the law, an autonomous vehicle is "a motor vehicle that uses artificial intelligence, sensors and global positioning system coordinates to drive itself without the active intervention of a human operator."[20] "Artificial intelligence," in turn, was defined as "the use of computers and related equipment to enable a machine to duplicate or mimic the behavior of human beings."[21] The law was the product of consultation with automakers, Google, insurance companies, and consumer groups, and has been generally well received, with regulations implementing the statute unveiled on February 12, 2012.[22]

Such academic and legislative interventions remain outliers. As a society we lack an awareness of the impending revolution in robotics, much less any concrete understandings about how the law should regulate or even understand robots. This is a problem because uncertainty about (for example) liability caused by robots could hamper innovation and the widespread consumer adoption of these useful and potentially transformative technologies. How could we develop such an understanding? We would suggest that the law and robotics project should look to the lessons of other bodies of law, which have grappled with varying degrees of success with the problem of regulating new digital technologies for decades.[23] The experience of cyberlaw and other areas of technology-influenced jurisprudence has revealed one particularly important lesson for technologically-sophisticated applications of law - when it comes to new technologies, applying the right metaphor for the new technology is especially important. How we regulate robots will depend on the metaphors we use to

---

[19] Nev. Rev. Stat. Ch. 482A (effective March 1, 2012), available at http://www.leg.state.nv.us/Session/76th2011/Bills/AB/AB511_EN.pdf.

[20] Nev. Rev. Stat. Ch. 428A.020.

[21] Nev. Rev. Stat. Ch. 428A.030.

[22] Nev. Adopted Reg. Regulation R084-11, LCB File No. R084-11, Effective March 1, 2012, available at http://www.leg.state.nv.us/register/RegsReviewed/$R084-11_ADOPTED.pdf.

[23] See sources cited supra note 3.

think about them. There are multiple competing metaphors for different kinds of robots, and getting the metaphors right will have tremendously important consequences for the success or failure of the inevitable law (or laws) of robotics.

A classic example from the twentieth century illustrates the importance of getting the metaphors for new technologies right. It concerns how Fourth Amendment law came to understand the nature of government wiretapping under the Fourth Amendment, which requires the government to obtain a warrant before it searches its citizens' "persons, houses, papers, and effects."[24] In the 1928 case of *Olmstead v. United States,*[25] the Supreme Court was called upon to determine whether the police's warrantless wiretapping of a phone line leading into the home of notorious bootlegger Roy Olmstead constituted a "search" that would have required a warrant. Chief Justice Taft's opinion for the Court used a physical conception of a search rather than a broader understanding of the emerging capabilities of electronic technologies. The Court accordingly held that because wiretaps required neither physical trespass into the home nor the seizure of "tangible material effects," the Fourth Amendment was inapplicable to wiretapping.

By contrast, Justice Brandeis's dissent in *Olmstead* viewed the issue more broadly, and would not have required a physical intrusion from new technologies for the Fourth Amendment's protections to apply. Brandeis viewed existing law as establishing the principle that the Fourth Amendment protected against "invasion of the sanctities of a man's home and the privacies of life."[26] However, Brandeis warned that these protections were being threatened by emerging technologies like wiretaps that had enabled "[s]ubtler and more far-reaching means of invading privacy . . . . Discovery and invention have made it possible for the government, by means far more effective than stretching upon the rack, to obtain disclosure in court of what is whispered in the closet.[27]" Brandeis also suggested that science was likely to provide governments in the future with even more invasive and secret methods of surveillance beyond wiretapping. He warned that

---

[24] U.S. Const. Amd. IV.

[25] 277 U.S. at 466 (1928)..

[26] Id. at 465 (Brandeis, J., dissenting).

[27] Id. at 473-74 (Brandeis, J., dissenting).

> [w]ays may some day be developed by which the
> government, without removing papers from secret
> drawers, can reproduce them in court, and by which it
> will be enabled to expose to a jury the most intimate
> occurrences of the home. Advances in the psychic and
> related sciences may bring means of exploring
> unexpressed beliefs, thoughts and emotions."[28]

Brandeis conceded that the Court's position was a
straightforward reading of the text of the Fourth Amendment.
But this reading, he maintained, was deeply flawed because it
clung to a narrow and outmoded view of the Fourth
Amendment as protecting only tangible property, and thus
failed to grasp the nature of the threat that the new technology
posed. By failing to understand the nature of the new
technology -- by applying only a physical metaphor focused on
trespass rather than a broader one rooted in conceptions of
privacy -- the Court's position failed to protect important
values in the face of new technologies.

The subsequent course of search and seizure law has
vindicated Brandeis's position about the right metaphor by
which to understand wiretapping technologies. Soon after the
case was decided, Congress enacted section 605 of the Federal
Communications Act, which made wiretapping a federal crime.
In 1937, the Supreme Court held that federal agents could not
introduce evidence obtained as a result of an illegal wiretap in
federal court.[29] And in the 1967 case of *Katz v. United States*,
after a series of cases chipping away at the *Olmstead* trespass
metaphor, the Supreme Court finally changed course and
adopted the Brandeis position that the Fourth Amendment
applied to wiretaps.[30] This was the case because the Fourth
Amendment was not limited merely to physical invasions, but
rather protected people rather than places against
unreasonable searches and seizures.[31] In a famous
concurrence in the same case that later became the blueprint
for modern Fourth Amendment law, Justice Harlan suggested
that the applicability of the Fourth Amendment should turn on
"a twofold requirement, first that a person have exhibited an
actual (subjective) expectation of privacy and, second, that the

---

[28] Id. at 474 (Brandeis, J., dissenting).

[29] *Nardone v. United States,* 302 U.S. 379 (1937).

[30] 389 U.S. 347 (1967).

[31] Id. at 351.

expectation be one that society is prepared to recognize as 'reasonable.'"[32]

The *Olmstead-Katz* example is a helpful one for several reasons. It shows quite clearly how different understandings of how emergent technologies apply to human activity can have profound legal consequences. And it shows how misunderstanding a new technology -- either how it works or what values it threatens -- can have pernicious effects. The *Olmstead* court failed to recognize the threat to privacy that unregulated government wiretapping presented. It clung to outmoded physical-world metaphors for the ways police could search without a physical trespass. By contrast, Justice Brandeis understood the threat that the new technology presented to established values, and asked not whether the new police technology constituted a physical trespass, but a threat to the broader value of citizen privacy against the state. He asked a better question of the new technology, was willing to adapt the law to fit changed technological circumstances while preserving its old normative values, and generated a better legal answer as a result.

Interestingly, the trespass/privacy issue of Fourth Amendment law that produced the *Olmstead/Katz* line of cases remains vital today. In its recent *Jones* decision, the Supreme Court held that a GPS transponder placed on a criminal's car that was unsupported by a search warrant violated the Fourth Amendment. While the entire court agreed that the Fourth Amendment was violated, it splintered about why this was the case. Justice Scalia's bare majority invalidated the law on the old trespass theory from *Olmstead,* while Justice Alito and three other Justices would have adopted a broader privacy justification that would have prevented non-trespassory GPS monitoring. Justice Sotomayor seemed caught between both camps.[33] Like *Olmstead* and *Katz*, *Jones* illustrates that the metaphors we use to understand the technology and the law matter a great deal, and can have profound consequences to which new practices the law limits, and which it allows.

4. The Importance of Metaphors

---

[32] Id. at 361 (Harlan, J., concurring).
[33] United States v. Jones X U.S. Y (2012).

This example illustrates the importance of metaphors when law confronts new technologies. In designing and implementing new technologies, we must be mindful of the metaphors we use to understand the technologies. As these cases suggest, metaphors matter at several levels. At the conceptual design level, designers of cutting-edge technologies frequently understand the problem or the solution in terms of something else. The metaphorical choice (either implicit or explicit) to design a technology as a new version of an existing thing has real effects on how research questions are framed and pursued, expanding or limiting the range of possible results that can be tested and engineered. For example, a video streaming service might understand itself as a movie theatre, a bookstore, a library, or a television network. These understandings shape both the ways technologies are designed and the sorts of potential problems that engineers try to anticipate. In this regard, consider the differences between a software download service like iTunes, which designed itself as a bookstore with initially heavy digital rights management protections, and streaming services like Netflix, which allows users to "rent" videos for a limited time like a video store, or Spotify, which allows access to unlimited music like a radio station that charges a fee. Similarly, when technologies leave the laboratory and enter the outside world, both consumers and the legal system will use metaphors to try to understand the technology. To stay with the example of digital music, many early users of music on .mp3 files shared them freely in the tradition of the mix tape. By contrast, copyright holders (and increasingly the legal system) have viewed such sharing as theft, asserting that it is more akin to "piracy," with all the metaphorical baggage that term entails.

In the context of robots, these considerations are particularly important. How we think about, understand, and conceptualize robots will have real consequences at the concept, engineering, legal, and consumer stages. At the concept stage, how we think about robots (and their human operators) will affect their design. Do we want them to be virtual butlers? Virtual pets? Virtual children? The answers to such questions will affect not only how the robots are configured to solve particular problems, but also how they are physically presented. Butlers and children typically don't have offensive capability; some pets (like dogs) do. Children and pets are less autonomous than butlers, while children and butlers (but not pets) are anthropomorphic. Nor is the

butler/children/pet an exhaustive list of possibilities; we could conceive of robots as personal computers or gaming systems with wheels, as housekeepers, roommates, sexual partners, or even spouses.

The importance of metaphor continues at the legal level. Lawyers are trained from the first day of law school in "thinking like a lawyer": essentially the act of reasoning by analogy from one case to another.[34] Particularly in the context of changing technologies, the law almost always considers new technology as merely a new form of something else. Wiretaps, as we saw earlier, were first analogized to physical searches, and only later to broader invasions of an interest in "privacy." Websites were thought of as property (rather than, for instance, unclaimed land, or trademarks or telephone numbers, or something entirely new altogether), and subjected to a regime of anti-"cyber-squatting, wherein companies with existing trademarks similar to the Universal Resource Locators (URLs) of people who had claimed them already.[35] Electronic messages were analogized as "e-mail" and given heightened protection from snooping, even though the underlying packet-switching technology could have been viewed as more akin to postcards, whose contents receive much lower protection from surveillance under pre-existing law. For autonomous robots, the importance of metaphors extends not only to how the legal system will understand (and regulate) the robots themselves, but also how it will understand (and regulate) human operators of semi-autonomous robots which could come to market first to fill the gap in our current technical ability to produce fully autonomous robots.

Finally, the importance of metaphors matters at the consumer level. Numerous studies have shown that the way people react to technology in different ways depending upon how it is presented (or marketed). For example, previous studies have shown that people react differently to technology that is anthropomorphic in shape (or isn't), has visible eyes like a human (or doesn't), or speaks with a human voice (rather than text). These behavioral reactions appear to be hard-wired, but even if there is a social construction at work,

---

[34] E.g. Karl Llewellyn, *The Bramble Bush* (1930); Frederick Schauer, *Thinking Like a Lawyer* (2009).

[35] The Anticybersquatting Consumer Protection Act (ACPA), 15 U.S.C. § 1125(d) (1999).

the demonstrable effects of human-like versus non-human-like technology will have a real effect on how consumers react to, accept, and trust robots in their homes.[36]

Metaphors can constrain thinking, sometimes in an unnecessarily limiting way (if they rest on old social norms or technical limitations that are no longer applicable), and sometimes in a way that reflects the enduring wisdom of the past. One of our goals in the law and robotics project will be to see the big picture – the way that metaphors operate to show how people understand and react to public and domestic robotics, and to design the robots in ways that take advantage of good phenomena while avoiding undesirable outcomes.

5. The Android Fallacy

We must also beware seductive but dangerous metaphors. When we think of robots, we often picture them as anthropomorphic; C-3PO from Star Wars is a good example. Even when a particular robot is not shaped like a human, we find it hard not to project human-like features, intentions, and motivations onto it. Even in research labs, cameras are described as "eyes", robots are "scared" of obstacles, and they need to "think" about what to do next. This projection of human attributes is dangerous when trying to design legislation for robots. Robots are, and for many years will remain, tools. They are sophisticated tools that use complex software, to be sure, but no different in essence than a hammer, a power drill, a word processor, a web browser, or the braking system in your car. As the autonomy of the system increases, it becomes harder and harder to form the connection between the inputs (your commands) and the outputs (the robot's behavior), but it exists, and is deterministic. The same set of inputs will generate the same set of outputs every time. The problem, however, is that the robot will never see exactly the same input twice. No two camera images are ever the same, because of subtle lighting

---

[36] See M. Ryan Calo, Against Notice Skepticism in Privacy (And Elsewhere), 87 Notre Dame L. Rev. (Issue #2) (forthcoming Feb. 2012); M. Ryan Calo, People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship, 114 Penn St. L. Rev. 809, 849 (2010) (collecting such studies).

changes and measurement errors in the camera itself. Humans might not be able to see the differences, but the robot's software does.

The problem is that this different behavior in apparently similar situations can be interpreted as "free will" or agency on the part of the robot. While this mental agency is part of our definition of a robot, it is vital for us to remember what is *causing* this agency. Members of the general public might not know, or even care, but we must always keep it in mind when designing legislation. Failure to do so might lead us to design legislation based on the *form* of a robot, and not the *function*. This would be a grave mistake.

For example, if we fall into the trap of overly anthropomorphizing a human-shaped android, we might hold the designers less responsible for its actions than a more robotic robot. After all, it seems to have some limited form of free will, so how can we expect the designers to cover every eventuality? On the other hand, we hold car manufacturers to very high standards. If an automobile fails while on the highway due to a design oversight, it is the manufacturer's fault. A car is just a mechanism, and the designer should be able to predict what it will do in a given situation.[37]

Under these assumptions, if we are driving our car down the freeway and it fails to respond when we turn the steering wheel, it is unambiguously the manufacturer's fault. If an android is driving the car, and it's hands slip on the wheel while trying to make a turn, can we hold the robot-maker as accountable? Probably not. This means that the same outcome (the car leaving the freeway unexpectedly) is legislated differently, depending on who or what is driving the car.

This becomes problematic when we take the perception and reasoning technology in the android and embed it in the car itself, in a box under the hood. Now, since the technology is part of the car, it is legislated as a car. While it physically

---

[37] For example, we might be witnessing the first robotic wrongful death lawsuit right now, brought by a man whose daughter was allegedly killed by a negligent surgical robot. See http://robotland.blogspot.com/2012/04/did-da-vinci-robot-kill-24-year-old.html?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A

resides in the body of the android, it is legislated as an android. We have legislated the form, not the function. The same sensors and the same software generate the same result, but we have split it into two different cases from a legal perspective.[38]

Of course, this example is absurd. How could we be so easily misled? The android is clearly a machine, despite the anthropomorphic language we use to describe and think about it. We can expect people, even those with no technical background, to realize this and design legislation appropriately. Or can we? A recent study[39] has shown that people treat androids more like humans than machines. In an experiment where the android acted to deprive the human subject of a $20 reward, 65% of test subjects ascribed moral accountability to the android. This does not happen with vending machines that fail.[40] They are just machines that can fail. However, the android is something different and special, and is held to a different standard. While this experiment does not directly support our example above, it does show that humans are wired to anthropomorphize, ascribe, and project.

We must avoid the Android Fallacy. Robots, even sophisticated ones, are just machines. They will be no more than machines for the foreseeable future, and we should design our legislation accordingly. Falling into the trap of anthropomorphism will lead to contradictory situations, such as the one described above.

6. Complications: Deus Ex Machina

Figuring out how to think about and analogize robots is hard enough for systems that are clearly autonomous or

---

[38] There will be a small difference in the mechanical arrangement. The android will use its arms to turn the wheel, while the in-car system will probably use an electric motor. However, we claim that this difference is not relevant to our argument.

[39] "Do People Hold a Humanoid Robot Morally Accountable for the Harm It Causes?" Peter H. Kahn, et al. Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction, 2012.

[40] Or at least substantially not as often.

clearly tele-operated. Things get harder when we start to consider the new generation of shared autonomy systems. In these, a human operator (often at a remote location) collaborates with the autonomous software on the robot to control the system. The robot is neither fully autonomous nor fully teleoperated, and it will be difficult for an external observer to determine which mode (autonomous or remote-controlled) the system is in at any given time. This greatly complicates our choice of metaphors used to understand the system. We must also carefully choose the metaphors that we use to understand the operator's role, operating a system over which they have only partial control.[41]

Is the robot a portal or avatar for a remote expert (like a plumber), or is the human-robot system the "expert"? Where does liability lie if the human teleoperator issues the correct command, but the autonomous software on the robot carries it out poorly? What are the privacy implications of not really knowing if there is a remote human "inhabiting" your household robot? How can we provide effective privacy metaphors and safeguards for both the owner of the robot and the remote operator?

7. Conclusions and Final Thoughts

In this essay, we have advanced four basic claims about how the legal community should think about robots. Each of these claims is closely tied to the others, and we must consider all of them, and their interactions, if we are to design effective legislation and consumer protections for the coming generation of robots.

The first claim is that that we need to **think carefully about our definition of a robot**. While we are influenced by depictions of "traditional" robots in the popular media, this

---

[41] The remote operator is unlikely to directly control all of the joints of a sophisticated robot, because it is simply too hard to do so. Instead, they will give higher-level directions, such as selecting an object to grasp, and rely on lower-level autonomous software to carry out these commands. Thus, although that have good control over *what* the robot does, they have only loose control over *how* it does it.

definition is too narrow to be useful. Robots and robotic technology will creep into our lives in other forms, and our legislation must be uniform across these forms, and address the *function* of the system, rather than its *form*.

We also need to **understand the technical capabilities of current robots**, both in the world and in the laboratory. While most real robots fall far short of their fictional cousins, many research robots can do truly astonishing things and display a remarkable amount of intelligence. In order to create effective legislation, we must understand what robots are capable of, what they cannot do yet, and what they will never[42] be able to do. It is, of course, hard to say what is impossible, and we are forced to play a game of probabilities. However, a good working knowledge of the technology involved, and its limitations, allows us to make high-probability predictions. These predictions allow us to focus our (limited) effort on legislating for systems and problems that are more likely to occur in the coming years.

We should **draw on our considerable experience with cyber-law**, looking at how it drew analogies to existing technologies and legislation, where it succeeded, and where it failed. This will help to inform our choice of metaphor and analogy for robots and robotic technologies, along with the choice of regulatory tools where appropriate.

Finally, we should **avoid the Android Fallacy** at all costs. Not all robots are androids, and framing our analogies in highly anthropomorphized terms is dangerous. It will lead us into making false assumptions about the capabilities of robots, and to think of them as something more than the machines that they are, even if we try our best not to. This, in turn will lead us to use inappropriate analogies, and to design poor legislation.

The robots *are* coming, and they're coming soon. We need to be ready for them, and prepared to design appropriate, effective legislation and consumer protections for them. We

---

[42] Never is, of course, a long time. We adopt the pragmatic definition of "long after I'm dead". For example, robots will "never" (by our working definition) be able to read a human's thoughts without consent.

believe that we can only do this by understanding the technology, drawing on our recent experience with other disruptive technologies, and by avoiding seductive anthropomorphizations of our new metallic overlords.